Quality and Statistics:

Total Quality Management

Milton J. Kowalewski, Jr.
editor

ASTM STP 1209

STP 1209

Quality and Statistics: Total Quality Management

Milton J. Kowalewski, Jr., Editor

ASTM Publication Code Number (PCN):
04-012090-34

ASTM
1916 Race Street
Philadelphia, PA 19103
Printed in the U.S.A.

Library of Congress Cataloging-in-Publication Data

Quality and statistics: total quality management/Milton J.
Kowalewski, Jr., editor.
(STP 1209)
Includes bibliographical references and index.
ISBN 0-8031-1872-4
1. Quality control—Statistical methods. 2. Total quality management. 3. Process control—Statistical methods.
I. Kowalewski, Milton J., 1994– II. Series: ASTM special technical publication: 1209.
TS156.Q3Q292 1994
658.5'62—dc20 94-29563
CIP

Photocopy Rights

Peer Review Policy

Each paper published in this volume was evaluated by three peer reviewers. The authors addressed all of the reviewers' comments to the satisfaction of both the technical editor(s) and the ASTM Committee on Publications.

To make technical information available as quickly as possible, the peer-reviewed papers in this publication were printed "camera-ready" as submitted by the authors.

The quality of the papers in this publication reflects not only the obvious efforts of the authors and the technical editor(s), but also the work of these peer reviewers. The ASTM Committee on Publications acknowledges with appreciation their dedication and contribution to time and effort on behalf of ASTM.

Printed in Philadelphia, PA
September 1994

Foreword

This publication, *Quality and Statistics: Total Quality Management,* contains papers presented at the symposium of the same name held in Atlanta, GA on 4–5 May, 1993. The symposium was sponsored by ASTM Committee E-11 on Quality and Statistics. Milton J. Kowalewski, Jr., of E G & G Rocky Flats, Inc. in Golden, CO presided as symposium chairman and is also the editor of the resulting publication.

Contents

Overview

It was a warm and balmy time in Atlanta where 16 total quality individuals presented their dreams, messages, and stories to ASTM members and friends. Dorothy Savini and Scott Orthey provided a warm Atlanta welcome and southern style food in addition to their incomparable planning and coordination success. Nancy Trahey, the ASTM Chairman of the Board opened the E11 Symposium by awarding the ASTM Award of Merit and Fellow recognition to Dr. Ricardo Stone for his contributions and persistence. She was followed by E11 Chairman, Duncan McCune, who presented the Harold F. Dodge Award to Dr. Edward G. Schilling for reflecting the spirit and integrity of his late graduate advisor, Dr. Dodge.

These exciting award presentations set the stage for the 2 day marathon of presentations on Total Quality Management in standardization. Kumar-Misir opened the 16 paper series by describing his global-ready business model. His presentation was both informative and appropriate. Zott followed by telling the International Standards Organization (ISO) registration story as it relates to Underwriter's Laboratories (UL) in the United States and other accreditation bodies worldwide. Vardys set the foundation for the following 13 speakers by emphasizing the necessity to plan, control, and prepare documentation systems.

Locke, President of the American Association for Laboratory Accreditation (A2LA), opened the technical section of the morning by sharing his case-study of statistical measurement control for the evaluation of laboratory test results of automobile parts. Octogenarian, Daly, showed us the importance of formal design review, failure mode effects analysis (FMEA), and fault tree analysis (FTA) procedures. He was followed by Tulay who explained the Energy Program Research Institute (EPRI) guide for sampling procured items for the nuclear industry.

Since Mandel was visiting his homeland overseas, McCune aptly read and interpreted Dr. Mandel's paper on the ASTM Guide for Interlaboratory Studies (E 691) and its computer software capabilities. An additional step for enhancing the application of the standard was proposed for use. Lindow closed day one with examples of construction projects where Total Quality Control (TQC) was not used and should have been used.

Day 2 was opened by McCune and Levine presenting the do's and don'ts of control charts. They were followed by Moyer who shared the frustrations and need for quality standards in the manufacture of magnetic products that were identified as functional throughout our homes and automobiles. His interlaboratory study revealed startling results. He was followed by Schilling's transitional paper and easy to apply ABC plan for initiating, establishing, and maintaining both capability and control of processes.

Farrar and yours truly, the symposium chairman, lightened up the morning in preparation for lunch with a philosophical and practical combination of organizational frames and TQM in a road design agency. Bernstein helped us digest lunch by his after meal tour through the "belly" of the chemistry lab at Bridgeport Hospital in Massachusetts. Winding down the last quarter of the marathon of papers on day 2 was Yeung speaking on the results of surveys and quality assurance systems designed for small clothing establishments in Hong Kong where the expectation of the return to Chinese rule in 1997 has retarded the development of quality systems.

Lau spoke about the advantages of continuous process flow analyzers in the Canadian oil industry, and Ping paved the way for us to understand a research project conducted on

Texas highways.

My personal thanks are due the ASTM staff, the speakers, each of the authors, the superb service of the Hyatt Regency Hotel, the ASTM manager of Acquisitions and Review, Kathy Dernoga, for this special collection of leading edge information in the sciences of quality, statistics, and Total Quality Management.

Milton J. Kowalewski, Jr.

E. G. & G. Rocky Flats, Inc.,
Golden, CO;
Symposium chairman and editor.

Leslie M. Kumar-Misir (1)

INTERNATIONAL COMPETITIVENESS AND BUSINESS EXCELLENCE

REFERENCE: Kumar-Misir, Leslie M. **"International Competitiveness and Business Excellence"**, Quality and Statistics: Total Quality Management, ASTM STP 1209, Milton J. Kowalewski, Jr., Ed., American Society for Testing and Materials, 1994.

ABSTRACT: The road to international competitiveness and business excellence is paved with good intentions. Nevertheless, its elusive nature and mystery can be peeled away, firstly, by a fuller understanding of certain unrelenting economic forces that govern growth and development and the realities of watershed technology, and, secondly, by harnessing the power of the pillars of international competitiveness, viz., globalization of the enterprise, globalization of markets and marketing, the pursuit of superior processing and the practice of business excellence.

KEY WORDS: Unbalanced growth, watershed technology, global company, global marketing, external minima, processing advantage, CQI, Triple-CIM, concurrent engineering.

SEARCHING FOR THE WAY

In our rapidly changing global output and trade environment, the international competitiveness and business excellence of world industry is the key to the continued growth and prosperity of the group of More Developed

---

(1) Industry Analyst, Textiles, Apparel and Leather Directorate, Consumer Products Branch, Industry Canada, Ottawa, Ontario K1A 0H5, Canada.

Countries (MDCs), and at the same time , the gateway to economic growth, economic development, economic progress and never-ending prosperity for the Lesser Developed Countries (LDCs). Promoting significantly improved performance by a country's commercial-industrial complex in the global marketplace must be a central element of the mandates of governments, industry et al.

Everyone everywhere in the world is interested in understanding and achieving international competitiveness and business excellence. To meet this worldwide interest, a veritable deluge of information has been generated, appearing in text books, articles in learned journals, etc. But much remains to be clarified and made simple.

Here are some examples of this great interest and its various interpretations. In Canada,

> "*To take advantage of the immense opportunities opening up in the new global market-place, Canadians must improve their ability to compete. The "Action Plan for Canada's Prosperity" makes clear that the keys to achieving this are innovation and quality management practices combined with technological mastery and growing productivity*", (Inventing Our Future: An Action Plan for Canada's Prosperity, 9-60).

In the United States of America,

> "*Globalization is the effective deployment and utilization of worldwide resources, integrated with opportunities, to achieve competitive advantage and superior business results*", (Globalization: The Forces Behind, 5).

In Mexico,

> "*Observers point to four strategies to improve the prospects for long-term competitiveness and market power of the Mexican textile industry: cost reduction, foreign investment, strategic/ technological improvements, and expanded demand*", (Competitiveness of the Mexican Textile Chain, 40).

TYRANNY OF THE STATUS QUO

All situations default (or seem to default) to the status quo. Indeed, the status quo is a most loyal companion and, its tyranny is acknowledged repeatedly with the common observation: "...but we've always done it this way!"

How then does this gentle tyrant exercise its rule over industry?

Canada's great manufacturing industries are rich with examples of sectors remaining home-market bound. Indeed, it is frustrating to note that, despite Canada's long manufacturing history, currently two-thirds of Canadian companies do not sell their products outside Canada.

Canada's spunky textile industry is but one case. From rather humble beginnings, namely, the start-up of a small cotton yarn and fabric mill in Sherbrooke in 1844 and another in Montreal in 1853, the Canadian textile industry today is found across the land, busily producing an impressive range of textiles and textile products for domestic consumption and export markets. Currently, the industry comprises some 1 065 firms, employs around 51 000 workers and annually ships some $5.7 billion, of which a mere 19.3 percent is exported.

Table 1 summarizes the default condition. While external market shares generally have risen in the latest two year period shown below, except for floor tile, linoleum and coated fabrics, and manmade fibres and yarns, to date, the external market shares of this long-established industry are far from what are desirable for Canada to be considered a textile exporting country, i.e., a country with a global textile industry.

TABLE 1 -- NOT REALLY A GLOBAL INDUSTRY, TEXTILE MARKET SHARES BY SECTORS, CANADA, 1988 TO 1990

	External Markets		Domestic Markets	
	1988	1990	1988	1990
	-%-			
Floor Tile, Linoleum and Coated Fabrics	45.8	67.3	54.2	32.7
Manmade Fibres and Yarn	28.1	42.1	71.9	57.9
Miscellaneous Textiles and Tire Cord	15.5	23.0	84.5	77.0
Natural Fibre Processing and Felt Products	34.5	21.8	65.5	78.2
Wool Yarn and Woven Fabrics	12.8	19.3	87.2	80.7
Other Spun Yarn & Woven Fabrics	11.5	15.0	88.5	85.0
Narrow Fabrics	5.3	10.6	94.7	89.4
Carpet, Mats and Rugs	8.3	10.1	91.7	89.9
Textile Hygiene Products	3.1	4.5	96.9	95.5
Broadknitted Fabrics	1.6	4.4	98.4	95.6
Textile Home furnishings	17.9	3.2	82.1	96.8
Canvas and Related Products	2.0	1.7	98.0	98.3
Contract Dyeing and Finishing	0	0	100.0	100.0

SOURCE: Textiles and Leather Directorate, Consumer Products Branch, ISTC.

PRECONDITIONS FOR INTERNATIONAL COMPETITIVENESS

Preconditions for achieving goals, business and otherwise, abound. Two preconditions are especially relevant to the drive to international competitiveness and business excellence of nations. They are: (a) unbalanced growth and (b) technological watersheds.

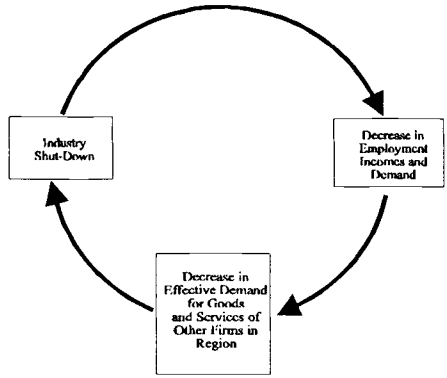
Unbalanced Growth versus Balanced Growth

Put simply, the experience of nations culminating in the rise of a smaller number of industrialized countries argue in favour of a theory of unbalanced growth, when analyzing the break-up of nations into more developed countries (MDCs) and less developed countries (LDCs), and, in turn, that of an MDC into rapid-growth regions and slow-growth regions.

To explain, unbalanced growth theory states that economic growth stems from the interplay of three forces: a) cumulative causation defined as a "circular constellation of forces tending to act and react upon one another in such a way as to keep a poor country in a state of poverty", b) back-wash effects defined as "all relevant adverse changes, caused outside that locality, of economic expansion in a locality", and c) spread effects defined as "certain centrifugal "spread effects" of expansionary momentum from the centres of economic expansion to other regions", (Economic Theory and Underdeveloped Regions, 11, 30-31). Indeed, economic change is not normally accompanied by countervailing changes but, instead, by supporting changes which not only move the system in the same direction as the first change but much further, (Figures 1 and 2).

FIGURE 1

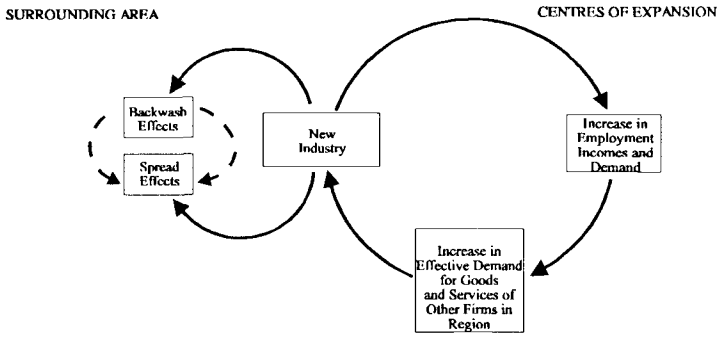
THE DECLINING REGION IN THE DRIFT TOWARDS REGIONAL ECONOMIC INEQUALITIES



SOURCE: Based on Gunnar Myrdal, Economic Theory and Under-Developed Regions (London: Gerlad Duckworth and Co. Ltd., 1957)

FIGURE 2

REGIONAL CENTRES OF ECONOMIC EXPANSION IN THE DRIFT TOWARDS REGIONAL ECONOMIC INEQUALITIES



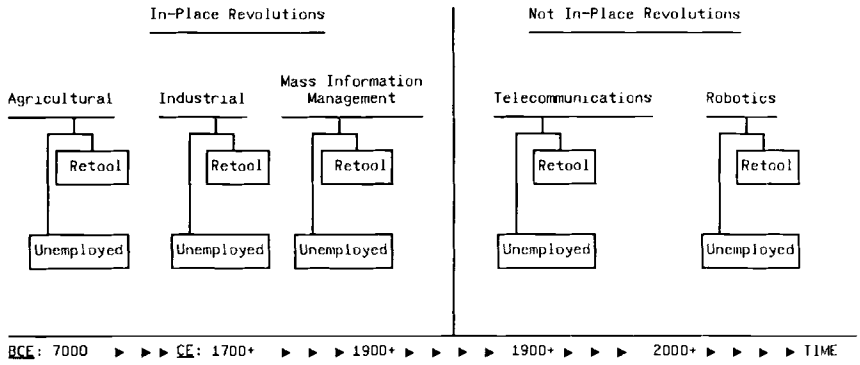
SOURCE: Based on Gunnar Myrdal, Economic Theory and Under-Developed Regions (London: Gerlad Duckworth and Co. Ltd., 1957)

TECHNOLOGICAL WATERSHEDS

From the dawn of civilization to the present day, technological watersheds continue to occur bringing decline certainly but more importantly, giving rise to the conditions for nations to achieve sustained international competitiveness and business excellence, (Figure 3).

FIGURE 3

TECHNOLOGICAL WATERSHEDS, 7000 BCE TO PRESENT AND BEYOND



BCE-Before the Christian era; CE-Christian Era
SOURCE:Author

Figure 3 suggests that socio-economic progress and the satisfaction of ever-changing human wants and desires are maximized when quantum leap-type technology revolutions, past and imminent, are recognized and incorporated fully into the business culture, i.e., when the retool phenomenon path is followed. Alternatively, when technology revolutions are ignored or served less than fully, the phenomenon of the "pools of unemployed" will prevail. Indeed, international competitiveness and business excellence may forever remain just beyond one's grasp.

INTERNATIONAL COMPETITIVENESS AND BUSINESS EXCELLENCE

The shift to a North American market, accelerated, on the one hand, by the U.S.-Canada Free Trade Agreement of 1989 and the actions by the United States, Canada and Mexico to implement a North American Free Trade Agreement (NAFTA), and the unrelenting globalization of enterprise and trade, on the other, demand that concerned countries everywhere complete effective reconstruction and development towards achieving sustained international competitiveness (IC) and business excellence (BE)

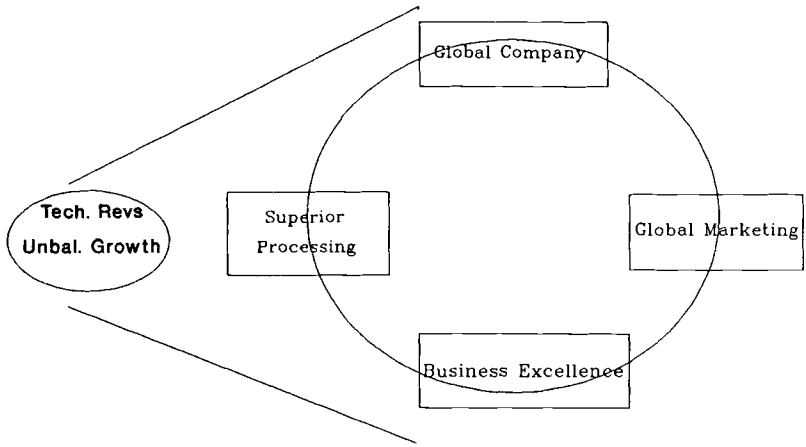
What is this thing called international competitiveness? What and how many are the pillars of international competitiveness? Do they converge to form a practical model of international competitiveness?

International competitiveness may be defined as the power of a global company to undertake sustained production and/or marketing of customer-quality driven products in one or more of the four smaller markets of the single global market, viz., local, regional, national and external, in which competing products, perfect or otherwise, may also be offered , or to be global market-ready, i.e., ready to enter on a sustained basis any smaller market of the single global market anywhere at any time.

Recalling the earlier overview of unbalanced growth and technological watersheds, as is the case with the rise of rapid-growth regions, a circular constellation of forces tend to act and react upon one another in such a way as to bring international competitiveness and business excellence to the responsive enterprise. These irresistible forces separate to yield four pillars of international competitiveness, namely, the global company, global marketing, superior processing and business excellence. Their workings are circular, interactive, interdependent and

convergent. Schematically,

FIGURE 4
THE FOUR PILLARS OF INTERNATIONAL COMPETITIVENESS



SOURCE: Author

The four pillars of international competitiveness are explained below.

Global Company

A global company is neither a multi-national enterprise (MNE) nor a subsidiary as currently (and historically) recognized. Rather, a global company is a newly evolved business enterprise (NEB) that:

1) produces and/or distributes an economic good or service which is in demand worldwide,

2) is engaged in sustained international marketing,

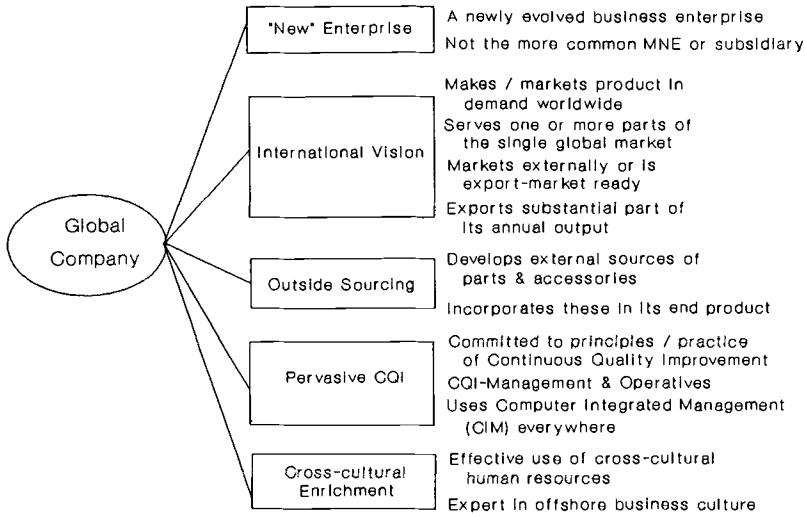
3) exports a substantial part of its annual output,

4) in its final product, increasingly incorporates parts and accessories manufactured elsewhere,

5) has a life-long commitment to the principles and practice of Continuous Quality Improvement (CQI) of management and operatives, and

6) in its corporate culture, organization and business operations, emphasizes the cross-cultural content and contributions of its diverse personnel complexion.

Figure 5--Main Elements of the Global Company in the Drive to International Competitiveness



SOURCE: Author

Global Marketing

Turning to Global Marketing, its main elements are:

1) product availability, accessibility and affordability over time and at any time,

2) unmatched quality of service to customers worldwide,

3) reciprocal marketing arrangements,

4) present/convince/sell techniques based on:
a) the well-known four P's (right product, right place, right promotion, right price), and
b) global marketing systems design answering accurately questions such as:
- Is there a market for the product?
  - If yes, who and where are the buyers?, etc.,

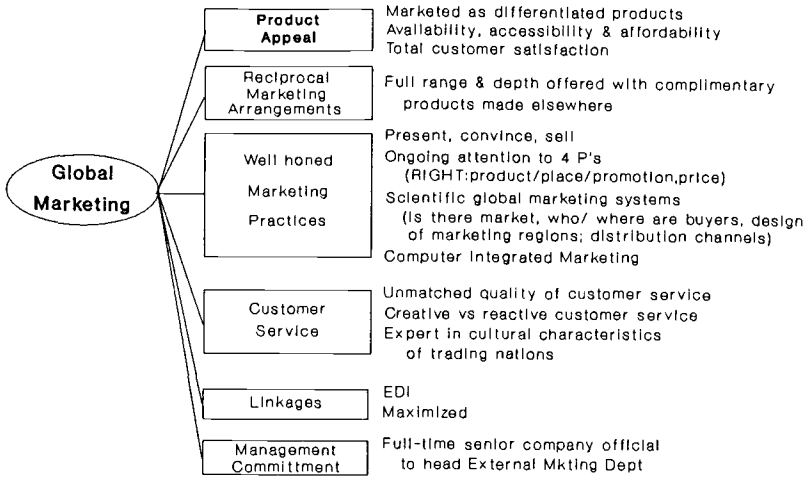
and:
- drawing up marketing regions within countries having both demand and purchasing power,
- determining "best" distribution channels.

5) maximization of computer linkages, and

6) senior management committment to external trade.

Schematically,

FIGURE 6--Main Elements of Global Marketing In the Drive to International Competitiveness



SOURCE: Author

Superior Processing

Turning to superior processing, its main elements are six:

1) a clear and definite product manufacturing mandate.

2) continuous modernization, i.e., from stuck-in-time technology (SIT) to state-of-the-art technology (SAT),

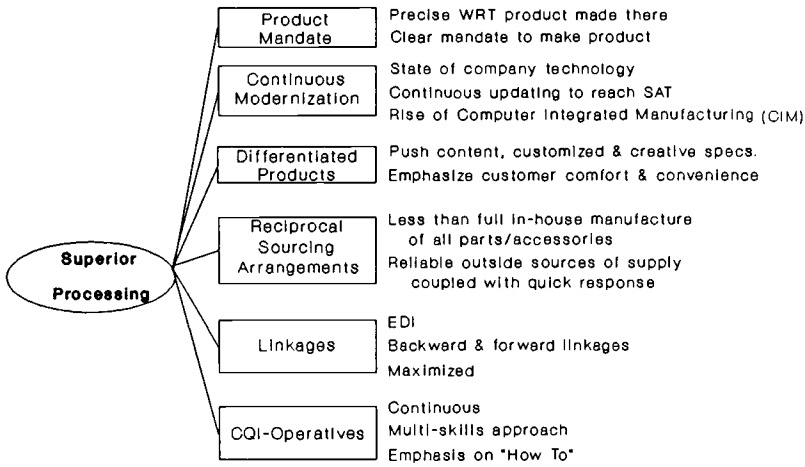
3) product differentiation by way of manipulating the full range of product specifications (content, customized, creative , customer comfort and convenience specifications), and right-the-first-time and every-time quality manufacturing,

4) reciprocal parts and accessories sourcing arrangements,

5) maximization of computer linkages, and

6) Continuous Quality Improvement of operatives, (CQI-Operatives), (Figure 7).

FIGURE 7--Main Elements of Superior Processing in the Drive to International Competitiveness



SOURCE: Author

Business Excellence

Finally, turning to Business Excellence, its main elements are four:

1) people excellence reinforced by a pervasive Continuous Quality Improvement programme designed to keep

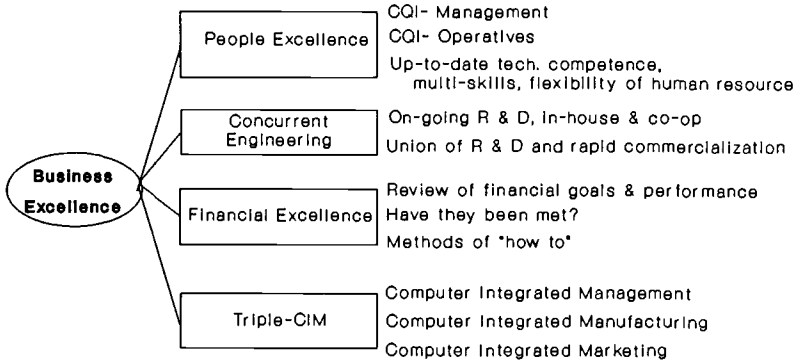
management and operatives efficient, effective, second-to-none technologically, multi-skilled and flexible,

2) concurrent engineering, i.e., Research and Development linked (R&D) to rapid commercialization,

3) financial excellence from never-ending attention to the commonly used financial ratios (liquidity, leverage, activity and profitability) [1], and

4) triple-CIM-ready, e.g., some enterprises may not have an immediate need for Computer Integrated Manufacturing, (Figure 8).

FIGURE 8-- Main Elements of Business Excellence in the Drive to International Competitiveness



SOURCE: Author

[1]Liquidity: current ratio, acid test: Leverage: debt to total assets,times interest earned, fixed charges coverage; Activity: average collection period, inventory turnover, fixed assets turnover; Profitability: profit margin on sales, return on total assets (ROI), return on equity.

CONCLUSION

> *"Changes can come from the power of many when the many come together to form that which is invincible- the power of one."*

International competitiveness and business excellence constitute such a power. Its power of many coming together to form one is unleashed by the convergence and interactive workings of its four pillars. Prepare it room, then, by paying mind to its underpinnings (growth theory, looming watershed technologies) and undertaking concerted action on its power of one:

a) pay proper mind to the contribution of unbalanced growth to the performance of the enterprise ,

b) take the enterprise boldly into watershed technologies, in place and imminent, viz., voice interactive micro-chip telecommunications and robotics,

c) customize the generic model of international competitiveness to the enterprise,

d) cultivate among management and operatives a passion for exellence,

e) globalize the entreprise, its mission, culture and organization,

f) globalize the market and marketing functions of the enterprise,

g) move increasingly to the substitution of superior processing,

h) let the content and pursuit of business excellence take the enterprise into the twenty-first century.

To strive, to seek, to find. Not to yield but to go boldly where no industrialist has gone before.

Yes, you can!

REFERENCES

(1) Steering Group on Prosperity. Inventing Our Future: An Action Plan for Canada's Prosperity. Industry, Science and Technology Canada, Ottawa, Ontario, Canada, Oct. 1992.

(2) Snyder, John W. "Globalization - The Forces Behind". Global Business Support - International Development, Du Pont. Paper presented at the INDA-TEC '90 Conference, INDA-Association of the Nonwoven Fabrics Industry, Baltimore, MD., U.S.A., June 5-8, 1990.

(3) Mandelbaum, Judah."Competitiveness of the Mexican Textile Chain". Textile Outlook International, No. 44, (November 1992). London, U.K., The Economist Intelligence Unit, 1992.

(4) Textiles and Leather Directorate. Textile Industry Profile. Industry, Science and Technology Canada, Ottawa, Ontario, Canada, 1992.

(5) Multifibre Agreements Directorate. Textile Industry Data Book, Industry,Science and Technology Canada, Ottawa, Ontario, Canada, 1992.

(6) Perroux, F. "Note sur la Notion des Pôles de Croissance." Economie Appliqueé, 8 (jan-juin 1953).

(7) Hirschman, Albert O. The Strategy of Economic Development. New Haven, CT.: Yale University Press, 1958.

(8) Myrdal, Gunnar. The Challenge of World Poverty, New York: Pantheon Books, 1970.

(9) Myrdal, Gunnar, Ed. Economic Theory and Underdeveloped Regions, London: Gerlad Duckworth and Co. Ltd., 1957.

(10) Kaldor, Nicholas. "The Case for Regional Policies." Scottish Journal of Political Economy, Vol. XVII.

(11) Richardson, Harry W. Elements of Regional Economics, New York: Penguin, 1970.

(12) Richardson, Harry W, Ed. Regional Economics: A reader. New York: St. Martins's Press, 1970.

(13) Richardson, Harry W, Regional Economics: Location Theory, Urban Structure and Regional Change. New York: Praeger, 1969.

(14) Richardson, Harry W, "A Critique of Regional Growth Theory." Paper given at the Regional Economic Development Conference, University of Ottawa, Ottawa, March 2-4, 1972, (mimeographed).

Ruta K. Vardys[1]

STANDARDIZATION OF POLICIES AND PROCEDURES AT COMMERCIAL TESTING & ENGINEERING CO.

REFERENCE: Vardys, R. K., "Standardization of Policies and Procedures at Commercial Testing and Engineering Co.," Quality and Statistics: Total Quality Management, ASTM STP 1209, Milton J. Kowalewski, Jr., Ed., American Society for Testing and Materials, Philadelphia, PA 1994.

ABSTRACT: Pressure from world-wide competitors, stricter government regulations, conformance to quality system standards such as ISO 9000, and other market conditions are requiring companies to look for better ways to address the quality of products and services and become more flexible and efficient in satisfying their customers' needs. U.S. companies are introducing or perhaps re-introducing Total Quality Management (TQM) concepts to respond to these changing conditions in the market environment. A fundamental element required for the successful implementation of a TQM system is formal documentation. Documentation is the corner stone of all quality systems. Without documentation, a quality system really does not exist. This paper will discuss how documentation and TQM are interwoven concepts, the immediate benefits of documentation, and what is involved in documenting an organization's quality system.

KEYWORDS: Total Quality Management (TQM), documentation, document control, quality management system, quality system, quality records

[1]Technical Auditor, Quality Management Services Group, Commercial Testing & Engineering Co., 1919 South Highland Avenue, Suite 210-B, Lombard, Illinois, 60148.

TOTAL QUALITY MANAGEMENT (TQM) IS BEING MADE A PRIORITY

Many U.S. companies have been struggling to survive under a constantly changing market climate caused by a number of factors. For example, for many U.S. industries, increased competition has been caused by the increase of new low-price producers from external sources. Many organizations are affected by new environmental regulations such as the Clean Air Act. Others are coming under pressure to conform to quality system requirements such as ISO[2] 9000.

Commercial Testing & Engineering Co. along with many U.S. companies, is deploying Total Quality Management (TQM) concepts to become more flexible in meeting the needs of the constantly changing market climates in the world today.

One key requirement for the successful implementation of TQM concepts is to formally document an organization's entire management system. Documentation defines, through an objective medium, the concepts that must be followed and activities that must be carried out to achieve the desired level of total quality in an organization. Documentation is the objective evidence that a quality system exists.

COMMERCIAL TESTING & ENGINEERING CO. (CT&E)

Commercial Testing & Engineering Co. (CT&E) is a member of the Société Générale de Surveillance S.A. Group (SGS), the largest control and inspection organization in the world. CT&E, with over 850 employees, operates over 50 laboratories and 35 sampling operations throughout the United States.

CT&E laboratories perform routine chemical and physical analyses used to evaluate coal, minerals, and other bulk commodities. CT&E laboratories also perform organic and inorganic environmental analyses as well as quality verification services.

Whatever type of service CT&E provides:

- The nature is PROTECTION for the client; and
- The function, which is INDEPENDENT verification and assessment.

CT&E is in the process of formally deploying Total Quality Management (TQM) concepts throughout the organization through training, auditing, and documentation. Although

---

[2] International Organization for Standardization

there is no specific way to deploy TQM, documentation is one area that must be addressed. There are several key areas to consider when developing documentation such as the organization of documents, document control, and handling of records.

TOTAL QUALITY MANAGEMENT AND DOCUMENTATION: INTERWOVEN CONCEPTS

Total Quality Management (TQM) is a set of management concepts that any organization should adopt at every management level. The ultimate goal of TQM deployment is producing quality products and/or services efficiently and economically that continue to self-improve and consistently meet/exceed client requirements/expectations.

The following outlines the TQM concepts that management must drive so that quality concepts and the quality system are effective:

- An operating philosophy of total commitment to quality in everything the organization does;
- Involvement of everyone and everywhere both as team members and as individuals;
- Consideration of everything the company does as a process;
- Application of data and scientific methods to measure and improve the organization's processes;
- Seeking perfection (always meeting/exceeding client expectations/requirements);
- Focusing on continuous process improvement;
- Seeking perfection in always meeting/exceeding internal client expectations/requirements;
- Verifying that goals and performance requirements have been met.

Documentation is directly related to each of the concepts outlined above and is also the medium that management must use to deploy these concepts in the organization (see Figure No. 1.0: Relationship between Documentation and Total Quality Management.) For example, documentation provides the format for communicating the goals or standards to which an organization is committed. Documentation provides the basis upon which an organization may make comparisons to verify that their goals have been met, etc.

TQM and Documentation: Interwoven Concepts	
Concepts to be Driven by Management:	Documentation provides:

*	Total commitment	⇒	Format to state the goals and standards to which the organization is committed
*	Total involvement		
*	Everything is a process	⇒	Format to define processes and for implementation
*	Application of data and scientific methods to measure and improve	⇒	Record of data measurements and methods for review for improvements
*	A focus on total client satisfaction	⇒	Definition and communication of client needs and expectations
*	Total satisfaction for internal as well as external clients		
*	Continuous process improvement	⇒	Objective evidence for comparison to a standard or benchmark to judge improvement
		⇒	Record of past performance, methodology, and future improvement
*	Verification that goals have been met	⇒	Objective evidence for verification of meeting goals or requirements

Figure No. 1.0: Relationship Between TQM Concepts and Documentation

STANDARDS, REGULATIONS, AND CERTIFICATIONS

ISO 9000, A2LA[3], the EPA[4]'s Clean Air and Clean Water Acts, Hazardous Emissions, and state ELAP[5]'s are all examples of the standards, regulations, and certification

---

[3] American Association for Laboratory Accreditation

[4] Environmental Protection Agency

[5] Environmental Laboratory Accreditation Programs

programs where compliance is required to conduct business or to be able to compete in related industries. Upon comparison of these programs, there are many similarities.

All of these programs require definition of management responsibilities, the identification of training needs and verification that training has taken place, verification of the satisfactory performance of equipment, and some type of identification, record, and action upon non-conformances so that they do not reoccur. In other words, each program resembles a quality system in one form or another.

One other common thread throughout these programs or quality systems, is documentation. Words such as document, define, data, procedures, record, and instructions are mentioned throughout the requirements outlined in the standard, certification, or regulation.

Documentation also serves as the "audit trail" or objective evidence an assessor uses to verify compliance to a standard, regulation, or certification program. For example, the assessor looks at records, which is one form of documentation, to verify that calibration of instrumentation has taken place according to schedule.

THE BENEFITS OF DOCUMENTATION

How many times does your client audit your facility? How often are employees relying on management to solve simple problems and trouble-shoot? How many times do you depend on "old weird George" for advice? How do you know if your organization's performance has improved? These are some of the issues where documentation is integral to the operation of your organization and will provide many benefits.

A documented quality system is objective evidence, not just anecdotal evidence, of the existence of policies and procedures describing the actions employees and organizations follow to accomplish their goals. Because they are documented, they can be reviewed by anyone wishing to audit the system -- internally or externally. Such documented procedures serve as the standards against which an individual's or organization's performance can be measured.

In turn, these documented policies and procedures can be measured against industrially-accepted or industrially-required quality system models to see if all pertinent quality system elements are present and appropriately controlled. Documented procedures can also be checked

against technical procedures (ASTM[6], ANSI[7], ISO, ASME[8], SAE[9], API[10], etc.) to be assured of technical competence. Such auditable documents also assist in providing a map or audit trail of the organization and related processes to make auditing a system easier and more efficient.

Documentation also provides uniformity and consistency assisting in communicating policies and procedures throughout an organization. For example, CT&E has recently expanded, through acquisitions, into different testing industries, including environmental. Often times when organizations are incorporated, purchased, or restructured there becomes the challenge to standardize policies and procedures to better communicate the standards, objectives, and expectations of the new organization.

A "functional expert" ("old weird George") is someone who usually has been with the organization for a number of years and is (or appears to be) highly knowledgeable of company operations. "Functional experts" are sought by other employees for their expertise because, usually, there is no other resource available. Many times, the qualifications of the "functional expert" are questionable and, should the expert leave the organization, valuable operating information is lost. Documentation eliminates the risky reliance on "functional experts."

Documentation helps to install confidence in employees which helps them make their own decisions and find their own answers. By having effective documentation, management has the freedom to perform administrative management functions. Management confidence is increased when personnel have the tools to make their own decisions in operations. Client confidence increases when they can see a documented system in place that assures that their supplier produces a consistent quality product or service. Clients do not feel the need to monitor sub-contractor or vendor operations as closely. For some, registration to an ISO standard is assurance enough.

Documentation is the Cornerstone

---

[6] American Society for Testing and Materials

[7] American National Standards Institute

[8] American Society of Mechanical Engineers

[9] Society of Automotive Engineers

[10] American Petroleum Institute

Though there are many benefits to documentation, the bottom line is that a quality system does not exist without it. Documentation can be compared to the cornerstone of an establishment. The cornerstone of an establishment identifies when the building was built, by whom, and the building name. Documentation provides the identification of a quality system describing where the system came from, who is responsible for its maintenance, and describes exactly what the system is.

Documentation Hindrances

There are no disadvantages in the documentation itself, but the process used to develop documentation can be destructive if not approached properly. As stated in one of the TQM concepts, "everyone, everywhere, must be involved both as team members and individuals" and there must be "an operating philosophy of total commitment to quality in everything the company does." When employees sense that there is a lack of commitment and have seen examples of a lack of total commitment to quality in daily operations, it is tough to convince those employees and their management that documenting the entire system is necessary. The commitment to document the quality system may be demonstrated through, for example, the commitment of resources to complete the job properly.

Whatever hindrances there may be to the documentation process, documenting the quality system will forever save time for existing and future processes forever after.

DOCUMENTATION DEVELOPMENT

Historically, CT&E has documented, in the form of procedures, mostly the technical functions of the organization. Documentation that reflected the entire quality system was either incomplete or did not reflect the entire system.

The following briefly outlines the steps that have been taken to develop effective quality system documentation at CT&E:

- Determine status of current documentation;
- Develop a documentation plan, coordinating across different functions; Assign responsibilities;
- Develop document organization (hierarchy);
- Establish standardized procedure for developing CT&E documentation; Develop uniform format for all CT&E documentation; Establish document coordinator;
- Develop the policies and objectives of the quality system;
- Develop the supporting documentation for the policies with procedures and work instructions.

Status of Current CT&E Documentation:

Having identified the requirements established by the ISO 9000 standards and the standards set by corporate objectives to deploy TQM concepts, it was necessary to evaluate documentation currently available. This documentation included policies, procedures, manuals, work instructions, video tapes, and any other media that was or is being used to demonstrate methodology.

The documentation was measured against the following criteria:

- Is the quality system adequately documented?
- Do authorized personnel approve documentation before implementation?
- Are documents controlled (such that we are able to determine that the necessary personnel have the latest document available for their use)?
- Which documents have business value?
- Are documents produced in an efficient manner?
- Are the appearance and organization of the documentation consistent and easy to understand and use?
- Does the documentation in its current state effectively communicate CT&E policies and procedures?

By reviewing current documentation before producing new documentation, it was possible to determine what areas have already been satisfactorily completed to avoid duplication.

Developing the documentation plan

It is essential to develop a plan for producing documents efficiently and producing documents that would best communicate organizational policies, responsibilities, requirements, and activities. Based upon the initial review, a plan was developed to complete the necessary documentation. This included assigning responsibilities for documenting each area and holding regular status meetings to measure progress.

Document Organization

In order to organize the large amount of information that was to be documented, the documentation was structured into a four-tier hierarchy. Each tier of documentation would draw its core substance from its related document in the next higher tier. Each layer of documentation would progress from more general, as in policies, to more specific, such as work instructions (see Figure No. 2: CT&E Document Hierarchy.)

The highest document in the hierarchy is a quality policy

that briefly describes the beliefs and concepts that we aspire to as a company.

CT&E DOCUMENT HIERARCHY:
(LEVELS I - IV)

QUALITY POLICY
↑
POLICIES AND OBJECTIVES
(QUALITY MANUAL)
↑
PROCEDURES
↑
WORK INSTRUCTIONS

Figure No. 2.0:
CT&E Documentation Hierarchy

The second highest tier of documents in the hierarchy are the policies or objectives of the company. The policies will give a brief description of the company's requirements, objectives, and standards for the different elements of our quality system. These policies are generally compiled into a manual called the quality system manual.

The third tier of documentation are procedures. Procedures are written based upon the content of the relative policy and outline who, how, when, where, and at what standard an activity should be performed without location-, equipment-, or client- specific details.

The last tier of documentation are the work instructions. Work instructions provide details and supplement a relative procedure(s) or policy(-ies). Work instructions are also location-, client-, or equipment-specific or specific to one position. There may also be supplemental work instructions that provide temporary or alternate guidelines from routine procedures.

Using this four-tiered organizational approach, CT&E was able to define the structure of its documentation system.

Quality Records

To verify that all of the activities stated within the documentation have taken place, a large volume of records were to be maintained. Under the documentation structure, quality records fall under the fourth tier of documentation. A procedure was established on how to file, retrieve, and index all quality records.

Developing and Controlling Documents

Simultaneous with establishing a document structure for CT&E documents, a standardized procedure for developing and controlling documents was produced. This procedure established guidelines and indicated how to write a policy or procedure. A review process was established to ensure that each document is technically correct, to check for readability, and to allow input from more than one source. The review process developed is similar to the process used to develop ASTM standards. Policies and procedures are established by consensus among a review team. Should one person disagree with the content of a procedure, the originator would have to resolve the disagreement before the procedure could receive final approval and be implemented. When a revision or cancellation is required for a document, the same process would be used to review and approve the revision or cancellation.

Controlling Documents

Equally important to developing a method for reviewing and approving documents, a method was developed to control documents. The document control process assures that the documents that affect the quality of the product or service are distributed to the people that implement them and must be aware of them. This process also assures that these people have received the most current revision of the document and that superseded and cancelled documents are removed from use.

The document control process includes establishing a uniform format for all documents, identifying each document produced with a unique number, numbering copies of controlled documents, recording the number and recipient of each controlled document, and using document receipt forms to verify that the controlled copy was delivered as desired. Controlled documents are also printed on water-marked stationary for copying security.

Establishing a Document Coordinator

As a part of the documentation plan, a "document coordinator" was established at each location. This person's responsibilities would include maintaining a

master list of all documents and their status, ensuring all documents underwent the established review process, and controlling all documents. This person would be the centralized location for satisfying document needs.

Development of Policies and Objectives

Once a uniform document format and development process were instituted, the quality system proceeded to be documented in a structured and organized manner. The policies or objectives of the quality system were defined first. The supporting documentation was developed based upon the content of the policies and objectives.

These policies and objectives are based upon the TQM system concepts previously outlined. CT&E used the ISO 9002 quality system standard as a guideline to determine which areas needed to be described by a policy. These policies were then organized and placed collectively within a manual that would be called the CT&E Corporate Quality System Manual (CQSM.)

Development of Supporting Documentation

Using the company policies, supporting documentation was be developed. This supporting documentation describes how, when, where, and by whom the policies are to be carried out and at what standard. Generally, this supporting documentation is best described in a written procedure and, when necessary, a complementing work instruction. The application of these documents are being used to develop job descriptions, implement training programs, explain auditing practices, and many other quality-related activities.

Involvement of all Employees

The commitment of management to Total Quality Management concepts will determine whether the documentation will be successfully developed and effectively implemented or not. However, without the initial involvement of all employees, the documentation will not accurately describe the organization's quality system and will bring considerable expense in training and implementation to the organization.

In order to have effectively documented the quality system, the documentation must accurately reflect what the quality system does. The common phrase regarding this is "Document what you do, and do what you document." The employees know the operation first-hand and are able to provide the insight required to accurately document the processes of the organization.

Once documentation is developed, especially if the documentation describes a new or revised process or program, all employees must be made aware of the changes and new processes. This requires training and a period of time for implementation until the new or revised procedure is working efficiently. These costs can be considerably reduced by initially involving all employees in the documentation process.

An invaluable part of employee involvement is the awareness the employees gain regarding the organization's quality system. The employees will increase their understanding of the quality system. The employees will be more apt to maintain the system and be able to revise the system to better satisfy and exceed client needs and expectations. Involvement also provides ownership of the quality system by the employees and all of the associated benefits.

SUMMARY

Many U.S. organizations are implementing Total Quality Management (TQM) concepts to better adapt to an ever-changing market climate. An integral part of the TQM implementation process involves documentation.

Documentation can be an immense effort on the part of organizations to develop and implement. However, CT&E has found that by establishing a hierarchial documentation structure, establishing a standardized method for review, approval, and control of documents, and a document coordinator that the process is much more efficient.

Ultimately, the success and effectiveness of your documentation will depend upon the involvement of all employees in its development. Initial involvement of all employees saves costs in training and implementation of the documentation. It also gives employees an invaluable amount of awareness regarding the operation of the quality system.

REFERENCES

[1] "ISO 9000 Series, Quality Management Systems," a presentation by SGS Yarsley Quality Assured Firms and Commercial Testing & Engineering Co., April 30, 1992, Birmingham, AL.

[2] Gitlow, H., Gitlow, S., Oppenheim, A., Oppenheim, R., Tools and Methods for the Improvement of Quality, Richard D. Irwin, Inc., Homewood, IL, 1989.

John W. Locke[1]

STATISTICAL MEASUREMENT CONTROL

REFERENCE: Locke, John W., **"Statistical Measurement Control"**, Quality and Statistics: Total Quality Management, ASTM STP 1209, Milton J. Kowalewski, Jr., Ed., American Society for Testing and Materials, Philadelphia, 1994.

ABSTRACT: Improvements in quality often mean a reduction in tolerances of components in a manufactured item. With the introduction of statistical quality control techniques in the manufacturing process, reduced tolerances are attainable. But the measurement processes are subject to the same statistical variations. As manufacturing quality improves, the errors due to measurement become more and more a factor in continuing improvement. This paper addresses the problem of how laboratories can establish their own statistical evidence of their capability -- the accuracy of their measurements. A combination of customer need, quality function deployment, test method process charts, cause and effect considerations, and control charting are integrated into a concept of statistical measurement control.

KEYWORDS: laboratory, statistical measurement control, accuracy and precision of measurement, quality function deployment, cause and effect diagrams, test method flow charts, and statistical control charts.

Laboratories are implementing statistical techniques to demonstrate the capability of their measurement processes. Use of these techniques must be based upon the needs of the laboratory's customers who use the data for a variety of purposes. Typically, the data are used for: the evaluation and disposition of raw materials and purchased components; quality control in a manufacturing process; and the analysis of failures to determine the cause of these failures.

A key to considering when to use statistical techniques is the accuracy of the data needed by the customer. Other factors, such as the characteristics of the test methods, will also come into play. For a very expensive, complex test set-up, it may not be suitable to organize a statistical comparison because of the expense of setting up for each test.

[1]. President, American Association for Laboratory Accreditation (A2LA), 565 Quince Orchard Road, Gaithersburg MD 20878.

In this paper, the principles of Quality Function Deployment are suggested for use in categorizing the needs. Process flow diagrams are used to determine potential control points. Statistical methods are presented suitable for the laboratory personnel to use.

A decision process to prioritize the application of statistical measurement control for lab tests based on Dr. Deming's Plan-Do-Check-Act Improvement Cycle can be used as the basis for implementation of a statistical measurement control program in the laboratory:

1. Flow chart the standard method.
2. Flow chart the method as run in the laboratory.
3. Determine the need to adjust for differences in 1 and 2 above.
4. Proceed with statistical measurement control analysis one test method at a time in order of priority based on the following considerations:

- Quality Function Deployment
- Time required for the test
- Cost of the test
- Needs of the customer
- Specifications/tolerances required
- Statistical capability possible in making the measurement

5. Perform cause and effect analysis -- where could variation occur in the method?
6. Use design of experiments procedures to determine which variables most directly affect the data from the test method.
7. Create control charts of tests performed over a period of time.
8. React to charts: calibration; maintenance; improvements in the testing protocols where necessary.

THE NEED

User needs, in context of their use of test data, must be considered first in describing the need.

Types of Testing

Measurements are needed in a wide variety of applications. The need for accuracy varies with the type of testing and the purpose for which it is being done. Specific examples are:

Quality Control Testing--Perhaps the most significant operation today is testing done to know and measure the variation of materials and components in the production process and for evaluating the variations in various stages of the production process itself.

End Item Testing--One of the most significant operations in the past was testing to assure that the end-products met required specifications based on customer needs and wants.

Incoming Inspection Testing--Testing performed on incoming materials and components from which products are to be made is another significant need. This testing may be done at the supplier's facility or at the purchaser's facility.

Problem Solving Testing--Testing performed to find the causes of problems.

Calibration--Testing done in the laboratory to check or verify other testing. Calibration or verification of the performance of equipment is an example.

Research and Development--Finally, there is testing done of a research nature, where ideas are examined and new concepts are developed.

For quality control purposes, the testing process is very discrete and is often streamlined so that the data can be developed very quickly. In some cases, production is held up until the data are received and analyzed, so time and precision are extremely important.

In research, one test method will normally be used to obtain data and be followed by a completely different test method to obtain the same kind of data. If the data agree when approaching the measurement in two different ways, the measurement is considered accurate within the precision of the methods.

In all other instances, there is no regular volume of testing and the need for accuracy and precision varies.

Testing Procedures

The following types of testing are to be considered as candidates for statistical measurement control:

Chemical (& Physical):

Spectrography: AA, EDAX, Emission, IR, X-Ray Diffraction, X-Ray Fluorescence
Standard: Alloy Steel, Aluminum, Cast Iron, Stainless Steel
Elemental Analyzer: Carbon, Chromatograph, Colorimeter, Nitrogen, Oxygen, Sulfur
Wet Chemistry: pH, Qualitative, Quantitative, Specific Gravity, Titration, Volume, Weight
Physical: Coating Weight, Color Match, Flammability, Melt Index, Melting Point, Moisture Content, T.G.A., Viscosity
Coating Adhesion. Humidity, Immersion, Tape Test
Weathering: Ozone, QUV, Weatherometer
Thermal: Cycle, Freezer, Oven
Artificial Aging: Air, ASTM Fluid

Mechanical Testing:

Strength: Adhesion, Compression, Elongation, Modulus, Peel, Tear, Tensile, Yield
Hardness: Brinell, Durometer, Microhardness (DPH, Wallace), Pencil, Standard Rockwell, Superficial Rockwell
Polymeric: Burst, Circulation, Dynamic Rate, Efficiency, Kink, Rheometer, Vacuum Collapse
Other: Abrasion, Fatigue, Gavelometer, Impact, Torque-Tensile, Torque, Vibration, Weld, Pressure
Corrosion: CASS, Neutral Salt Spray
Metallography: Case hardness, notch sensitivity, nodularity, grain size, coating thickness.

Nondestructive Testing (Evaluation):

Radiography: Fluoroscope, Isotope, X-Ray
Superficial: Acid Etch, Dye Penetrant, Magnetic Particle, Zyglo
Special Purpose: Eddy Current, Magnetic, Sonic, Ultrasonic
Coating Thickness: Anodize, Film, Paint, Plate

Quality of Testing Operations

Quality is judged in the eyes of the customer. For that reason, quality can be considered to consist of at least six attributes.

1. Is the accuracy of the data adequate to serve the customer's need?
2. Is the precision of the data within desired tolerance?
3. Are data interpreted correctly by the laboratory?
4. Are data received in a timely fashion?
5. Are costs comparable with alternatives?
6. Does the laboratory provide appropriate consultation?

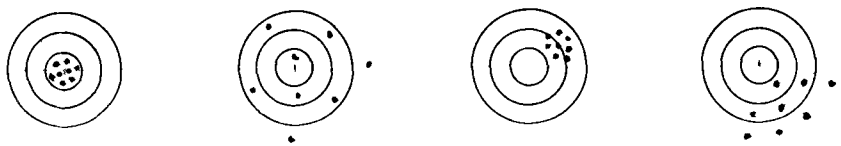
The thrust of this paper is on the development of statistical measurement control addressing the accuracy and precision capability of a laboratory.

Perhaps the most efficient way of finding problems in the accuracy and precision of measurements in the laboratory is to use quality control charts on a regular basis. The advantage is that problems can be found quickly and corrections to the measurement process made very near to the time that the problems occur. The disadvantage is that additional testing of control samples is required in volume and frequency sufficient to have meaning when subject to statistical rules of significance.

Precision is made up of repeatability of a measurement (one operator over a short period of time) and the reproducability (full range of alternative measurement arrangements over a long period of time). Within laboratory variation is sometimes difficult to measure so reproducability measured over a longer period of time as described in the statistical aspects of this paper are sometimes better estimates of precision.

Accuracy is the closeness of agreement between an observed value and an accepted reference value. Ideally, in statistical measurement control the control charts depict accuracy. But in order to determine accuracy, there must always be a "true" value obtained by measuring a reference standard with its known variation expressed statistically.

Accuracy is made up of two elements, precision and bias. (Bias is sometimes referred to by its opposite, trueness). Precision is the closeness of agreement between randomly selected individual measurements or test results. Bias is the systematic error that contributes to the difference between a population mean of the measurements or test results and an accepted reference or true value. The Target Analogy shown in Figure 1 illustrates the relationship between precision and bias.



Accurate: Precise, True | Inaccurate: Imprecise, True | Inaccurate: Precise, Biased | Inaccurate: Imprecise, Biased

Fig. 1--Precision, bias, trueness, and accuracy

When there are no suitable reference materials or reference methods, bias, strictly speaking, cannot be determined. However, in many instances a reference value established by the equipment manufacturer or by data from round-robin tests is often used as the true value. Where there are not any available reference materials with known accuracy or a consensus true value, a laboratory can manufacture their own materials and perform statistical analyses using a significant

number of samples to give confidence in the quality of the data produced. An ASTM Standard, EXXXX.93 is being developed to provide guidance in this area.

In statistical measurement control, experiments should be designed to measure accuracy if possible. If this is not possible, then the experiments will measure precision.

OVERALL CUSTOMER REQUIREMENTS PLANNING MATRIX

Users of test data are striving continually to improve the quality of their operations. As their quality improves, the quality of the test data which they use to make decisions needs to improve also. Denis Swytt of NIST (Reference 1) describes this phenomenon in reporting on remarks make by Taguchi, describing one aspect of Japanese automobile manufacture:

> According to the account, quality in the operation of an automobile door as perceived by customers correlates with variations in the force required to open the door and a comparison showed that to open them US cars required forces of 76 $\pm$ 58 N (17 $\pm$ 13 lbs) while Japanese cars required 31 $\pm$ 9 N (7 $\pm$ 2 lbs), the Japanese companies thus holding a six-to-one advantage over U.S. companies. The variations in force correspond to variations in the dimensions of door-assemblies which in turn correspond to variations in the dimensions of the dies from which panels are stamped. With a 3σ norm of 1 mm on door assemblies, Japanese cars have half the variation in such dimensions as the US-made cars.

A major component of test data quality is its accuracy (precision and bias). Many customers do not know the accuracy that they need. Often, the laboratory must interpret the need and provide data commensurate with that interpretation. In the past, no formal statement of data accuracy was recorded; the routine performance of the laboratory was presumed to provide the needed accuracy. This is changing.

Quality Function Deployment

In order to formalize customer requirements, Quality Function Deployment (QFD) concepts are proposed. QFD is a philosophy, specification, and mechanism for deploying customer desires vertically and horizontally throughout the company. It has been designed for use in the manufacturing process. It has been adopted for use in the laboratory as a way to put the laboratory quality issue in context.

Controlling the Measurement Process

The American Society for Quality Control (ASQC) has a committee developing a relevant standard (Reference 2). Two methods are described for controlling the measurement process, the "Program Controls Method" and the "Measurement Assurance Method".

In effect, the "Program Controls Method" implies that if the measurement process (ensemble) can measure with an accuracy which is significantly more than required by the customer, no statistical analysis is necessary. An accuracy ratio (required tolerance over the measurement uncertainty known for the process) of four is often used as a guideline; if the measurement process has an accuracy at least four times better than the required measurement tolerance, it may be considered as not affecting the measurement process.

The second method is called the "Measurement Assurance Method" which implies that something specific is known about the measurement process and its uncertainty and that one way to do this is through the use of quality control charts. It is this approach which we will focus on in this study of Statistical Measurement Control.

This paper is intended to guide laboratories in exploring when and where to use quality control charts to demonstrate that the laboratory testing system is in control.

Customer Requirements for Accurate Data

Following the overall guidance of QFD, and using the principles from the pending ASQC standard, the laboratory must specify the required accuracy (measurement tolerances) for each customer for each test. If the tolerances are not available from the customer, the laboratory should make an estimate after understanding the customer's use of the data.

Next, the laboratory must indicate the accuracy it can achieve for each test. It is conceivable that the accuracy attainable for a given test can be different for different groups of customers using the same test, but normally, one estimate of accuracy is all that will be available. ASTM E691 (Reference 3) is often used a basis for determining this accuracy. This value for the accuracy can also come from designed experiments when such studies have been done. More likely, this value of accuracy is an estimate, based on claims of equipment manufacturers or on experience gained over the years.

Notation should be made as to the use of the data; whether as input to a production process, problem solving, component and materials acceptance, etc. Figure 2 has been prepared to collect the required user data for each test method.

CUSTOMER ACCURACY OR PRECISION REQUIREMENTS
LABORATORY _____

Test Method Designation _____

Title _____

Customer	PARAMATER # 1	# 2	# 3	Comments
1. XYA Corp.				Lab. Est; Customer Spec; or ? SMC; Equip. Manu.; Exp.; or ? Use; SPC; Incoming Accept
2. XYB Corp.				
3. XYC Corp				

Fig. 2--Customer accuracy and precision requirements.

PROCESS DESCRIPTION FOR EACH TEST METHOD

Once the tolerances have been determined and the capability estimated for each test, it is necessary to decide which tests should be considered for possible statistical measurement control determinations. If the laboratory capability is always at least four times better than the tolerances, then statistical measurement control may not be necessary. But, if the estimated capability is less than four times the tolerances, then the test should be considered for statistical measurement control. Most test methods should be amenable to statistical measurement control. That is, suitable reference materials are available, and the tests do not require an inordinate amount of time to set up or conduct.

A key step is to identify possible control points which should be used in making measurements to control the testing process. One way to do this is to develop process descriptions or flow charts of how each test method is being implemented in the laboratory. These process descriptions can take on many different formats.

There is no one way to describe the process that is used in a given laboratory. The main point is to look at the process in enough detail to identify those parts of the process which might affect the test performance. Control charts will then measure results at those parts of the process to determine if they are in control. It is often useful to compare the process as performed in the laboratory with the process as described in the test method. This may give additional clues to where control charts may be used to monitor testing operations.

From the process diagrams, intuitive judgment can be used to identify those control points most likely to be important in statistical measurement control. Intuition may not be very reliable because the major causes of error may not be identified in the process description if there are second order effects or because intuition is not that trustworthy. A cause and effect analysis may be helpful in identifying the second order effects. A designed experiment or ruggedness testing may be helpful in verifying intuitive judgment.

Cause and Effect Analysis

Cause-and-Effect Diagrams (sometimes known as Fishbone Charts) permit a more formal assessment of the factors important to the measurement which may go beyond the process flow diagram. For each control point, one specifically questions the factors which can influence the value at that point. Brainstorming is used to identify factors which can affect that control point and the effect those factors could have on the value of the control point.

An example of such an analysis for tensile testing is shown in Figure 3. The branches of this diagram are the factors that affect the quality of the test. Personnel performing the test are in the best position to identify as many possible effects of quality. A brainstorming by staff should be used to identity all possible effects on the quality of the test. The cause-and-effect diagram gives people a structure for working together on the same test.

Once all of the elements of the diagram have been identified, it is necessary to identify those which can be readily controlled. measured results.

In this way, it is possible to identify the critical component characteristics and the most likely control points for each element in the process (for performing a given test).

The Quality Functional Deployment procedures call for the development of a process plan chart. This chart will identify the control points, monitoring methods and frequency for each test method which will possibly be subject to statistical measurement control.

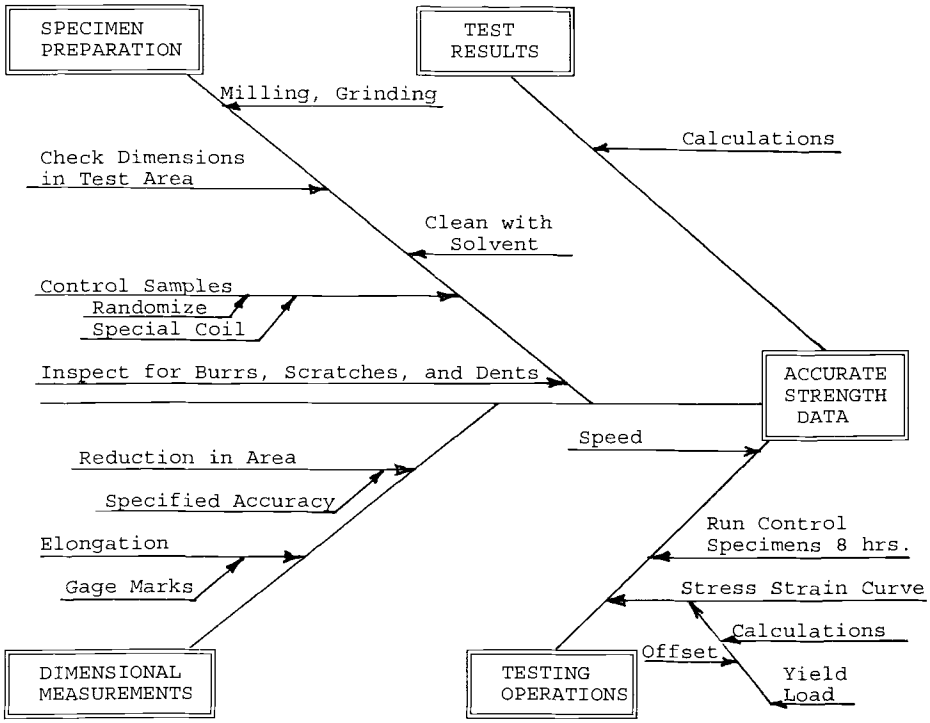


Fig. 3-- Cause and Effects (Fishbone) Diagram

Controlled Experiments

If it is difficult to determine the most important elements in the testing process to control, designed experiments such as are suggested by ruggedness testing (Reference 4) should be used. The purpose of a ruggedness test is to find the variables that strongly influence the measurements provided by the test method, and to determine how closely these variables need to be controlled. The ASTM reference uses the "Plackett-Burman" designs with two levels per variable. The designs require the systematic change of the levels of all of the variables, and the determination of the effects of each of the variables on the

QUALITY CONTROL PLAN FOR DATA COLLECTION AND ANALYSIS

There are two types of data, variables (typically numbers) and attributes (typically visual verification of existence of some element). Most laboratory data will be of the first type. The measurement process which produces the variable data can be evaluated at one point to determine the critical variables which need to be controlled by conducting a ruggedness test [Sometimes called a Reproducibility and Repeatability ("R & R") study]. Each element in the measurement process is systematically varied to determine those which have a significant effect on the data obtained from the measurement process. An R & R study will not, by itself, provide a prediction of what will happen under the varying conditions of use or whether the same level of

variation observed at one time will occur in the future. However, it will identify the most important variables for control during the testing operations. For the control of these important variables in the measurement process over time, a Shewhart control chart must be used. Various types of control charts which may be used are described in Reference 5.

Data Collection for Measurement Control.

When a test measurement is critical to the user, the measurement process must be of sufficient accuracy to support the user's need. If standard reference materials, which have a known accuracy, are available, they can be used to determine the accuracy of the laboratory's measurement. If relevant standard reference materials are not available, the best that can be determined by statistical measurement control is the precision possible in the laboratory, based on the use of internal reference materials.

The accuracy/precision comes from taking repeated measurements over time. The terminology can be illustrated in the following example:

Suppose two readings were taken under the same conditions and these readings are repeated at successive points in time (time is really the surrogate for changing conditions):

Time 1	Time 2	Time 3
X	X	X
X	X	X

Suppose X-bar and R charts are constructed from the resulting data. The following results and corresponding conclusions might occur:

R chart in control: Variation in repeated readings under the same conditions is predictable -- "accuracy/precision" is quantified

R chart out of control: Variation in repeated readings under the same conditions is not predictable -- the measurement process has no quantifiable degree of accuracy/precision.

X-bar chart in control: The average reported value for the internal reference material is stable (not changing with a change of conditions over time). The measurement process has a degree of bias to the (possibly unknown) external standard which is predictably the same.

X-bar chart out of control: There is no such thing as precision or bias (or accuracy). The potential amount of deviation from the standard (be it internal or external) varies and depends upon when (under what conditions) a measurement is made.

The idea is that if there has been no study of the measurement process over time (and therefore conditions), and no other information about whether the measurement process is predictable, then nothing can be said concerning future precision or bias of the method. If the measurement appears to be stable, then we can say that we can rely on the measurement process internally. If we have an standard reference material, then we can go further (if we have a stable process) and quantify the degree of bias relative to the external standard.

When we do not have a standard reference material, we can still determine the degree of imprecision. The key reference samples can be controlled samples produced internally from material selected for its uniformity in mechanical properties. Samples such as these can be used to measure and control the precision of the measurement process, but not its accuracy. The use of Standard Reference Materials produced by the National Institute for Standards and Technology (NIST) or other sources which are traceable to the National Measurement System is ideal in that these provide the basis for determining accuracy. Often such samples

are not available, in which case the laboratory can determine only the precision of its measurements and must rely on process controls and calibration procedures to minimize any possible bias in the final test measurement result.

Even though a standard reference material is not available for the particular kind of sample provided by the customer, a reference material can be used in the measurement system to calibrate the equipment and in this way provide some form of traceability to national standards attained. One must be sure not to change the calibration of the equipment when the data obtained are within the limits of the reference standard. Such changes would only increase the uncertainty of the measurements.

A Quality Control Data Collection Plan should be established to describe how the laboratory will proceed with the measurement control program for a specific test and specific samples. Every plan needs to identify the frequency and conditions of measurement. Ideally, these conditions will be identical to those in the routine testing process. Data collection sheets must be developed and should include:

Date	Type of Measuring Equipment
Time of Day	Lot Number of Control Samples
Laboratory Location	Purpose of Data Collection
Testing Machine Control Point	
Operator	Lab Conditions (Temp., Humidity, etc.)

The data on these sheets will form the basis for evaluation of each control point.

Data Analysis.

The following sections describe some standard data analysis tools which have proven useful in properly controlling the accuracy/precision of laboratory measurement processes. See the references for complete details of these procedures.

Ruggedness Analysis. Reference 3 describes an analysis procedure for the determination of the sensitivity of a test method or variations in control parameters or variables. A ruggedness analysis is essentially a screening design of experiments. It may be used to supplement or replace the Pareto chart analysis discussed below. Other design of experiments procedures such as those of Taguchi, can be used if found to be more suitable. It is recommended that such an analysis be performed for each test method to identify those control parameters/variables which exert the greatest influence on the test measurement response. Any such designed experiment will only assess the effects of each control variable under existing conditions. When operating procedures are changed, such as by the introduction of new instrumentation, testing machines, etc., this analysis must be repeated.

Graphical Analysis. Pareto charts are constructed to show which of the known control points have the largest effect on the measurement process. These charts order the control points on the basis of their contribution to accuracy, precision, cost, etc., so that those control points which are the most influential, the "vital few", may be determined.

Control Charts. Control charts for data on variables are used to show the measurements in time sequence. The X-Bar chart shows the variability in subgroup averages which may indicate a lack of control relative to an overall average value due to assignable cause, and the Range Chart which, when in control, measures the random or common cause variation in the measurement process.

The data are arranged in time sequence and summarized in rational subgroups. The underlying concept of the Shewhart Control Chart is that within a subgroup there is only chance, random, or common cause variation; while all assignable or systematic variation will occur between subgroups. Each measurement process must be examined relative

to the timing of each measurement and laboratory operation to determine the proper subgrouping. For example, the measurements from two different machines or two different operators must not be included in a single rational subgroup. The subgroup size is often 3 to 7 measurements and should be constant for a given control chart. However, there may be situations when measurements are obtained so infrequently that they can be logically grouped and a control chart for individuals or a moving average would be more suitable. In such a case, a moving range would be used for the R Chart. There are other forms of Control Charts, for example The Lot Plot, CumSum, etc. and these may be used if found suitable instead of the standard Shewhart X-bar and R charts. (See References 4 and 5 for further details.)

X-bar, R Charts.

The following procedures are recommended in preparing the X-bar R charts:

Collect Preliminary Data--It is recommended that the initial control chart be based on a least 25 subgroups if these can be obtained in a short time period. If not possible, a chart may be constructed with as few as 8-10 subgroups and the limits later modified as more data become available. Individual observations may be plotted in time sequence as a run chart, however, such a chart does not have the same ability to detect assignable causes as does a chart for the rational subgroup averages.

Calculate the Range, R for Each Sample--For each rational subgroup of 3 to 7 observations, called a sample, the range is calculated as the signed difference between the largest and the smallest observation in the subgroup.

Calculate the Average Range, R-bar--The centerline of the Range Chart will be R-bar, the average of all the ranges.

Determine the control Limits for R--The upper and lower control limits for R are determined as follows:

$$UCL_R = \text{R-bar} \times D_4$$
$$LCL_R = \text{R-bar} \times D_3$$

The factors D_3 and D_4 are given in the Figure 4.

Subgroup Size (n)	Factor for X-Bar Chart A_2	Factors for Range Chart Lower Limit D_3	Upper Limit D_4	Factors for Estimating σ~ d2
2	1.880	0	3.267	1.128
3	1.023	0	2.574	1.693
4	0.729	0	2.282	2.059
5	0.577	0	2.114	2.326
6	0.483	0	2.004	2.534
7	0.419	0.076	1.924	2.704
8	0.373	0.136	1.864	2.847
9	0.337	0.184	1.816	2.970

Fig. 4 Factors for control charts

Construct the R Chart--The individual R values for each sample are plotted horizontally in time sequence. Each point should be identified as to time, lot number, etc. Consecutive points are connected to indicate their time sequence. Add the center line and control limits.

Interpret the R Chart--The references on Control Charts (4,5,6,and 7) have extensive lists of visual tests for a chart to show lack of control. However, the following two may be sufficient for most work:

Control Limit Test: Do any of the R values fall outside the control limits?

Runs Test: Examine the chart and see if there are runs of seven or more consecutive points above or below the center line or in a consistent increasing or decreasing sequence.

Where local rules for detecting lack of control have been documented, these should be used.

If the R chart passes these tests, it can be said that the measurement process is in control with respect to the range. If, on the other hand, any points fail these tests, lack of control is indicated and an investigation to determine the reasons must be made. It is not sufficient just to plot the chart. Control limits for the X-bar Chart are not meaningful unless the Range Chart shows control.

Construct the X-bar Chart--The subgroup averages are plotted horizontally in time sequence. The center line is X-Double Bar, the average of all the X-Bars. The control limits are calculated as:

$$UCL_{X\text{-}bar} = \text{X-Double bar} + A_2 \times \text{R-Bar}$$
$$LCL_{X\text{-}bar} = \text{X-Double bar} - A_2 \times \text{R-Bar}$$

Plot the X-bars and draw the center line and control limits.

Interpret the Chart--The same two tests for lack of control given above are used to test the X-bars for lack of control. Any lack of control shown in the chart is an indication of some assignable cause and must be investigated. Again, it is not sufficient just to plot the chart.

Continuation of the X-bar and R Charts--Once the preliminary limits have been calculated as above and both the X-bar and R charts show control, they are to be used for further monitoring the measurement process until there is some reason to change limits. One such reason would be a significant improvement in the process indicated by a sustained run below the average in the Range Chart. In those cases where the preliminary chart shows any lack of control, each point should be investigated for assignable causes. Where assignable causes are identified, appropriate corrective action is taken and the limits recalculated removing this subgroup. After the chart has been established, corrective action is taken for any point out of control and the limits revised only when significant change in the process has been achieved. A periodic review of the control limits should be made on a regular basis to verify that the limits truly reflect the present state of measurement control.

Determination of Process Capability--The natural tolerance (NT) or capability of the measurement process may be determined when both the X-Bar and the Range charts are in control.

An estimate of the standard deviation of the process, σ, can be calculated as:

$$\sigma = \text{R-Bar}/d_2$$

Where d_2 is based on the subgroup size as shown in the table of control chart factors below.

The natural tolerance or process capability is given by:

$$NT = 6\ \sigma$$

This natural tolerance may be compared with the customer's required measurement tolerance, perhaps and engineering tolerance (ET). Note

that if the natural tolerance is compared with an appropriate engineering tolerance, NT must be very small, say 10-20 percent compared with ET to allow product/process to be within variation prescribed by its specification. If the ratio of NT/ET is not acceptable, problem solving techniques must be used to determine how the process may be improved in order to provide acceptable precision data. Again note that both the X-Bar and Range charts need to be in control in order to determine NT or process capability. If there are significant shifts in the subgroup averages indicating assignable causes and these are not explainable, the interpretation of the process capability measure may be meaningless.

When the ratio of the natural tolerance to the customers's measurement tolerance is quite small, indicating a measurement process which is very capable of meeting the engineering tolerance requirements, the frequency of control testing may be reduced to an audit level. The control chart should be maintained, however, as a visual monitor of the measurement process.

For additional information on the presentation of data and control chart analysis, see References 6 and 7.

CONCLUSION

An orderly procedure for organizing and applying statistical measurement control in a laboratory has been described. The need for this kind of quality control data is growing rapidly.

REFERENCES

[1] Swyt, Dennis A., "Challenges to NIST in Dimensional Metrology: The Impact of Tightening Tolerances in the U.S. Discrete-Part Manufacturing Industry", NISTIR 4839, dated January 1992.

[2] _____, "American National Standard for the Quality Control of Measurements", ANSI\ASQC Q4-19XX, Draft dated 24 April, 1989.

[3] _____, "Standard Practice for Conducting and Interlaboratory Study to Determine the Precision of a Test Method", ASTM E 691-92

[4] _____, "Standard Guide for Conducting Ruggedness Tests", ASTM E 1169-89

[5] Grant, Eugene L. & Leavenworth, R. R., "Statistical Quality Control", McGraw Hill Book Company, New York, Fifth Edition, 1974.

[6] Committee E-11 on Quality and Statistics, "Manual on Presentation of Data and Control Chart Analysis: Sixth Edition" ASTM Manual MNL 7, 1992.

[7] Kowalewski, Milton J. and Tye, Josh B., "Statistical Sampling, Past Present and Future Theoretical and Practical", ASTM STP 1097, 1990.process as described in the test method. This may give additional clues to where control charts may be used to monitor testing operations.

Thomas A. Daly [1]

WHERE REAL PRODUCT QUALITY STARTS

REFERENCE: Daly, T. A., **"Where Real Product Quality Starts,"** Quality and Statistics: Total Quality Management, ASTM STP 1209, Milton J. Kowalewski, Jr., Ed., American Society for Testing and Materials, Philadelphia, PA 1994.

ABSTRACT: Current master plans for "Total Quality Management" and the "Deming Approach" devote rleatively minor attention to a key operation- the actual design of the product. This paper discusses techniques and procedures for management of the design function to assure optimum product quality and reliability, with the objective of complete customer satisfaction. Areas covered are: Design Personnel, Product Design Specification, Formal Design Review, Fault Tree Analysis, Failure Mode and Effect Analysis, and Field Performance Data Collection and Use.

KEYWORDS: Product design, quality, reliability, engineering management, design review, concurrent engineering,failure analysis, computer-aided design

INTRODUCTION

The much-respected management consultant Peter Drucker has observed that "Nothing is less productive than to make more efficient what should not be done at all!" (1). This statement might serve as a word of caution to the devotees of such all-encompassing programs as Total Quality Management, Deming's Fourteen Points, and the Malcolm Baldrige National Quality Award.

There is no doubt that these programs promote some very worthwhile practices, but they are also perfect examples of the need for application of Pareto's Law. This principle can be paraphrased briefly by stating that a relatively small number of problems in a given area will be found to be accountable for a large percentage of the resultant costs.

In other words, before embarking on a company-wide TQM program it might be wise to conduct a Pareto analysis and establish priorities for those areas which have the most potential for improvement. The

[1]Consultant in Management and Reliability, 5024 Northlawn Drive, Murrysville, PA 15668.

hard cold fact exists that today in most companies neither money nor manpower is available to scrutinize every department and institute an all-out attack calling for continuous improvement forever.

However, some organizations are trying to do exactly that, and this has reulted in a feeding frenzy by all sorts of consultants in fields from corporate accounting to psychoanalysis. For example, before it was recognized and rolled back, the TQM program at Florida Power and Light had swelled to a total of 75 "quality consultants".(2). At the Aluminum Company of America, one of the first moves of CEO Paul O'Neil after taking over his new position was to drastically curtail the corporation's massive TQM effort. (2).

In the midst of these actions, we find some discrepancies in the master plans for "total" quality. One of these is the relatively minor amount of attention paid to a key operation--the engineering design of the product. Of the total of 1000 points which can be awarded in the examination of a company for the Malcolm Baldrige National Quality Award, only 40 points are allocated to the "Design and Introduction of Quality Products". (3). Similarly, while Deming's Fourteen Points, which are regarded by enthusiasts as the Ten Commandments of the TQM movement, call for "constancy of purpose", the adoption of "a new philosophy", and the elimination of "dependence on inspection to achieve quality", no mention is made of "design" or "engineering".

MANAGING FOR PRODUCT DESIGN QUALITY

Fifty years ago it was common to see signs in many plants that read "Quality Cannot Be Inspected Into The Product". This was tacit admission that quality had to be designed and built into the product, yet years went by when the major efforts of quality control people seemed to be concentrated on sampling plans, percent defective reports and quality costs. Of course, the majority of quality control experts in those days were inspectors and statisticians with little or no experience in the actual engineering design of products. Furthermore, there was a tendency for the product to be designed wholly within the walls of the engineering departmentand then thrown over the fence to the manufacturing people, who had to figure out how to produce it.

No longer can a product design, even the simplest, be the result of a single engineer's talent and effort, even though some engineers might still like to have it that way. Engineering operations must be controlled by procedures and schedules that ensure that , as early as possible in the development cycle, all essential input is provided to the design team. This means all of the customer's requirements, the environmental limitations, manufacturing capabilities, part and material requirements, past field performance of similar designs, economic factors, etc. Some recommended practices and techniques follow.

PERSONNEL QUALITY

An effective recruiting and training program should be instituted that ensures that those given design responsibility are of sufficiently high calibre. Engineering education in the universities, it is now generally agreed, has been far too limited in actual design work. Undergraduate engineering courses are presently undergoing the first stages of a metamorphosis, based on the realization that design has been given short shrift in the traditional engineering curriculum. Engineering students should be required to develop a product concept, do the calcul ations and the actual drafting , and select the latest manufacturing processes and materials. (4).

However, until there is more evidence that such changes are having an effect, a large part of an engineer's design experience will continue to be acquired through on-the-job training, supplemented by courses in special procedures and techniques such as those outlined below.

PRODUCT DESIGN SPECIFICATION

Seldom are the original specifications supplied by the customer or by the marketing research department sufficiently detailed to provide all the information needed by the design team to come up with a complete, highly reliable, customer-pleasing product. The design specification should include as a minimum the following:

1. Customer uses,life expectancies, and environments
2. Required functions and reliability goals
3. Constraints (weight, space, power, etc.)
4. Competitive product characteristics (benchmarking)
5. Human engineering and safety considerations
6. Maintenance and logistics requirements
7. Transportation and packaging demands
8. Purchased parts quality and reliability
9. Cost objectives
10. Test and inspection requirements
11. Technical documentation needs (instruction manuals, etc.)
12. Field performance of similar designs
13. Customer satisfaction with previous designs
14. Applicable standards (American National Standards Institute/ American Cociety for Quality Control Standards Q90 to Q94- "Quality Systems"; International Electrotechnical Commission Standard EC 300- "Reliability and Maintainability Management"; etc.)

A complete, detailed specification provides assurance that every one on the design team is working with all the information available and with the same information, before large amounts of time and money are spent on the project.

FORMAL DESIGN REVIEW

Much has been written about the use of "quality circles" in the office and on the shop floor, but one of the most critical and productive areas for a team effort is the design operation. Here

the decisions are made that will determine how long the product will last, how well it will operate, and whether it will live up to all of the customer's expectations.

Embryonic efforts in this direction started over 40 years ago, when reliability engineers set up design reviews in military contracts. In the Sixties these meetings evolved into Formal Design Reviews (FDR) (5). In these reviews all involved disciplines and departments were represented, and team members presented comments on the design from the aspect of each of their specialties. Detailed checklists were developed for electrical, electronic, hydraulic, mechanical and other elements of the design to ensure that the reviews would be comprehensive.

An organized FDR program spells out in written procedures the method of conducting the review, the information to be supplied to each team member in advance, the documentation to be kept on all items discussed, and calls for fillow-up on all action items. The project schedule identifies in advance the points where conceptual, briefing, preliminary, intermediate and final FDR's are to be held. (Fig. 1).(6).

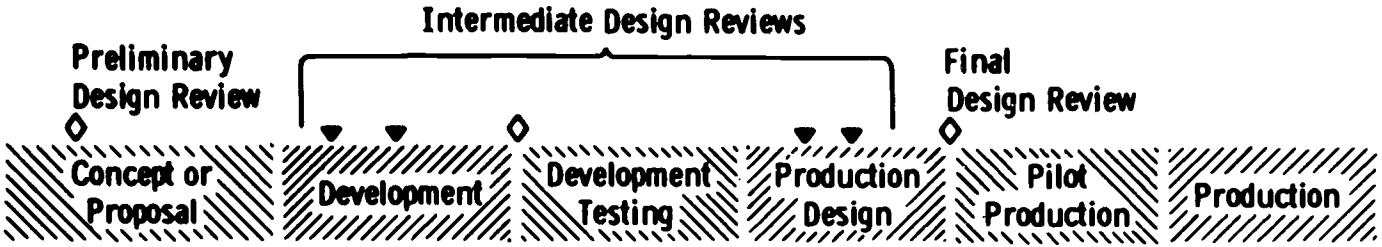
A Formal Design Review program, when properly activated, is one of the most powerful means of getting maximum input to and concerted effort from all participants in a design project from conception to final production. It can be seen from the foregoing description that FDR programs were the forerunners of, and present an organized approach to what is currently being called Concurrent Engineering.

From an economic standpoint, studies of a representative group of FDR's and of the field performance of the end product have shown estimated savings of $10 to $20 for every dollar spent on the program.

FAILURE MODE AND EFFECT ANALYSIS

Failure Mode and Effect Analysis (FMEA) is another technique for improving product quality and reliability that was pioneered by engineers working on military contracts. (7). It is a systematic method for (a) considering all the possible modes of failure of each component of a product or system, and (b) developing the effect of each of these failures on the ultimate performance of the designed item. These effects are then assigned degrees of criticality and probabilities of occurrence. Corrective action is recommended where necessary and tracked to ensure that assignments are followed through to completion. (Fig. 2).

FMEA offers a valuable technique available to engineering managers to provide additional assurance that a design incorporates a high degree of reliability. It provides a set of tools for finding and assessing failure possibilities and for eliminating them prior to final production. Granted, it requires additional time and manpower but it may save many times the cost of that effort by eliminating field trouble-shooting, redesign and retooling expense, plus the resultant customer dissatisfaction.



Product life cycle and design review schedule

FIG. 1

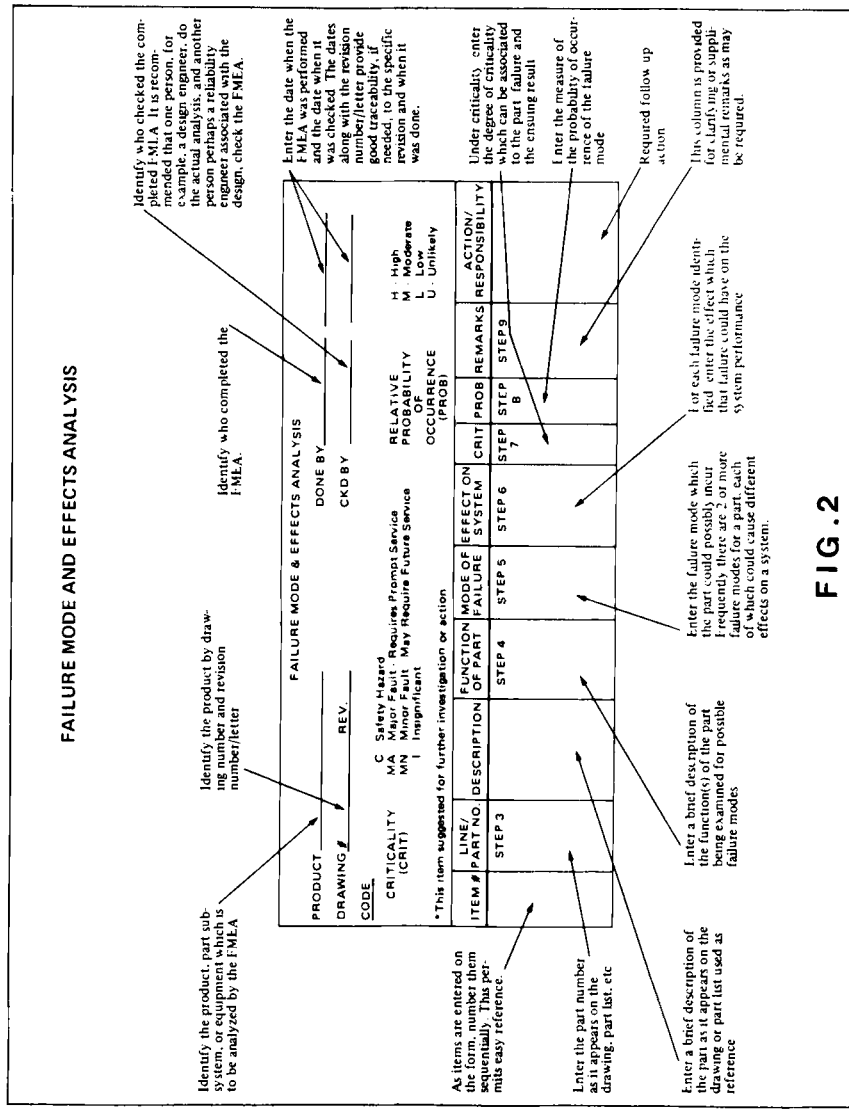

FAILURE MODE AND EFFECTS ANALYSIS
FAILURE MODE & EFFECTS ANALYSIS
PRODUCT
DRAWING #
REV.
CODE
DONE BY
CKD BY
CRITICALITY (CRIT)
C Safety Hazard
MA Major Fault - Requires Prompt Service
MA Major Fault - May Require Future Service
I Insignificant
RELATIVE PROBABILITY OF OCCURRENCE (PROB)
H High
M Moderate
L Low
U Unlikely
* This item suggested for further investigation or action
ITEM # | LINE/ PART NO. | DESCRIPTION | FUNCTION OF PART | MODE OF FAILURE | EFFECT ON SYSTEM | CRIT | PROB | REMARKS | ACTION/ RESPONSIBILITY
STEP 3
STEP 4
STEP 5
STEP 6
STEP 7
STEP 8
STEP 9
Identify the product, part subsystem, or equipment which is to be analyzed by the FMEA
Identify the product by drawing number and revision number/letter
Identify who completed the FMEA.
Identify who checked the completed FMEA. It is recommended that one person, for example, a design engineer, do the actual analysis, and another person perhaps a reliability engineer associated with the design, check the FMEA.
Enter the date when the FMEA was performed and the date when it was checked. The dates along with the revision number/letter provide good traceability, if needed, to the specific revision and when it was done.
As items are entered on the form, number them sequentially. This permits easy reference.
Enter the part number as it appears on the drawing, part list, etc
Enter a brief description of the part as it appears on the drawing or part list used as reference
Enter a brief description of the function(s) of the part being examined for possible failure modes
Enter the failure mode which the part could possibly incur. Frequently there are 2 or more failure modes for a part, each of which could cause different effects on a system.
For each failure mode identified enter the effect which that failure could have on the system performance
Under criticality, enter the degree of criticality which can be associated to the part failure and the ensuing result
Enter the measure of the probability of occurrence of the failure mode
Required follow up action
This column is provided for clarifying of supplemental remarks as may be required.

FIG. 2

Computer-aided design (CAD) programs have been developed that are aimed at reducing the time required to perform an FMEA. However, some of these are so sophisticated that a manual approach may be preferable for domestic products. (8).

FAULT TREE ANALYSIS

In systems which present safety hazards to personnel and the possibility of liability litigation, it is particularly important to assess in depth the reliability of the product during the design and development phase. One of the risk assessmant methods developed by safety and reliability engineers is known as Fault Tree Analysis(FTA). This technique has three principal uses:

(a) To detect and correct weaknesses in a new design,
(b) To study a past failure and determine the sequence of events which contributed to the malfunction, and
(c) To develop a trouble-shooting chart for incorporation in maintenance manuals.

The construction of the fault tree begins with the listing of a system failure in the top rectangle (Fig. 3). Next, all events which could directly cause this failure are shown in rectangles in the level below, and connected to the top box by logic symbols. An "and" symbol is used if more than one event must occur in order to cause the failure, and an "or" symbol if a single event precipitates it.

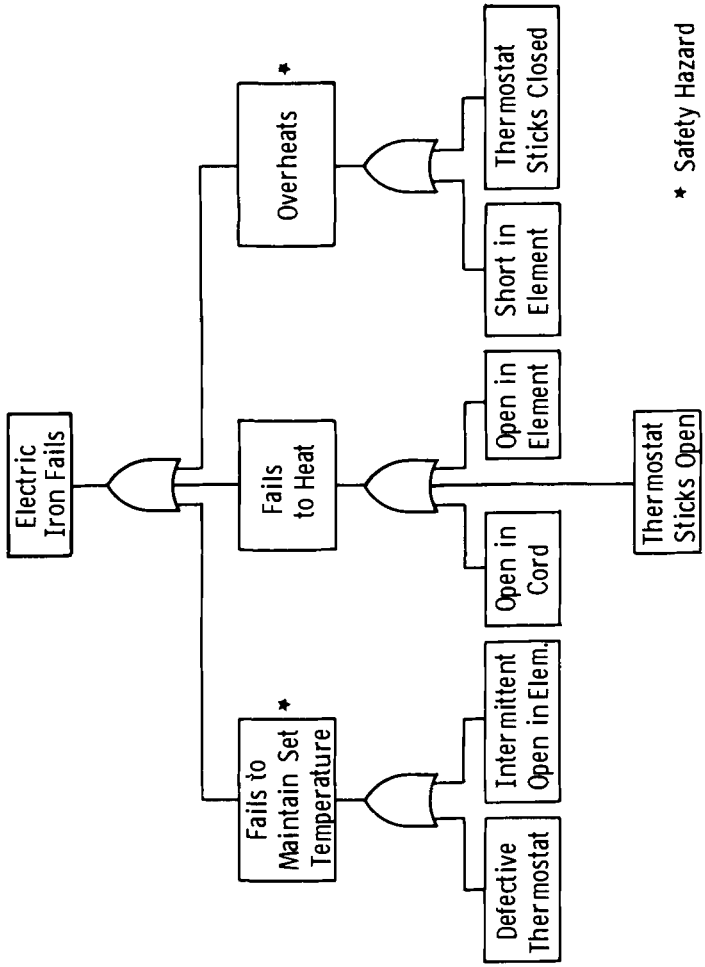
Each event in the second level is then studied , and all events which could lead to its occurrence are identified in the same manner on the third level. This process is continued until the lowest assignable is reached for each branch. As in the FMEA, degrees of criticality and probabilities of occurrence can be assigned to each event, and recommendations for corrective action recorded. (9).

of the two techniques described above, the FMEA is the most intensive, since it starts by considering every component of the design and its possible modes of failure from the bottom up, while the FTA starts with the premise of a product failure and works down through those components which could be involved in that specific malfunction.

FIELD PERFORMANCE DATA COLLECTION AND USE

Engineering management for high quality designs cannot be done in a vacuum--it requires accurate and complete knowledge of the performance of previous and similar products in the field. No "total quality" program can succeed without an adequate system for the collection and analysis of performance history from the customer, from field service engineers and from any other source available.

The data collection system must provide forms which facilitate detailed reporting of all significant troubles and customer complaints. Such forms should be so organized that the data can be coded for entry into a computer data base. The information can then be processed to provide valuable guidance with regard to failure rates of components



FAILURE MODE ANALYSIS FAULT TREE

FIG. 3

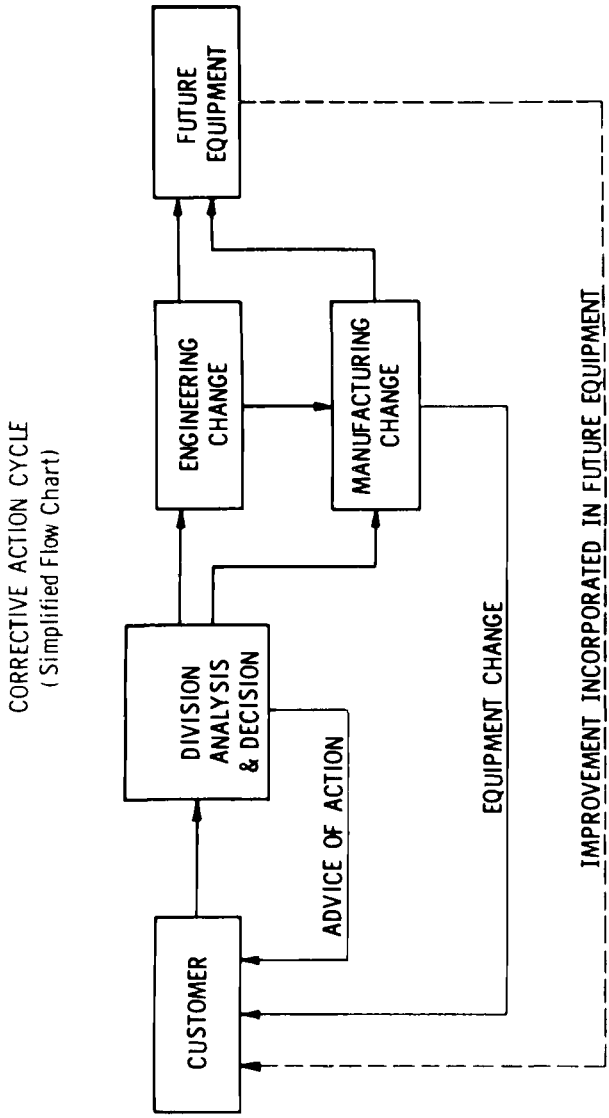

CORRECTIVE ACTION CYCLE
(Simplified Flow Chart)
CUSTOMER
DIVISION ANALYSIS & DECISION
ENGINEERING CHANGE
MANUFACTURING CHANGE
FUTURE EQUIPMENT
ADVICE OF ACTION
EQUIPMENT CHANGE
IMPROVEMENT INCORPORATED IN FUTURE EQUIPMENT

FIG. 4

and materials in actual environments and customer usage. Obviously, the more accurate and complete the reports are which come in from the real world, the more valuable this information will be to the FDR team and those who conduct FMEA's and FTA's. (Fig. 4).

CONCLUSION

It is noteworthy that the Food and Drug Administraion, which exercises control over all medical products and devices by means of a "Good Manufacturing Practices" regulation, has recently acknowledged the vital importance of the design function in determining the ultimate quality of any product. The Safe Medical Devices Act adds to the above regulation a requirement for "preproduction quality assurance" , which emphasizes the need for devoting more attention to the design and development operations through the use of design reviews, failure analysis techniques and detailed determination of customers' requirements. (10).

REFERENCES

(1) Drucker,Peter F. , Management Tasks,Responsibilities,Practices, Harper & Row, 1973.

(2) Sharman, Graham, "When Quality Control Gets in the Way of Quality", Wall Street Journal,February 24,1992.

(3) 1992 Award Criteria, Malcolm Baldrige National Quality Award, National Institute of Standards & Technology, Gaithersburg,MD.

(4) "Higher Grades for Accreditors", IEEE Spectrum, Institute of Electrical & Electronics Engineers, August, 1992.

(5) "Formal Design Review", Corporate Reliability Program Manual, Westinghouse Electric Corporation, February, 1964.

(6) "Design Review Guidelines", Westinghouse Electric Corporation, June, 1968.

(7) "Failure Mode & Effect Analysis", Engineering Manual Section 1310, Westinghouse Plant Apparatus Division , June, 1983.

(8) Bell, Daniel et al, "Using Causal Reasoning for Automated Failure Mode & Effect Analysis 9FMEA)", Proceedings,Annual Reliability & Maintainability Symposium, Las Vegas, Nevada, January,1992.

(9) "Fault Tree Analysis Procedure", Document No. 89051-00-PP-06, Medrad, Inc., November,1989.

(10) Hudson, Bertram J., "Preproduction Quality Assurance--Quality Intrinsic to Design", Medical Design & Material, April, 1991.

Michael P. Tulay,[1] Frank J. Yurich[2]

A SAMPLING PLAN FOR COMMERCIAL PRODUCT TESTING AND VERIFICATION ACTIVITIES

REFERENCE: Tulay, Michael P., and Yurich, Frank J., "**A Sampling Plan for Commercial Product Testing and Verification Activities,**" Quality and Statistics: Total Quality Management, ASTM STP 1209, Milton J. Kowalewski, Jr., Ed., American Society for Testing and Materials, Philadelphia, 1994.

Electric Power Research Institute, "Guidelines for the Utilization of Sampling Plans for Commercial Grade Item Acceptance (NCIG-19), 1992.

ABSTRACT: Acceptance sampling of incoming items is a common activity conducted by purchasers in many different industries. In the commercial nuclear industry, procurements of replacement parts for operating plants typically involve small quantities purchased in isolated lots. This paper discusses a newly issued guideline, EPRI NP-7218, prepared specifically for acceptance sampling of commercial grade items for safety related use. Unlike most sampling publications, a major emphasis is placed on the technical selection factors which should be considered before selecting a sampling plan.

KEYWORDS: Sampling, Product Acceptability, Verification, Nuclear, Lot Size, Destructive Testing, Acceptance Criteria

INTRODUCTION

EPRI NP-5652, "Guideline for the Utilization of Commercial Grade Items in Nuclear Safety Applications" was published in June 1988 by the Electric Power Research Institute. The preparation of this guideline was sponsored by a nuclear industry group called NCIG. The guideline provides guidance to utilities and suppliers on how to accept (dedicate) commercial grade items for nuclear safety related use. As a result of a Nuclear Management and Resources Council initiative, nuclear utilities were committed to implement EPRI NP-5652 by January, 1990.

The acceptance process outlined in EPRI NP-5652 essentially requires the selection and verification of a sufficient number of critical characteristics to provide reasonable assurance the item received is the item specified in the procurement document. Reasonable assurnace is defined as a justifiable level of confidence based on objective and measurable facts, actions, or observations which infer adequacy.

These verifications can be accomplished using one or more of the four acceptance methods discussed in the guideline. In the case

[1]Professional Engineer, Project Manager, Gilbert/Commonwealth, Oak Ridge, TN 37830

[2]Senior Program Manager, Gilbert/Commonwealth, Oak Ridge, TN 37830

of Acceptance Method 1, Special Tests and Inspections, and Method 3, Source Verification, the degree of sampling necessary to provide reasonable assurance a lot is acceptable is an important decision.

In 1990, NCIG authorized the preparation of a companion guideline to EPRI NP-5652 to provide specific guidance on commercial grade item acceptance sampling. EPRI-7218 "Guideline for the Utilization of Sampling Plans for Commercial Grade Item Acceptance", was prepared by NCIG and issued by the Electric Power Research Institute in mid-1991.

PHILOSOPHY OF GUIDELINE

The approach of the guideline has some significant differences from that of other acceptance sampling plans.

Most published plans do not distinguish between acceptance by a product manufacturer as opposed to acceptance by a purchaser. The product manufacturer is responsible for establishing the necessary controls to assure a manufactured item meets the producer's published specifications. These controls can include the use of sampling techniques to 100 percent verification when performing inprocess/final tests and inspections.

For commercial grade item dedication, the purchaser's responsibility is to perform a secondary verification to assure the supplier has fulfilled its contractual commitment. In addition to the acceptance process, there are other activities performed by utilities both prior to and after the acceptance process that provide added assurance only acceptable items are installed in a nuclear power plant. Because of these additional considerations, the degree of acceptance sampling suggested by the guideline is based on reasonableness and not statistical certainty.

When selecting a sampling plan, NP-7218 identifies a number of factors that should be considered before arriving at a sampling plan decision. Proper consideration of the selection factors will assure the appropriate sampling plan is chosen.

PROCESS

The following are the key steps in the acceptance sampling process described in the guideline:

1. Select the items's critical characteristics.
2. Select the acceptance method.
3. Form the lots to be sampled.
4. Select the sampling plans.
5. Implement the sampling plans.
6. Document and trend results.

The first two steps are essentially accomplished as part of the EPRI NP-5652 acceptance process. (NP-7218 emphasizes that acceptance sampling is really only applicable for Acceptance Methods 1 or 3.) Steps 3 and 4 will be the emphasis for this paper. NP-7218 should be consulted for more information on Steps 5 and 6.

LOT FORMATION

Lot formation is an extremely important part of any acceptance sampling process. Intelligent upfront planning can justify smaller sample sizes if it can be demonstrated the lot is homogenous. The

degree of lot homogeneity is a major qualitative input in determining the appropriate sampling plan to utilize.

The guideline addresses the four following types of lot formations:

Production Traceability--These lots provide specific traceability to a product manufacturer's heat, production lot, or batch number. This type of lot formation provides the highest confidence that items in a lot are homogenous.

Line Item/Single Product Manufacturer--These lots are traceable to a specific purchase order line item and the lots are from a single product order manufacturer provided directly or through a distributor. This type of lot formation provides some evidence of homogeneity because the same product manufacturer has produced the item. A logical assumption is similar type product controls should be exercised on all production runs.

Line Item/Multiple Product Manufacturing--These lots are traceable to a specific purchase order line item but either different product manufacturers may have produced the item or product manufacturer traceability does not exist. These lots are typically provided by distributors. Greater risks exist in assuming homogeneity because variations in product conformance can exist between different product manufacturers even though they have provided like items. Greater assurance through sampling is normally necessary for these types of lot formation.

Multiple Line Items/Single Product Manufacturer--These lots are composed of multiple purchase order line items from a single product manufacturer. When the line items have one or more similar critical characteristics, the overall sample size to verify a similar critical characteristic can be reduced by grouping the line items rather than treating each line item as an individual lot.

SELECTION FACTORS

A key feature of the guideline is the recognition that there are many different factors that should be considered before selecting a specific sampling plan. Certain factors may indicate the need for more assurance while other factors may indicate that less assurance is needed. Specific selection factors to be considered when deciding on a sampling plan follow:

Product/Supplier Selection Factors

1. Acceptance history of supplier's products.
2. Formed lot (indication of degree of homogeneity).
3. Item performance history.
4. Complexity of the item.
5. Applicability of industry standards to the item.

Testing or Inspection Factors

1. Acceptance method chosen.
2. Whether verification technique is nondestructive or destructive.
3. Number of other characteristics being verified.
4. Cost effectiveness of the test or inspection.
5. Correlation between nondestructive and destructive tests.

The facts associated with each relevant selection factor are identified or investigated. Once the information on all the relevant

selection factors is collected, engineering judgment is used to select the appropriate acceptance sampling plan that will give sufficient confidence that a conforming lot is received.

NONDESTRUCTIVE TEST AND INSPECTION SAMPLING PLANS

The guideline provides a recommended set of three sampling plans when nondestructive tests and inspections are required. A normal, a reduced, and a tightened sampling plan are provided. The plan selected is determined after an evaluation of the selection factors. The reduced plan requires, on the average, .5 as much inspection as the normal plan. The tightened plan, on the average, requires 1.5 times as much the inspection as the normal plan.

Normal Sampling Plan--When selecting a plan, the normal plan will be considered if:

1. It is expected the lot will be acceptable based upon available knowledge of the product manufacturer or distributor has consistently had a satisfactory product history.
2. The lot is expected to have a sufficient homogeneity that a randomly selected sample will be representative of the lot

Reduced Sampling Plan--In many situations, there will be factors that support the use of a reduced sampling plan. Factors that may justify the use of the reduced sampling plan for a specific critical characteristic include:

1. Acceptance trending provides objective evidence that the product manufacturer or distributor has consistently had a satisfactory product history.
2. Lot formation is established based on a product manufacturer's heat number, production lot number, or batch number.
3. Multiple critical characteristics are being verified on items in the formed lot produced by a single product manufacturer.
4. Satisfactory item performance history exists based on documented evidence obtained from either internal (e.g. maintenance) or external (e.g. INPO NPRDS) sources.
5. The item is a standardized product manufactured to a national standard.
6. The cost effectiveness of the test/inspection is low. The guideline defines cost effectiveness as the value of information provided by a specific sampling plan compared to the sum of the (1) sample costs and (2) test or inspection costs.
7. The item is simple.

Tightened Sampling Plan--In other situations, certain factors may support the use of the tightened sampling plan. Factors that may justify the use of the tightened plan for a specific critical characteristic follow:

1. Based upon available knowledge of the product manufacturer, distributor, or item, there is concern that the lot is unacceptable. An example could be the rejection of products from a product manufacturer on previous shipments.
2. The lot consists of like items from multiple or unknown product manufacturers.
3. The homogeneity of the lot needs to be assessed to justify smaller sample sizes for other critical characteristics.
4. The item is not produced to a national standard.
5. The cost effectiveness of the inspection/test is high.

Obviously for a given critical characteristic, different

selection factors may point to the use of different sampling plans. Engineering judgment is used to decide on the appropriate sampling plan. Acceptance trending and collection of supplier/product information should be utilized to determine if a different sampling plan should be used on future procurements or similar items.

An excerpt from the Recommended Set of Nondestructive Test and Inspection Sampling Plans is shown below (Table 1).

TABLE 1--Sampling Plan Excerpt

NORMAL		REDUCED		TIGHTENED	
Lot Size	Sample Size	Lot Size	Sample Size	Lot Size	Sample Size
26	7	26	4	26	12
27	7	27	4	27	12
28	7	28	4	28	12
29	8	29	4	29	12
30	8	30	4	30	12
31	8	31	4	31	12
32	8	32	4	32	13
33	9	33	4	33	13
34	9	34	4	34	13
35	9	35	4	35	13
36	9	36	4	36	13
37	9	37	4	37	13
38	9	38	4	38	13
39	9	39	4	39	14
40	9	40	4	40	14
41	9	41	4	41	14
42	10	42	5	42	14
43	10	43	5	43	14
44	10	44	5	44	14
45	10	45	5	45	14
46	10	46	5	46	14
47	10	47	5	47	15
48	10	48	5	48	15
49	10	49	5	49	15
50	10	50	5	50	15

SAMPLE SIZES FOR DESTRUCTIVE TESTS

Unlike many other sampling publications, EPRI NP-7218 recognizes that different sampling approaches are needed for destructive tests and inspections. The guideline emphasizes that the use of small representative sample sizes is an accepted practice used in both material testing standards and equipment qualification testing.

The guideline, however, also emphasizes that there should be sufficient justification to permit the use of the small sample sizes. The selection factors should again be used to determine if a small sample size is permissible. A number of upfront activities can be performed to provide technical justification for small sample sizes.

Some of the selection factors that should be considered to justify a small sample size when a destructive test or inspection is necessary follows:

1. Lot formation based on a product manufacturer's heat number, production lot number, or batch number. This type of lot formation essentially assures the lot is homogenous and sample results will be representative of the lot. When this type of lot formation exists,

only one destructive test or inspection sample is considered necessary.

When destructive tests/inspections are required, upfront planning should be performed to determine if production traceability can be obtained.

2. If the supplier has a record of providing a consistently conforming product, a small sample size can be justified.

3. If the lot is from a single product manufacturer, the successful verification of other nondestructive critical characteristics provides additional confidence in the destructive test or inspection results from a small sample.

4. If there is a correlation between a nondestructive test and destructive test. Where a correlation exists, successful results from testing the nondestructive test characteristic can justify only a small sample size for the destructive test.

5. A satisfactory item performance history often provides evidence the supplier has been providing items meeting the destructive test or inspection acceptance requirements.

6. The item is produced to a national standard which specifies the critical characteristic's acceptance requirements.

7. The cost effectiveness of the test/inspection considering the consequences if a defect is not detected, it low.

Because of the necessity to consider the selection factors, a specific destructive test/inspection sampling plan is not provided. The guideline emphasizes the need to justify the basis for the small sample size selected. Special research and upfront efforts should be performed to provide the proper technical justification for the sample size specified.

ACCEPTANCE CRITERIA FOR THE SAMPLING PLANS

A sampled item is classified as defective if one or more critical characteristics do not meet the established acceptance criteria. The guideline specifies zero acceptance numbers for both the nondestructive and destructive acceptance sampling plans. The lot acceptance basis is therefore as follows:

1. Accept the lot if the sample has no defectives.
2. Reject the lot if the sample has one or more defectives.

The guideline recognizes that a "rejected" lot is more correctly a nonconforming lot. Several actions can be taken by the purchaser that may ultimately result in items within the lot being accepted. These actions can include sorting of the lot or an engineering disposition to use-as-is.

CONCLUSION

EPRI NP-7218 should be a valuable tool for utility procurement organizations because industry use of this guideline will result in intelligent commercial grade item acceptance sampling and consistency in sampling approaches. The fundamental principles of this guideline may also be well suited for sampling activities in other industries. Proper consideration of the selection factors will improve the effectiveness of sampling decisions from both a technical and cost standpoint.

John Mandel[1]

ANALYZING INTERLABORATORY DATA ACCORDING TO ASTM STANDARD E691

REFERENCE: Mandel J., **"Analyzing Interlaboratory Data According to ASTM Standard E691,"** Quality and Statistics: Total Quality Management, ASTM STP 1209, Milton J. Kowalewski, Jr., Ed., American Society for Testing and Materials, Philadelphia, PA 1994.

ABSTRACT: ASTM Standard Practice E691-92 deals with the design and the analysis of the data of an interlaboratory study of a test method. This paper explains the procedure of E691 and goes beyond it in a discussion of the linear model analysis of interlaboratory test data. The explanations are illustrated in terms of a real set of interlaboratory data.

KEYWORDS: Interlaboratory testing, h-graph, k-graph, Linear model, Outliers, Proficiency testing, repeatability, reproducibility

There are two types of interlaboratory studies. The first type is concerned with the monitoring of laboratories. It is often referred to as "proficiency testing". Its objective is to attempt to achieve good agreement between the results obtained by different laboratories, by allowing each laboratory to compare its results with those obtained by other laboratories and take remedial action, if necessary. The second type is concerned, not so much with the laboratories, as with the method of measurement. Standardization organizations, such as ASTM, desire to associate, with each method of measurement, measures of repeatability and of reproducibility. The former refers to repeated measurements of the same characteristic in the same laboratory; the latter to the agreement between different laboratories. E691 [1] is a standard dealing primarily with the second type of interlaboratory testing, that is, with the evaluation of methods of measurement in terms of repeatability and reproducibility.

Of course, even the second type of interlaboratory testing involves the collaboration of different laboratories, but these laboratories, which must be qualified, are not the main objective. They are only tools for the measurement of repeatability and reproducibility.

Design of an Interlaboratory Study

A typical design is illustrated in Table 1, which presents the results of a study of an analytical method for the determination of oxygen in steel [2]. Twelve samples of steel, ranging in oxygen content from 5 to about 55 parts per million (ppm) were tested by sixteen laboratories. Each laboratory made 3 replicate determinations on each

[1]Guest Scientist , Chemical Science and Technology Laboratory, National Institute of Standards and Technology, Gaithersburg, Maryland 20899

Table 1

Oxygen in Steel

	Level											
Lab	A	B	C	D	E	F	G	H	I	J	K	L
1	4.5	7.2	7.3	11.7	13.1	13.8	21.8	24.1	35.4	27.4	50.0	55.4
	4.8	6.8	8.2	10.6	13.2	14.7	22.6	26.6	35.7	25.8	47.5	53.6
	4.3	7.4	7.4	11.0	13.1	13.8	22.1	24.4	36.3	25.1	47.1	54.7
2	4.2	4.8	6.4	10.8	11.2	14.2	21.6	20.4	34.6	38.8	44.0	51.2
	3.8	4.6	7.2	11.6	11.4	13.6	21.8	17.8	34.6	32.2	43.4	52.0
	3.4	5.4	6.6	9.8	11.8	14.2	19.0	22.0	35.6	35.2	48.8	54.6
3	4.3	7.2	14.1	14.8	14.8	15.8	28.8	28.6	41.7	65.5	56.0	72.2
	5.4	6.6	11.2	16.0	15.0	17.5	30.0	29.6	46.3	57.4	57.8	74.0
	5.5	6.7	14.3	15.7	14.1	18.4	31.1	26.3	44.0	58.2	57.5	76.1
4	5.1	5.0	5.8	10.8	13.8	12.3	21.3	26.4	34.0	17.7	45.3	51.6
	5.1	5.6	6.0	10.5	14.8	12.6	21.7	26.1	34.5	24.0	44.4	51.0
	4.2	6.1	5.9	9.6	13.6	12.1	21.1	26.5	34.7	33.0	45.9	49.8
5	5.2	6.0	6.8	11.8	12.7	15.4	21.3	21.8	36.7	52.5	46.2	56.0
	5.9	6.3	4.5	11.2	10.8	14.6	20.5	18.7	37.6	46.8	50.6	55.3
	5.9	4.8	4.4	10.5	11.7	15.7	19.8	20.1	39.3	58.8	42.2	54.9
6	3.0	3.2	4.9	8.4	8.7	13.2	19.0	25.2	34.6	39.0	41.9	48.6
	4.6	3.6	4.5	9.7	10.8	13.1	20.8	21.8	33.9	33.7	41.0	49.4
	4.0	4.4	5.9	8.7	8.6	11.0	20.3	21.1	33.0	39.5	43.5	50.9
7	6.6	4.6	7.3	16.0	11.0	26.0	14.0	22.0	41.0	50.0	46.0	51.0
	7.5	3.8	6.5	12.0	12.0	13.0	19.0	24.0	37.0	27.0	42.0	55.0
	4.6	3.9	6.5	12.0	11.0	13.0	16.5	22.0	37.0	34.0	41.0	56.0
8	3.6	2.3	5.1	8.0	10.1	11.8	18.9	17.0	39.7	49.3	45.4	50.6
	5.0	2.0	5.1	10.1	11.2	11.5	21.4	22.3	37.2	18.3	45.2	51.6
	1.3	0.0	9.3	7.6	12.1	11.6	22.8	23.4	34.3	26.4	39.8	53.7
9	6.0	4.0	9.0	14.0	13.0	12.0	17.0	44.0	32.0	26.0	21.0	52.0
	5.0	6.0	7.0	12.0	14.0	12.0	19.0	44.0	32.0	32.0	21.0	50.0
	7.0	6.0	10.0	16.0	13.0	5.0	15.0	42.0	32.0	34.0	21.0	53.0
10	3.2	7.0	8.6	11.5	17.4	13.7	26.8	32.9	37.8	19.4	47.0	56.3
	5.3	5.6	6.4	18.1	13.6	16.7	20.6	26.8	39.3	21.9	48.9	59.9
	6.3	6.1	8.4	13.9	22.7	14.1	26.5	26.8	40.4	25.3	46.0	54.3
11	6.9	8.8	5.9	14.3	13.0	13.9	22.7	24.6	34.5	26.0	39.2	55.3
	5.4	9.1	7.7	12.2	12.0	14.8	21.4	24.8	35.5	28.2	43.1	56.4
	5.2	9.9	8.1	12.3	12.7	16.7	31.7	26.3	35.3	25.4	36.4	54.4
12	3.0	5.0	7.0	12.0	12.0	13.0	20.0	20.0	37.0	50.0	47.0	57.0
	3.0	4.0	6.0	10.0	13.0	12.0	21.0	24.0	35.0	42.0	48.0	61.0
	4.0	3.0	7.0	10.0	12.0	12.0	21.0	20.0	36.0	60.0	46.0	59.0
13	4.0	4.0	4.0	10.0	6.0	8.0	21.0	22.0	34.0	36.0	43.0	53.0
	3.0	5.0	5.0	9.0	7.0	12.0	21.0	24.0	37.0	36.0	47.0	59.0
	6.0	6.0	8.0	10.0	9.0	13.0	21.0	25.0	35.0	32.0	45.0	59.0
14	7.0	10.0	13.0	15.0	19.0	16.0	27.0	24.0	34.0	50.0	43.0	47.0
	9.0	10.0	16.0	15.0	15.0	16.0	24.0	28.0	32.0	52.0	41.0	50.0
	9.0	10.0	18.0	18.0	16.0	17.0	24.0	23.0	32.0	47.0	41.0	49.0
15	15.0	2.6	7.6	8.2	10.8	13.1	19.5	25.5	34.5	43.4	44.9	54.9
	3.0	3.2	5.7	8.6	11.6	14.4	19.5	23.2	36.9	37.0	44.9	55.1
	2.1	2.8	5.8	12.0	11.2	16.4	20.3	22.6	33.2	32.7	43.8	56.0
16	4.0	4.0	14.0	14.0	13.0	16.0	21.0	19.0	28.0	30.0	35.0	56.0
	4.0	8.0	13.0	10.0	14.0	11.0	24.0	22.0	30.0	29.0	31.0	55.0
	7.0	8.0	12.0	10.0	11.0	13.0	22.0	24.0	31.0	28.0	34.0	50.0

Table 2

Calculations for Level A (j=1)

Lab (i)	Replicates			Average (y_{i1})	Std. Dev. (s_{i1})	h	k
1	4.5	4.8	4.3	4.533	0.252	-0.432	0.120
2	4.2	3.8	3.4	3.800	0.400	-0.983	0.190
3	4.3	5.4	5.5	5.067	0.666	-0.031	0.317
4	5.1	5.1	4.2	4.800	0.520	-0.232	0.247
5	5.2	5.9	5.9	5.667	0.404	0.419	0.192
6	3.0	4.6	4.0	3.867	0.808	-0.933	0.385
7	6.6	7.5	4.6	6.233	1.484	0.845	0.706
8	3.6	5.0	1.3	3.300	1.868	-1.359	0.889
9	6.0	5.0	7.0	6.000	1.000	0.670	0.476
10	3.2	5.3	6.3	4.933	1.582	-0.131	0.753
11	6.9	5.4	5.2	5.833	0.929	0.545	0.442
12	3.0	3.0	4.0	3.333	0.577	-1.334	0.275
13	4.0	3.0	6.0	4.333	1.528	-0.582	0.727
14	7.0	9.0	9.0	8.333	1.155	2.423	0.549
15	15.0	3.0	2.1	6.700	7.202	1.196	3.427
16	4.0	4.0	7.0	5.000	1.732	-0.081	0.824
Material Average			x_1 =	5.1083	2.102 (pooled) = s_r		
Std. Dev.			sL =	1.3311			

Level Average = 5.1083
s_L = 1.331
s_r = 2.102

$$s_R = \sqrt{(2.102)^2 + \left[(1.3311)^2 - \frac{(2.102)^2}{3}\right]} = 2.172$$

Table 3

Precision Parameters for Data of Table 1

Level	Average	s_r	s_R
A	5.108	2.102	2.172
B	5.575	0.906	2.303
C	8.008	1.374	3.333
D	11.771	1.598	2.664
E	12.554	1.489	2.827
F	13.848	2.478	2.986
G	21.754	2.079	3.690
H	24.931	2.054	5.817
I	35.710	1.425	3.473
J	36.644	6.550	12.397
K	43.140	2.145	7.857
L	55.038	1.940	6.010

steel. The large range of oxygen content values for the twelve samples is deliberate. The steels are arranged in increasing order of oxygen content.

In general, the number of laboratories is represented by p, the number of levels by q and the number of replicates by n. In our example, p = 16, q = 12, n = 3.

In Table 1 we see that the design is simply a rectangular array of laboratories (rows) and materials or levels (columns) with n-fold replication in each cell of the p x q table. It is important, in interlaboratory studies, to cover the entire range of material levels (oxygen content in this case) for which the test method is intended.

In what follows, the materials have been rearranged in increasing order of their average values over all laboratories.

Analysis of the Data: Phase 1

The analysis of a table of data such as that of Table 1 involves three phrases:

(1) the determination of s_r (standard deviation of repeatability) and of s_R (standard deviation of reproducibility), as functions of the oxygen levels in materials.

(2) examination for outliers.

(3) the examination of the quantitative interrelationships between results from different laboratories. This may be referred to as the "structure" of the data.

ASTM standard E691 and its companion software package [3] provide for (1) and (2), but not (3). This paper deals with all three phases.

E691 Analysis: Phase 1

According to E691, we first look at each level separately, and then at the global picture of all levels combined. We illustrate the procedure for the first material level of Table 1 (level A or j=1). Table 2 shows the calculations, which are very simple. For each laboratory, the average y_{ij} (for our illustration, j = 1) and the standard deviation of replicates, s_{ij}, are calculated. The average of averages is represented by x_j. The standard deviations are squared, the squares are averaged, and the square root of this average is s_r, the pooled standard deviation of repeatability. The standard deviation among the y_{ij} is represented by s_L. The standard deviation of reproducibility, s_R, is calculated by the formula:

$$s_R = \sqrt{s_r^2 + \left(s_L^2 - \frac{s_r^2}{n}\right)}$$

The quantity in parentheses, inside the square root, is occasionally negative. In that case, it is replaced by zero. The columns labeled h and k will be discussed in the section dealing with Phase 2 (checking for outliers).

Table 3 is a summary of the x_j, the s_r-values and the s_R-values for all levels. Fig. 1 is a plot of s_r and s_R versus x_j. This completes phase 1 of the analysis. Phase 1 may occasionally be invalid, because of the presence of outliers.

E691 Analysis: Phase 2

Much has been written about tests for outliers. A widely advocated procedure is to make statistical tests of significance on the data at each level separately and to discard observations that show significance in these tests. Rejection procedures are used in ISO Standard 5725 [4] in Youden and Steiner [5] and, as a result of the latter, in numerous publications of the Association of Official Analytical Chemists (AOAC) [6]. Tests for the rejection of outliers are

statistically and intiutively invalid, because they don't allow us to look at all the data before rejecting some of them.

A more reasonable approach is used in E691 in the form of h graphs and k graphs. The h-graph examines the relation between laboratories in the following way. Let

$$h_{ij} = \frac{y_{ij} - x_j}{(s_L)_j} \tag{1}$$

This quantity measures the deviation of the results y_{ij} in cell i,j from the "consensus-value" x_j for the level in question, standardized by dividing by s_L for that level. The pxq values of h_{ij} (listed in Table 2 for j=1) are assembled by laboratories and plotted as shown in Fig. 2. The figure is highly revealing. We see for example that laboratory #3 obtained systematically higher results than the other laboratories for all levels; laboratories 6 and 8 tend to be low; laboratory 14 is high with three exceptions; laboratory 9 goes from very low values to very high values; etc.

Tests of significance would have eliminated some, but not necessarily all, results of some of the laboratories, but this would have been totally unreasonable. The subject-matter specialist (not the statistician) must decide whether further action should be taken with respect to systematically different laboratories. It is of the utmost importance to realize that action taken in regard to the data (such as "rejection" of outliers) is useless and in fact misleading, unless it is accompanied by appropriate action taken in regard to the source of the data. By rejecting outliers our data become "prettier", but the method of measurement itself is totally unaffected unless we study, and remedy, the source of the outliers. This source may be faulty practice in a laboratory, lack of proper calibration, etc. If these shortcomings are not found and corrected, the mere elimination of outliers contributes nothing to the improvement of measurement precision, and may in fact be viewed as harmfully condoning the status quo.

The k statistic addresses differences, not between laboratories, but rather between replicates in the same laboratory. The formula for k is:

$$k_{ij} = \frac{s_{ij}}{(s_r)_j} \tag{2}$$

It is a positive quantity measuring how large a particular s_{ij} is when compared to its pooled value for the same level. The values of k_{ij} for j=1 are listed in Table 2. Fig. 3 is a plot of all k_{ij}, grouped again by laboratories. This graph also singles out some laboratories as having inconsistent cells. Here again, action should have been taken in terms of what could have occurred in the laboratories, and how to correct it, rather than merely discarding data. If any rejection or modification of data occurs, phase 1 has to be redone on the modified data set.

Analysis of the Data: Phase 3

As mentioned above, E691 deals with phases 1 and 2 only. Nevertheless it is often worth while to go a step further and examine what mathematical model, if any, underlies the data [7].

We first observe that in properly designed interlaboratory studies, the range of levels is generally large, while the differences among laboratories are generally relatively small. It is found that very often the following model applies:

$$y_{ij} = A_i + B_i x_j + error \tag{3}$$

Here x_j is the consensus value for level j (average over all laboratories for level j), y_{ij} is the average of replicates in cell (i,j), and A_i, B_i are constants for laboratory i. This is called a

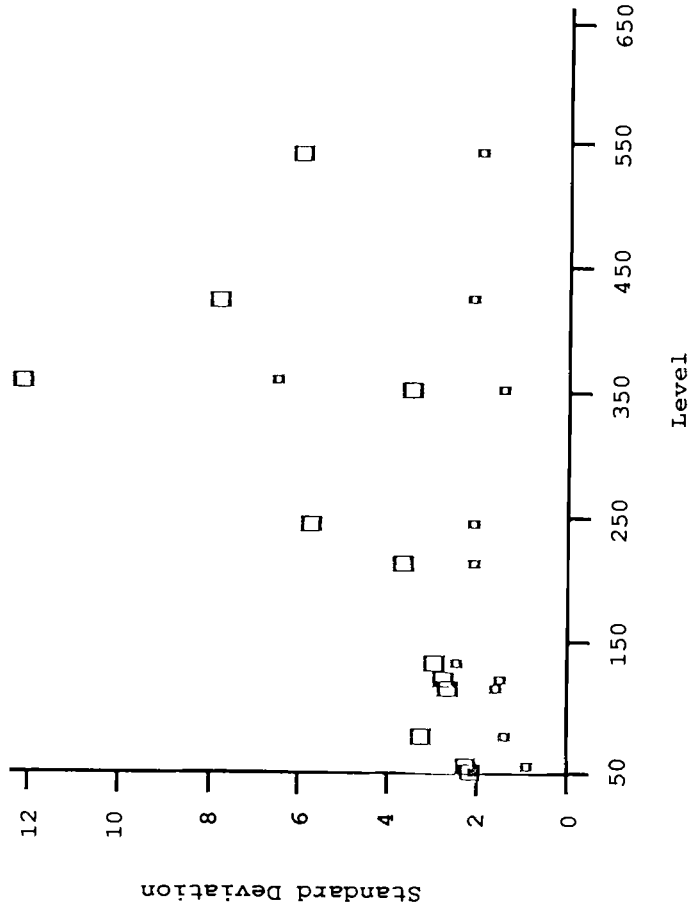


Fig. 1-- s_r and s_R Versus Level

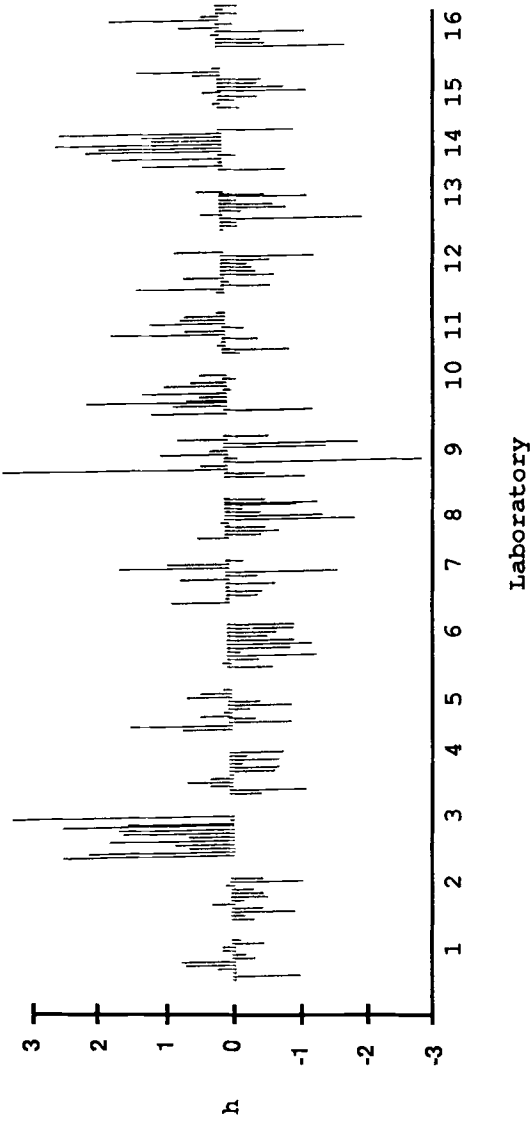


Fig. 2--Graph of h-Values by Laboratories

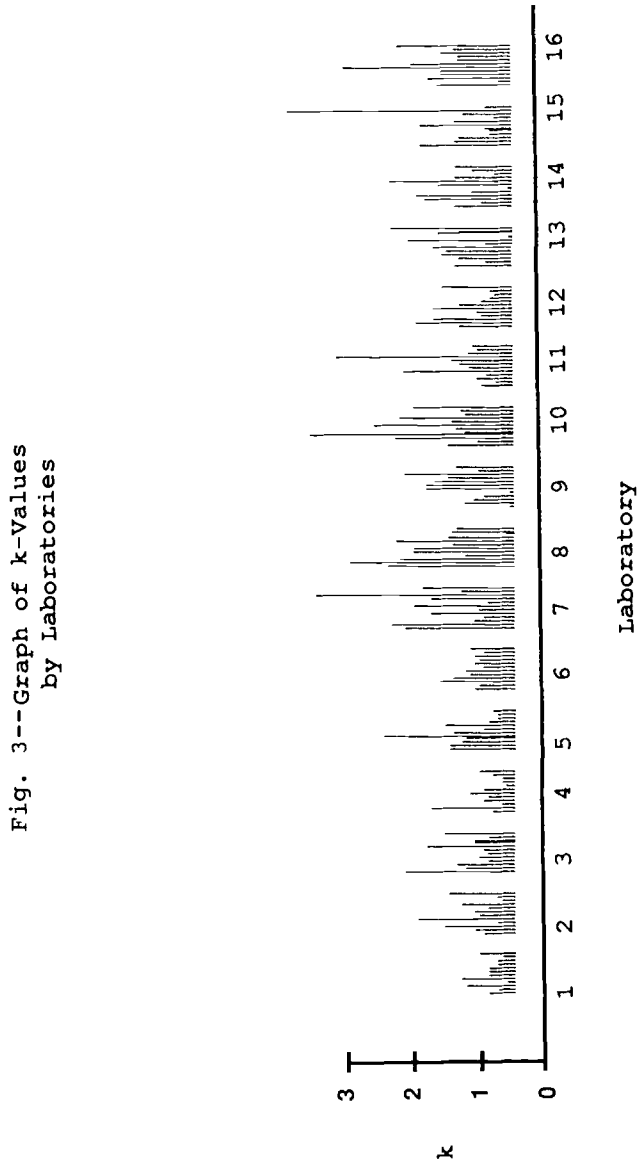


Fig. 3--Graph of k-Values by Laboratories

Table 4

Parameters in Row-Linear Representation of Data of Table 1

Laboratory	Height	Slope ($_b$)	s_e	s_b	$\hat{\rho}$
1	22.4583	0.9766	3.71	.068	.976
2	21.7667	0.9999	1.72	.032	.995
3	30.5139	1.3855	4.25	.078	.984
4	21.3306	0.9258	3.63	.067	.975
5	23.9806	1.1183	5.01	.092	.968
6	20.7639	0.9874	1.54	.028	.996
7	22.5500	0.9970	2.47	.045	.990
8	20.7000	1.0132	2.01	.037	.993
9	21.0556	0.7761	8.66	.159	.839
10	23.7639	0.9638	5.37	.099	.951
11	22.5028	0.8971	3.37	.062	.977
12	23.6667	1.1706	3.98	.073	.981
13	21.6667	1.0787	1.57	.029	.996
14	25.4722	0.8454	4.56	.084	.954
15	22.2778	1.0390	1.44	.026	.997
16	20.9722	0.8250	3.14	.058	.976
Std. Dev. of Slopes		0.1466			

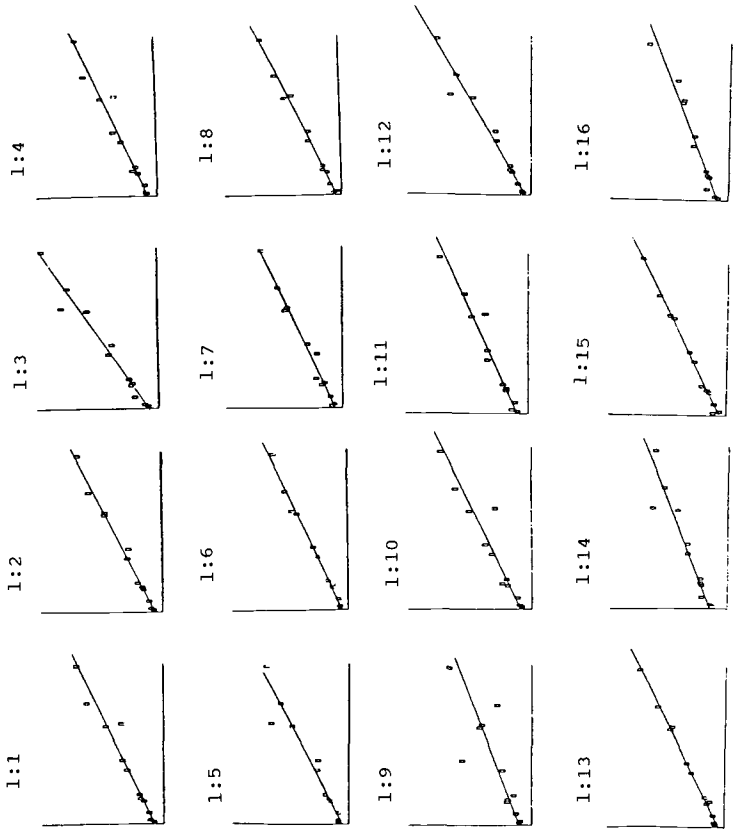


Fig. 4--Individual Graphs for All Laboratories
Laboratory Value Versus Average of All Laboratories

"row-linear" model, because geometrically it means that the cell-averages in each row (i) are linearly related to the corresponding column averages x_j, (x_j represents the level of the j th material). In Fig. 4, plots are shown for the values of each of the laboratories, when plotted against x_j. Clearly the values for each laboratory fall essentially along a straight line, but the lines representing the different laboratories are not parallel.

This is apparent in Table 4, which is a display of the estimated parameters in a modified form of Eq. (3). If $\bar{x}$ is the average of all x_j, then Eq. (3) can be written in the form:

$$y_{ij} = H_i + b_i (x_j - \bar{x}) + d_{ij}$$

where H_i, the "height", is the value of y_{ij}, at a value of x_j equal to $\bar{x}$ (the ordinate at that value of x). It can be shown that H_i is the average of Y_{ij} over j. The quantity s_e in Table 4 is the standard deviation of the d_{ij}, i.e., of the scatter of the points around the fitted straight line. The symbol s_b denotes the standard error of the slope b, and is a measure of its uncertainty. Finally, $\hat{\rho}_i$ is an estimate of the correlation coefficient between x_j and y_{ij}, for a given laboratory i.

At the bottom of the column of slopes, the standard deviation between these values is given. It is larger, with one exception, than all s_b values. This suggests that the slopes differ from one other, indicating that a simple "additive" model (the usual Anova model) is not applicable. The exception occurs for laboratory 9, for which the value of $\hat{\rho}$ is also the smallest. Fig. 4 clearly shows that laboratory 9 exhibits greater scatter than all other laboratories. It should have been investigated in the course of this round robin. Fig. 4 also shows a number of outlying points: such points are seen to occur in about half of the laboratories. They should have been investigated.

We see that Phase 3 is extremely useful for the complete elucidation of the data set, and is therefore a worthwhile addition to phases 1 and 2.

CONCLUSIONS

We have shown that the analysis of interlaboratory data can be carried out in a straight-forward manner by means of simple calculations and simple graphs. The h and k statistics, when properly plotted, are extremely useful in evaluating the behavior of the participating laboratories and for the detection of outlying observations or laboratories. On the basis of these graphs, the committee in charge of the study should decide what further action, if any, is necessary. Phase 3, which is not part of standard E691, is extremely useful for a better understanding of the data set.

REFERENCES

[1] ASTM Designation E691-92, Standard Practice for Conducting an Interlaboratory Study to Determine the Precision of a Test Method, American Society for Testing and Materials, Philadelphia, 1992.

[2] Green, W.B., Jr., Diamondstone, B.I., and Hoo, J.J.C., "Current Status of Round-Robin Testing of Oxygen Content in Bearing Steels," Effect of Steel Manufacturing Processes on the Quality of Bearing Steels, ASTM STP 987, J.J. Hoo, Ed., American Society for Testing and Materials, Philadelphia, 1988, pp. 198-210.

[3] ASTM Interlaboratory Data Analysis Software E691, User's Guide, American Society for Testing and Materials (1992).

[4] International Organization for Standardization, Draft International Standard ISO/DIS 5725-A, Accuracy (Trueness and Precision) of Measurement Methods and Results, Part 2. A Basic Method for the Determination of Repeatability and Reproducibility of a Standard Test Method, International Organization for Standardization (ISO), 1990.

[5] Youden, W.J., and Steiner, E.H., Statistical Manual of the Association of Official Analytical Chemists, Association of Official Analytical Chemists, Washington, DC.

[6] Association of Official Analytical Chemists, PO Box 540, Benjamin Franklin Station, Washington, DC 20044.

[6] Mandel, John, Evaluation and Control of Measurements, Marcel Dekker, Inc., 270 Madison Avenue, New York, New York 10016, 1991.

Barry H. Levine[1], Duncan C. McCune[2]

ALL YOU EVER WANTED TO KNOW ABOUT CONTROL CHARTS: BUT NEVER ASKED

REFERENCE: Levine, B. H., and McCune, D. C., **"All You Ever Wanted To Know About Control Charts: But Never Asked"**, Quality and Statistics: Total Quality Management, ASTM STP 1209, Milton J. Kowalewski, Jr., Ed., American Society for Testing and Materials, Philadelphia PA, 1994

ABSTRACT: Statistical Control Charts developed over 60 years ago have seen many applications in process analysis and control in the past decade. To some users, the control charts are a recent invention of the Japanese, to many there is a confusion about the underlying assumptions and where and how this analytical tool may be applied. The effect of some of these assumed assumptions for control charts is discussed followed by suggested areas of application.

KEY WORDS: Control Charts, Assumptions for Control Charts, Rational Subgroups, Control Charts for Individuals, Exponentially Weighted Control Charts, Measurement Precision, Short Production Runs.

Historical Introduction

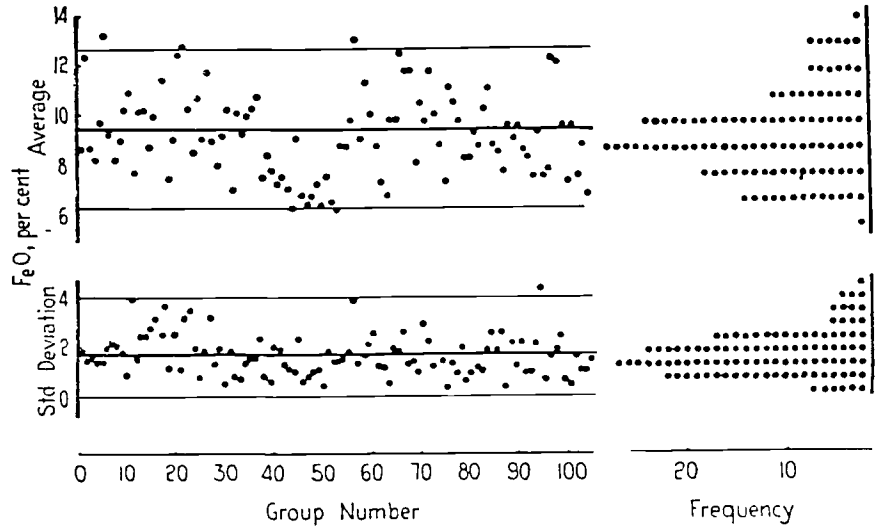
The concept of what we know today as Control Charts was developed by Walter Shewhart at Bell Telephone in the 1920's and 1930's and first appeared in 1931 [1]. As W. Edwards Deming states in his Dedication in the 1980 republished edition of Shewhart work,

> *"To Shewhart, quality control meant every activity and every technique that can contribute to better living, in a material sense, through economy in manufacture..... Economic manufacture requires achievements of statistical control in the process and statistical control of the measurements." [2]*

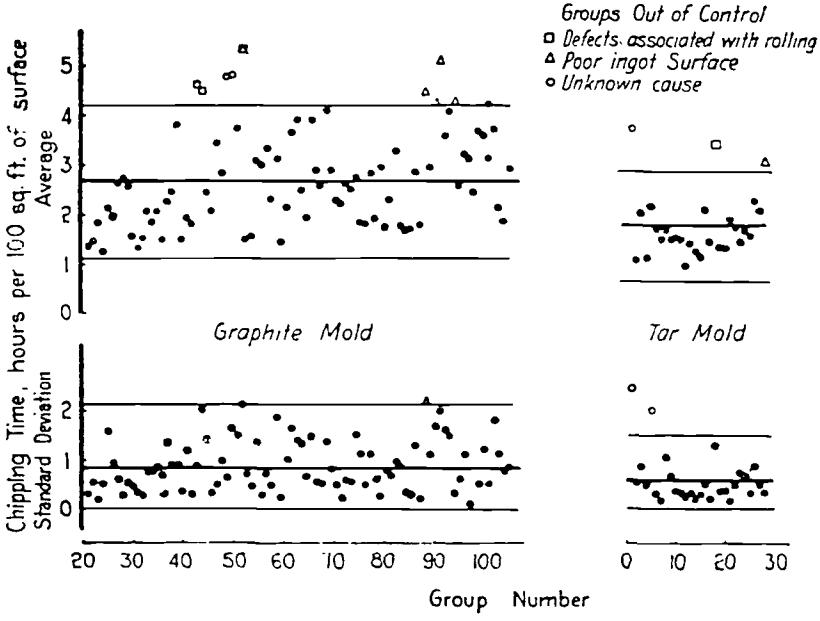
The concept of the Control Chart was not something which US industry learned from the Japanese as some of the newspapers and magazines would have one believe. The Japanese had been introduced to these concepts by W. E. Deming in the 1950's. However, there are examples of Control Charts from Japan dating from WWII. Control charts were used by American industry before World War II. ASTM Proceedings of 1934 contained a paper by W. C. Chancellor on the application of statistical methods, including control charts, to metallurgical problems in a steel plant [3]. See Figure 1 for two examples from this paper. Today, the points on these control charts would be connected to indicate the time sequence in the data so as to better identify patterns in the data. In the authors' opinion, there are

[1] Brisch, Birn and Partners, V.P. Quality Management Systems, 1075 SE 17th Annex, Fort Lauderdale, FL 33316

[2] Consultant, Quality/Statistics, 1236 Second Street, Beaver, PA 15009-2594



Control Chart for FeO in Open-Hearth Slag.



-Control Chart for Chipping Time.

Figure 1: Two examples of control charts from 1934.[3]

far too many points plotted on these charts. It is only recently that some users of Control Charts have added the practice of constructing a dot plot at the edge to form a frequency histogram of the data. Note also that the sample standard deviation is used rather than the range for the within subgroup variation chart.

In 1933, ASTM Committee E-11 published STP 15, ASTM Manual on Presentation of Data, to which was added Supplement B on the control chart method in 1935. Supplement B was later incorporated in 1941-1942 in American War Standards on Quality Control (Z1.1, Z1.2 and Z1.3). In 1946 this entire publication became ASTM STP 15C and is now in its sixth edition and known as ASTM Manual 7 [4]. During WWII, many engineers and plant management were instructed in Control Charts by the War Production Board. Harold Freeman of MIT published his text Industrial Statistics [5] in 1942. In England, L. H. C. Tippett published Technological Applications of Statistics [6] in 1952. Also in 1952, Eugene L. Grant published his classical text, Statistical Quality Control [7] which is still in print as a 6th edition. The Statistical Quality Control Handbook [8] written by a committee headed by Bonnie B. Small and published 1956 by Bell Telephone, after being used internally for several years, and is still available.

So much for historical background.

What are the assumptions behind a Control Chart and why does it work ?

Ask many persons about the underlying assumptions for control charts and you will hear some of the following.

1) *The data must follow the "Normal" Distribution.*
   or
   The Control Chart works because of the "Central Limit Theorem."
2) *The limits are always set at ± 3σ.*
3) *The individual observations must be independent.*
4) *The data must be grouped into sets of 5.*
5) *The data must be brought into control prior to using a Control Chart.*
6) *Take no action on an "out of control" point but wait until the next data point is taken and if this is outside the control limits, then take action to correct the process.*

Let us look at each of these 6 "assumptions" to see if they are correct and/or necessary.

(1) Anyone who has read Shewhart's 1931 book will quickly learn that he demonstrated through simulation studies that the Control Chart works even when the basic measurement came from non-normal distributions. The Central Limit Theorem states that the distribution of a sample average from any distribution will tend towards the Normal distribution as the sample size, n, increases. In a normal distribution, 99.7% of the individual measures will be between ± 3σ, 95.5% between ± 2σ and 68.3% will be between ± 1σ. Wheeler & Chambers [9] show that these percentages can be 98 to 100%, 95 to 100%, and 57 to 87% respectively for the distributions of individuals which are far from normal.

Thus, let us not get hung-up on a 99.7% syndrome.

(2) Why do we use ± 3σ as limits ? Shewhart [10] states

".. we usually choose a symmetrical range characterized by

limits

$$\bar{\theta} \pm t\,\sigma_\theta$$

"Experience indicates that t = 3 seems to be an acceptable economic value."

"Hence the method for establishing allowable limits of variation in a statistic θ depends upon theory to furnish the expected value $\bar{\theta}$ and the standard deviation σ_θ of the statistic θ and upon the empirical evidence to justify the choice of limits

$$\bar{\theta} \pm t\,\sigma_\theta$$

Shewhart's studies showed that it was more economic to use the $\pm 3\sigma$ limits than to be making unwarranted corrections or adjustments to the process. If one were to use $\pm 2\sigma$ limits, such unwarranted corrections to the process would be made 1 time in 20. Again to quote Shewhart [11],

"The fact that the criterion which we happen to use has a fine ancestry of highbrow statistical theorems does not justify its use. Such justification must come from empirical evidence that it works."

(3) An examination of most production processes will show some amount of autocorrelation in the so-called independent measures. If two observations are not independent, x_{t+1} may depend upon the previous value, x_t. In this case we say that there is autocorrelation in the x-series. An example of this autocorrelation can be observed in test measurements where the instrument may drift over time. Consecutive readings will be correlated. Data from production processes within the same setup may be correlated while those from differing setups may not be.

The Control Chart process has worked well for over 60 years based on $\pm 3\sigma$ limits. The maker of the Control Chart should be aware however that such autocorrelation may affect the limits. The purpose of the chart is to give insight into the behavior of the process and one should not be overly concerned about the effects of the lack of independence of consecutive measures. As Shewhart said "Such justification must come from empirical evidence that it works." And it does work in those situations where the data are autocorrelated.

(4) The concept that data must be grouped into subsets of 5 probably has led to the failure of more applications of Control Charts than any other reason. Shewhart established the concept of "rational" subgroups so that the random, chance variation would be contained *within subgroups* and the special or assignable cause variation would given by the *between subgroups* variation. For this reason, the measure of variation which is used is the within subgroup variation measured either as the average subgroup range or average subgroup standard deviation. It makes no sense to take one reading per day and then to combine the data into subgroups of 4 or 5 or any other number. Likewise to combine data from two different crews, two different machines, fails to recognize the concept of a "rational" subgroup. In applying a control chart to a laboratory instrument calibration process which is done once a day, it makes little sense to group the data in to weekly averages of 5 or 7 values. Shewhart's original charts used subgroups of 4 and theoreticians can prove that subgroups of 7 give the best response to certain process changes. Only careful consideration of the origin of the data set can determine the best rational subgroup size. In many cases the best subgroup size for a "rational" subgroup is one.

In some applications the use of sub-group averages may not be

appropriate, but this does not prevent the use of a Control Chart. The size of a rational subgroup may be *one*. There are many chart applications for product, process or laboratory measurements where a Control Chart for Individuals, a Moving Average Control Chart, a Cumulative Sum Control Chart or CuSum Chart, or an Exponentially Weighted Moving Average Control Chart, EWMA Control Chart, are applicable. (See Refs: [4], [7], [9] and [14] for further information.) In these particular chart applications it is *essential* that a measure of short-term variation be used to estimate the standard deviation rather than the standard deviation of all the data points. The most common method here is to use the moving range of two consecutive measurements. In this case the range is measured by consecutive pairs (x_1, x_2), (x_2, x_3), (x_3, x_4), etc, Such a range chart is oftentimes identified as MR2 for Moving Range of 2. The use of the standard deviation of all the data will include possible special or assignable cause variation and thus the control limits for the average may be larger than if the within subgroup variation was used.

There can be chart applications to a process where a trend is as important as a jump in the measurement average and it may be advantageous to use a Moving Average Control Chart or CuSum Control Chart in conjunction with or a replacement for the chart for averages. The less familiar EWMA Control Chart is a simple way to construct a Moving Average that gives a decreasing weight to each observation.

In an Exponentially Weighted Moving Average (EWMA),

$$\bar{x}_n = \alpha x_n + (1 - \alpha)\bar{x}_{n-1} \text{ where } 0 < \alpha < 1$$

Thus the new average is just the new observation, x_n, times α and the old average, $\bar{x}_{n-1}$ times $(1 - \alpha)$. There is a useful relationship between the variance of the EWMA average, σ^2_{ewma}, and the variance of the individual values, σ_x^2 .

$$\sigma^2_{ewma} = \sigma_x^2 \quad [\alpha/(2 - \alpha)]$$

or

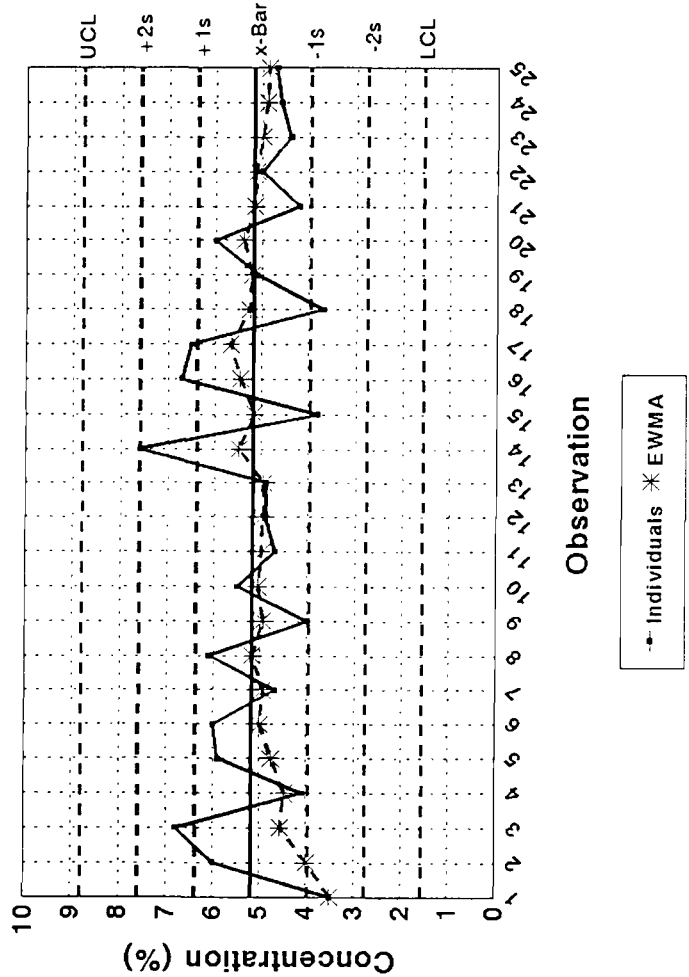
$$\sigma_{ewma} = \sigma_x \sqrt{\frac{\alpha}{2 - \alpha}}$$

If $\alpha = .2$, then the 3σ limits on the EWMA values will be $\pm 1\sigma_x$. It is then advantageous to plot such a EWMA along with the chart for individuals. See Figure 2 for an example. The $\pm 1\sigma$, $\pm 2\sigma$ and $\pm 3\sigma$ limits for "x" are sometimes called the bias, warning and action limits and are shown. The $\pm 1\sigma_x$ line applies to the EWMA trend line as $\pm 3\sigma_{ewma}$ limits. If 2 or 3 consecutive individual "x" values fall between the 2σ and 3σ limits, this is an indication of an increase in the measurement variability.

This EWMA Control Chart procedure can be used also for grouped data where a trend is suspected. An example would be tool wear or acid bath depletion. The only change would be that the subgroup averages are used to construct the EWMA line rather than single x values.

(5) The fifth assumption given above, that the process must be in control prior to using a Control Chart, ignores the primary purpose of the Control Chart. If this purpose of a control chart is to separate the assignable or special cause variation from the random chance variation, then one does not need a chart if there is no special cause variation. The process is always predictable.

Figure 2: Control Chart for Individuals with EWMA Trend Line



(6) The last of our assumptions is incorrect in that it fails to recognize that any data point plotted on the chart will vary at random above and below the process average. If the process average, the center line of the $\bar{x}$-Chart, shifts either upwards or downwards a full 3σ, there is still a 50-50 chance that the next point will still be within the control limits. If the shift is closer to $\pm 1\sigma$, the odds are about 19 out of 20 that the subsequent point will be in-control. Thus failure to take action when a point falls outside either control limit will probably result is the failure to detect an assignable cause of variation.

Thus we may conclude that none of the 6 assumptions are really necessary for the Control Chart to function as Shewhart intended.

Process or Product ?

Many of the early applications of Control Charts were similar to those being made today in that they measured the performance of a product. Even processes under computer control may not be "in-control". While much can be learned about product variation from such Control Charts, it is of much greater value to identify the key processing variables and to control them than charting the product values. Product characteristics may be charted to see if a problem exists and as a means of identifying the key process variables which need to be controlled. A point out-of-control on a product characteristic control chart signifies a change in the process but it may be too late to do much about it. If the goal is as stated by Deming, one of "never ending improvement", then it is only by identification and control of the key characteristics of the key process variables that this can be attained. Control Charts on product characteristics can impress a customer, but will have little effect on the quality of the product produced other than to measure the effect of the key process variables which must be identified and controlled.

In choosing a characteristic to be charted, some knowledge of the process is necessary. For example, to chart the width or deviation in width from an ordered value for slit or edge trimmed product usually will be useless as the slitting knives do not often move. Measuring the width of wide sheets or coils presents a problem of measurement precision unless a square or similar fixture is used to align the tape measure. A Control Chart based on some set-up parameters on the slitter knives would be more fruitful.

Often times, the product to be controlled is the result of two or more processing streams. Examples would be raw material feed belts coming from several storage bins for differing materials and a sample taken off the final belt. Probably only one of the several materials is key for control and the final belt sample may not characterize the real problem even if a true sample can be obtained from the belt. In molding processes there are multiple cavities in the mold possibly with differing dimensional capabilities. If the product of two primary units with different overall averages subsequently feeds into one final unit, the Control Chart on the final unit may show an up-down oscillation in the average with all the points near one and then the other limit or all the averages clustered about the center line depending on how the mix of the two primary products occurs. See Figures 3A and 3B for the same data re-arranged such that the subgroups of two contain two high or two low values (3A) or one high and one low values (3B). Notice the difference in the limits for the subgroup averages. Again it is better to control the key individual unit and not the composite. The use of a flow chart and a cause-and-effect diagram for the process can help to identify

Figure 3A: Control Chart for Multi-Stream Data
Subgroup Size = 2, Each Subgroup from Either Stream A or B

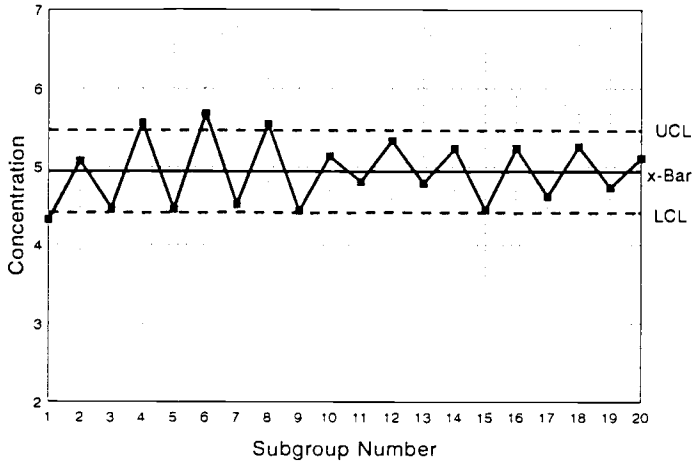
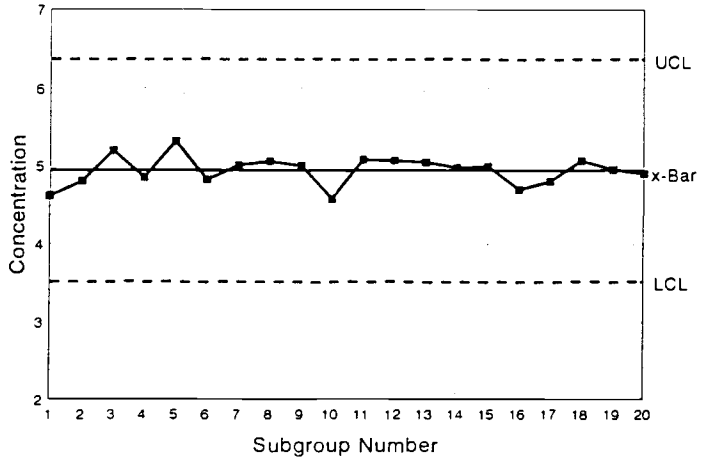


Figure 3B: Control Chart for Multi-Stream Data
Subgroup Size = 2, Subgroup from Both Stream A & B



systematically the key process variables for control and aid in the proper selection of subgroups.

Many new to the applications of Control Charts and Statistical Process Control, or SPC, think that continuous improvement will bring tighter control limits and thus more out-of-control situations. The opposite is the case, however, as a decrease in variability will not only reduce the variability in the range, R, but also in the subgroup averages, $\bar{x}$.

Control of Measurements

Many people do not think of a laboratory measurement procedure as a production process. It is just this, however, with the final measurement being the product. Too-frequent calibration of an instrument usually means more variation in the measurement. Deming in his book, Out Of Crisis [12], describes a funnel experiment in which a marble is dropped through a funnel centered over a target. When the marble does not hit the target, a common practice is to adjust the process, i.e. move the funnel. Deming gives several compensation rules which depict many calibration adjustment schemes used in most laboratories. The resulting effect on the variation is an eye opener.

In some laboratory procedures, there are standard reference materials; in others these may not exist. For laboratory control purposes it is not always necessary that standard reference materials be used and it is possible to build up a supply of internally generated test specimens which can be run on a systematic basis and charted to give a measure of control and signal when a calibration or process adjustment is required. The use of such samples can serve to reduce the variation in many measurement processes. See the paper by Levine, Kiszenia and Kyriakides [13] for one such an example. Another good reference would be Wheeler & Lyday [14]. It is the experience of the authors that the calibration frequency stated in manufacturer's manuals is not always applicable.

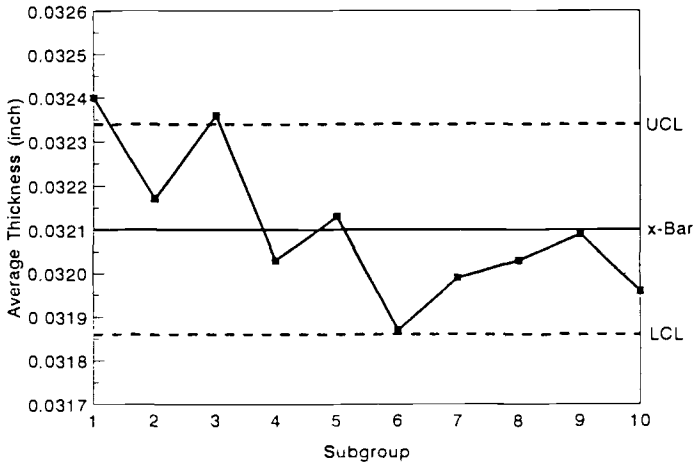
Measurement Precision

At times measurements are only reported to a single digit and the range has only a very limited number of values and a majority on zeros. In this case the calculated limits for the average may be too tight, due to the predominance of zeros in average range. A similar result is found where several digits are reported but the real variation is only in the last digit. An example of such a case may be seen in Figures 4A and 4B for the thickness of a tensile specimen. These are means of subgroups of size seven where the raw data were measured to 0.0001". This is shown as Figure 4A. The data were then re-scaled to 3 decimal places and plotted as the Control Chart in Figure 4B. Note the differences in limits and the absence the 2 of out-of-control points in the second plot (4B). There are just two values for the range in this plot, 0.000 and 0.001. The measurement process must be examined to see how another digit can be obtained. See Wheeler & Lyday [14].

Short Production Runs

In some manufacturing processes, the length of a production run may be short and the process switched to some other product before returning to the first product. For example, heavy zinc coated product may be only produced once a month and then for only a few days and then the galvanizing line changed to another product. In making

Figure 4A: Control Chart for As-Measured Thickness
X-Bar Chart - Subgroup Size=7



Range Chart - Subgroup Size=7

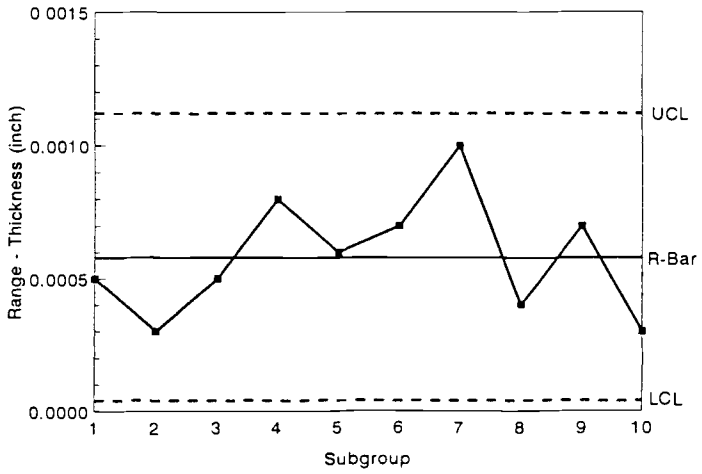
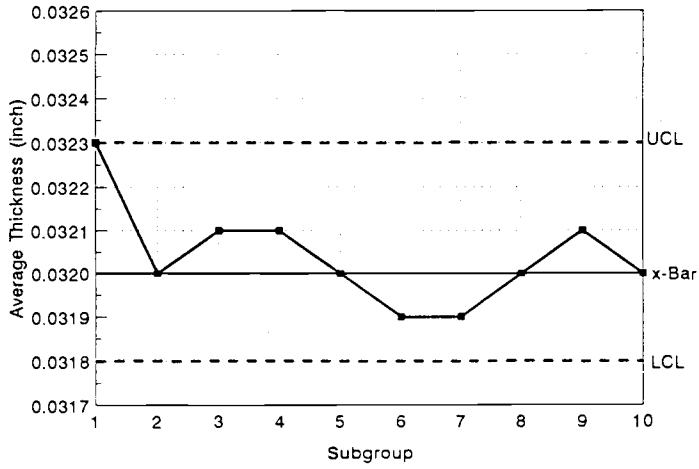
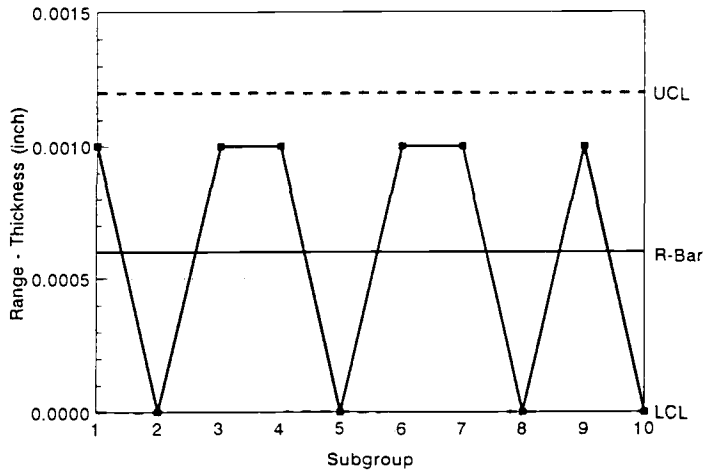


Figure 4B: Control Chart for Re-Scaled Thickness
X-Bar Chart - Subgroup Size=7



Range Chart - Subgroup Size=7



applications of Control Charts to such processes, special techniques are available. See, for example, Wheeler [15].

Summary

The authors have attempted in this short paper to enlighten the reader on some of the myths as well as some of the assumptions associated with control charts for variables. In addition a short history of control chart use was included.

The assumption concerning the normal distribution underlying Control Charts must remain as a guide, however, other distributions can and will produce similar results with similar risks in detecting special or assignable causes of variation.

The assumption that states that $\pm 3\sigma$ limits have performed both economically and practicably. The independence of each consecutive point on the Control Chart is achievable in most cases but can be ignored, with a change in the rules for detecting Out-of-Control points in situations where autocorrelation exists.

The most important point in the assumptions underlying the control chart is that rational subgroups be chosen with guidance that allows the process to speak to you. The control chart will tell you when to change the process variable settings to manufacture the product with the desired characteristics. Excessive, unwarranted, change in process variables will only serve to increase the product variability.

The authors point out that there are many applications for Control Charts for Variables where a rational subgroup of size other than 1 can not be justified. They propose that the Control Chart for Individuals be augmented by an EWMA trend line using $\alpha = .20$ to aid in detecting trends or drifts in the process average. This procedure is simpler than using a Moving Average.

Process variable control in preferable to product characteristic control. This is true because you can detect a problem and take corrective action sooner with the result of less out-of-control standard product. The control of measurements and precision is paramount to the entire system of control charts. Without reliable measurements and precision in these measurements you can not achieve control of the process variables necessary to manufacture product characteristics with the desired consistent properties.

In summary, if one uses the Control Chart with proper discretion, then the risk of making unnecessary changes to the process can be avoided. The resulting products will be produced with the lowest variation that can be achieved both economically and practicably.

References

1. Shewhart, Walter A., Economic Control of Quality of Manufactured Product, D. Van Nostrand Company, New York, 1931 (Republished by ASQC, 1980)

2. Shewhart, (1931) op. cit.

3. Chancellor, W. C. , "Application of Statistical Methods to the Solution of Metallurgical Problems in the Steel Plant", ASTM Proceedings, Vol 34, Part II, pages 891-909, 1931

4. Manual on Presentation of Data and Control Chart Analysis: Sixth Edition, MNL 7, ASTM, Philadelphia, PA, 1990

5. Freeman, H. A., Industrial Statistics, John Wiley & Sons, New York, 1942

6. Tippett, L. H. C., Technological Applications of Statistics, Williams & Norgate, London, 1952

7. Grant, E. L. & Levenworth, R. S., Statistical Quality Control, McGraw-Hill Book Co., New York, 6th Edition, 1988

8. Statistical Quality Control Handbook, Western Electric Company, 1956 (Now available from AT&T Technologies, Indianapolis, IN)

9. Wheeler, D. J., Chambers, D. S., Understanding Statistical Process Control, 2nd Edition,SPC Press, Knoxville, TN, 1992, pages 60-76

10. Shewhart (1931) op. cit. pp 147-148, 276-277

11. Shewhart (1931), op.cit., pp17-18

12. Deming, W. E., Out Of Crisis, MIT Press, 1986

13. Levine, B. H., Kiszenia, R., Kyriakides, S.,"Controlling The Tensile Testing Process Through SPC and Standardization", Second International Symposium on SPC and Sensors in the Steel Industry, CIM, August, 1988

14. Wheeler, D. J., Lyday, R. W., Evaluating the Measurement Process, 2nd Ed., SPC Press, Knoxville, TN

15. Wheeler, D. J., Short Run SPC, SPC Press, Knoxville, TN, 1991

Kenneth H. Moyer[1]

Quality Control of Powder Metallurgy (P/M) Parts for Magnetic Applications

REFERENCE: Moyer, K. H., **"Quality Control of Powder Metallurgy Parts (P/M) for Magnetic Applications,"** *Quality and Statistics: Total Quality Management, ASTM STP 1209,* Milton J. Kowalewski, Jr., Ed., American Society for Testing and Materials, Philadelphia, 1994.

ABSTRACT: Today's specifications for magnetic materials are based on the testing of material as supplied by the materials producer. These test methods are so designed that the specification qualifies the material only as purchased. However, additional operations, such as machining, welding, binding, etc., alter the magnetic properties from those of the raw material. The same problem arises with powder metallurgy (P/M) parts because, although their properties may be equivalent to a well-sintered part fabricated from the raw material, the parts are too small to evaluate using the test methods used to evaluate the raw material. Magnetic properties of sintered parts are predominantly a result of the sintering practice employed. Quality of these parts can vary appreciably as a result of the sintering practice employed in manufacture. Although the powders or the starting material may satisfy specification, there is no guarantee that the end user may receive a quality product after processing is complete. For P/M technology, rings can be fabricated from specified powders under conditions similar to parts being processed, and can be sintered in conjunction with the parts. The magnetic properties of the rings will be equivalent to the magnetic properties of the parts. A quality control procedure is proposed that will guarantee performance of P/M parts in accordance with currently existing P/M material specifications.

KEYWORDS: quality control, P/M magnetic parts, sintering furnaces, sintering atmospheres, phosphorus irons, ferritic stainless steels

Today powder metallurgy (P/M) offers processing that produces soft magnetic components of high purity, with superior magnetic properties [*1*]. Powder metallurgy processing of soft magnetic materials dates back to 1970. Although parts available at that time had good magnetic properties, today's improved processing has resulted in higher purity, superior quality parts available from a larger group of parts fabricators.

Magnetic parts fabricated by the P/M industry are relatively small, and generally multilevel. Owing to their complexity and size, magnetic testing of the parts themselves to assure quality is an almost impossible task. In addition, most of the P/M community is not readily versed in magnetic concepts and available test methods. As a result, parts fabricators are still in the early stages of developing systems that assure magnetic properties of consistent quality. Most parts fabricators sell on density only and do not guarantee magnetic properties. If high-quality parts requiring consistent performance are to become commercially viable, a realistic approach to assure quality must be developed. The purpose of this paper is to suggest such an approach.

[1] President, Magna-Tech P/M Parts Labs, 4 Green Briar Lane, Cinnaminson, NJ 08077.

Test Methods

Three basic test methods are commonly employed to determine DC magnetic properties (ASTM Test Method for Direct Current Magnetic Properties of Materials Using D-C Permeameters and the Ballistic Test Methods [A 341], ASTM Test Method for Direct-Current Magnetic Properties of Materials Using Ring Test Procedures and the Ballistic Methods [A 596], and ASTM Test Method for D-C Magnetic Properties of Materials Using Ring and Permeameter Procedures with D-C Electronic Hysterisigraphs [A 773]). Two methods, the hysteresigraph and the ballistic method, require ring specimens. Use of rings provides an ideal magnetic flux path, provided that the ratio of the mean diameter to radial width is not less than 10 to 1 (ASTM Practice for Procurement Testing, and Sampling of Magnetic Materials [A 34]). With this criteria satisfied, these test methods delineate the true magnetic properties of the material. For production, the hysteresigraph is faster, but more expensive as a capital expenditure. Although slower, equipment for the ballistic method is much less expensive and yields data equivalent to that obtained with the hysteresigraph.

Using permeameters to test small, irregular P/M magnetic parts is highly discouraged. Permeameters available commercially are generally employed to test rod, tube, wire, and strip. In general, long specimens of the order of 250 mm (10 in.) are required to assure low leakage loss to the surroundings. For smaller specimens, with a less uniform magnetic field, demagnetizing factors must be considered. These factors are geometry dependent. With small specimens, such as P/M parts these factors may be significant, and discourage confidence in the quality of P/M parts, owing to unaccountable losses of flux. In addition, with long specimens, mechanical clamping strains to the yoke may be considered as insignificant. However, with short P/M magnetic parts, the mechanical clamping strains can significantly affect the magnetic properties. On top of all these factors, different commercial permeameters may not agree with one another, or with the other specifications that employ rings for testing.

The most logical choice therefore is to settle on rings as the test specimen of choice. Since rings are known to provide true magnetic properties, and are of the size of the P/M parts that are being manufactured, they become a practical choice. If processed under identical conditions as the parts, rings provide meaningful data that would be truly representative of the quality of the parts sintered as one lot.

Processing Conditions

Factors of concern in processing include compacting, or green, density. This factor is important because the magnetic induction depends solely upon this property. As the density increases, so does the magnetic induction [*2*].

Since P/M parts for magnetic applications may be multilevel and complex in geometric shape, the green density may be variable throughout the part. It is therefore imperative that a critical density of the part be established between purchaser and supplier. Rings to serve as test specimens for the lot should then be compacted to the specified critical density. These rings should then be sintered in conjunction with the parts required to satisfy the order, thus experiencing the same sintering conditions that the parts receive. It has been shown that if the rings receive the same sinter as the parts, the magnetic properties are equivalent to those of the parts [*3*].

The next question focuses upon what constitutes a satisfactory sinter. The sintering conditions that are established determine both the sintered density and the carbon, nitrogen, oxygen and sulfur contamination levels of the material. Both sintered density and contamination levels depend upon the sintering temperature. A higher sintering temperature results in a higher sintered density, and higher magnetic induction. For iron and phosphorus ions,

temperatures may range from 1120 to 1260°C (2050 to 2300°F). The higher temperature (1260°C) ensures higher purity and higher sintered density, ensuring optimum magnetic properties. Since at least 1% shrinkage, not necessarily uniform, results, dimensions to tight tolerances are difficult to control. Ferritic stainless steels, 3% silicon iron and 50 Ni/50 Fe alloys must be sintered at temperatures of the order of 1260°C (2300°F) or higher in order to assure reduction of oxides and to provide grain growth to improve performance.

Second, the choice of time is a variable to be considered. P/M parts are fabricated from prealloyed or admixed powders. To permit occurrence of diffusion over surfaces and assure densification and bonding of the powder particles, at least 30 min at temperature is necessary for reduction of surface oxides and diffusion of atoms across particle surfaces. However, it is not uncommon for parts to be held longer at temperature to achieve grain growth and to reduce interstitial contaminants.

Other than density, purity and grain growth control structure sensitive magnetic properties. These properties include remanent magnetization, permeability, and coercive force. Therefore not only sintering temperature requires careful selection, but also the atmosphere selected to protect the parts from contamination. For optimum reduction and maintenance of low contamination levels, either a reducing atmosphere, such as hydrogen, is preferred, or a nonreacting atmosphere, such as vacuo. In no case should atmospheres containing nitrogen be considered. As with carbon, nitrogen has a low solubility in iron and its alloys, and readily reacts to form nitrides in ferritic stainless steels. This being the case, depending on the sintering conditions, nitrogen may be in solid solution or may have reacted to form nitrides [*4*]. Since the nitrogen cannot exist in solid solution, over time it precipitates to form nitrides, which in turn affect magnetic performance of the material. This process is known as aging, and can limit performance of the device over a period of time.

In summary, the green density, sintered density, sintering temperature, time at temperature, sintering atmosphere, and dew point within the hot zone, should all be carefully selected and reported to the customer along with the magnetic data determined. Should any secondary operations, such as grinding, sizing, or other processing which can produce mechanical stress be required, the parts should be annealed before shipment. Similarly, if the sintering atmosphere contains nitrogen, an aging treatment should be required to assure that parts have not been contaminated.

Commercially Available Powders

Powders available to fabricate P/M magnetic parts include iron, phosphorus iron mixtures, ferritic stainless steels, 50 Ni/50 Fe, and 3% silicon iron mixtures. Most of the sintered products produced from these powders have specifications that guarantee the quality of the sintered or the annealed parts (ASTM Specification for Phosphorus Iron Fabricated By Powder Metallurgy Techniques [A 839] and ASTM Specification for 50 Nickel-50 Iron Powder Metallurgy (P/M) Soft Magnetic Alloys [A 904]). In addition, a new specification for ferritic 410L and 434 stainless steels is presently being balloted for inclusion. A round robin is also being established to prepare a specification for 3% silicon iron.

Examples

Phosphorus Iron Development—The predictability of magnetic properties of well sintered phosphorus iron parts was understood as early as 1982 [*5*]. At this time, it was well established that the magnetic induction and residual induction were linear functions of the sintered density. As the sintered density increased, so too did these properties. Similarly it was shown that structure sensitive properties, such as permeability and coercive force, were largely a

function of the grain size, provided that sintering practice that would maintain interstitial contaminants, such as carbon, nitrogen, oxygen, and sulfur, at extremely low levels was followed. The work performed at that time permitted construction of predictability curves, through use of regression analysis, to accurately predict the magnetic properties of pure iron and a 0.45% phosphorus iron. Engineers used these curves to design magnetic components for applications that were emerging for electrical components. These equations were further expanded by additional work to include a 0.8 phosphorus iron alloy [6]. Either these curves, or magnetization curves of rings sintered at the supplying parts fabricator, are now customarily used to design new components and to establish their magnetic performance. Quality control of the parts is then established based upon the magnetic data furnished by the individual parts fabricator. More frequently today quality control is established based upon existing ASTM specifications (ASTM A 839 and A 904). These predictor equations are included as Figs. 1 through 4. Several equations are designated as new. These equations resulted from a second study of improved sintering practice.

Belt Furnace Sintering of P/M Parts

The demand for quality soft magnetic components has developed to the extent that more than 50% of production is dedicated toward fabrication of quality P/M magnetic parts. In these cases, quality is best maintained by sintering rings with the ordered parts. The magnetic properties of these rings are determined and submitted as a quality control report certifying magnetic performance of the supplied parts.

Many parts fabricators use conventional belt furnaces that are dedicated to sintering only magnetic parts. In this fashion, carbon, nitrogen, and oxygen content can be maintained at low levels, provided proper maintenance and control of the furnace is exercised. Sintering temperatures in belt furnaces are normally limited to 1120°C (2050°F). Time at temperature

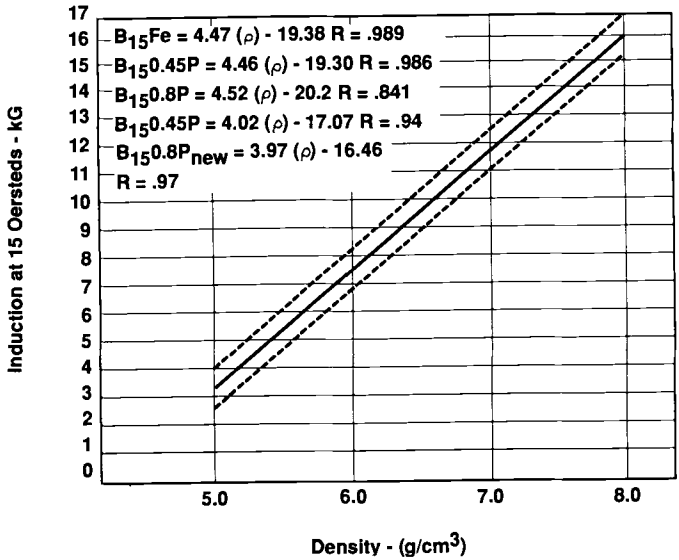


FIG. 1—*Magnetic induction of 0.45% and 0.8% phosphorus iron as a function of density.*

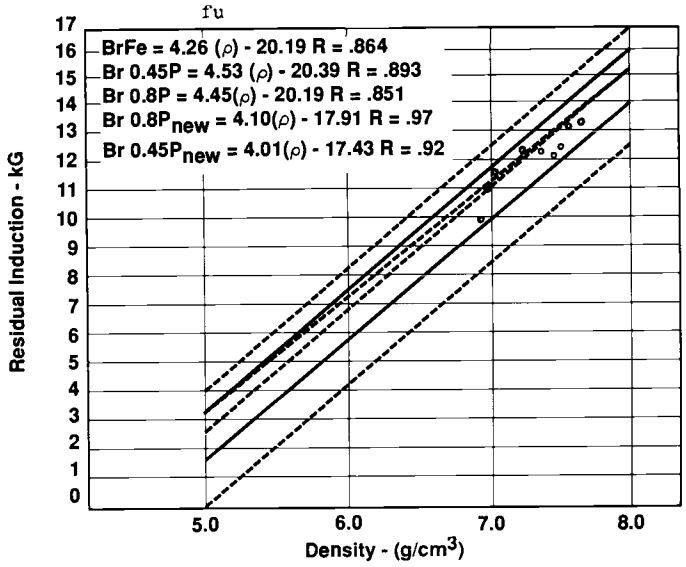


FIG. 2—*Residual induction of 0.45% and 0.8% phosphorus iron as a function of density.*

is established by zoning of the particular furnace, and the belt speed determined by management for maintenance of consistency and quality of parts as well as for reasonable economy. Parts normally experience 30 min at temperature. To maintain low carbon, nitrogen, and oxygen contamination levels, parts fabricators are now using pure hydrogen to reduce contaminants to low levels and enhance sintering.

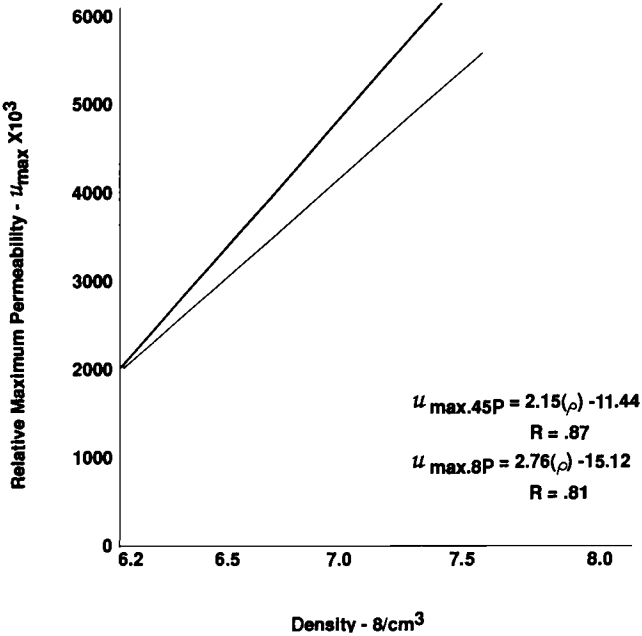


FIG. 3—*Relative maximum permeability of 0.45% and 0.8% phosphorus iron as a function of density.*

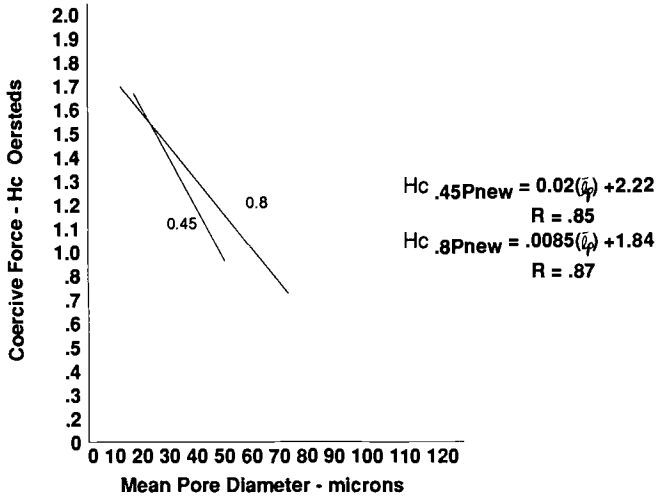


FIG. 4—*Coercive force of 0.45% and 0.8% phosphorus iron as a function of mean pore size.*

One such study of parts, which used rings for quality control of magnetic properties, is included as an example of the quality at which production can be maintained. It is normal for parts fabricators to shut down for furnace maintenance during the weekend and then exercise three-shift operation throughout the week. In this particular week, the sintering temperature was maintained at 1120°C (2050°F). The furnace atmosphere was pure hydrogen. The rings and parts experienced 30-min sintering time at temperature. A series of three rings was sintered in conjunction with the parts and extracted each shift of each day of operation. The magnetic properties that resulted from this week of sintering are shown in Fig. 5. All test data are in conformance with ASTM Specification A 839. Furthermore, the precision of the magnetic measurements is within the precision specified in ASTM Test Method A 596. The data demonstrate that quality control based on rings sintered with production parts is a viable procedure for assurance of quality of the P/M parts. Validity is established because magnetic properties of both parts and rings are established by the designated sintering cycle, provided that the rings' densities are equivalent to the parts.

Vacuum Sintering of Phosphorus Iron Parts

Modern technology has demonstrated that exceptional quality P/M parts can be sintered either in hydrogen or in vacuo [7]. In this case, a vacuum furnace designed to provide a partial pressure of hydrogen to reduce surface oxides is described. Before production, a cycle consisting of burn-off at 540°C (1000°F) and a sinter of 1120° (2050°F) for 1 h was developed within the laboratory. To accomplish this work, the vacuum furnace was first surveyed to determine uniformity of properties. The conditions established for this work are shown in Fig. 6. Ring specimens were located front, center and back on each side of the furnace. Two massive pipes were centered at the front and back in the middle of the fixture to determine the effects resulting from mulitple layering of parts. Work thermocouples were located at the front, middle, and rear of the load to monitor the sintering cycle. The magnetic properties resulting from this test are included as Table 1.

The results of a similar laboratory furnace are included for comparison at the right of the table. As can be seen, there is little variation in the magnetic properties regardless of position

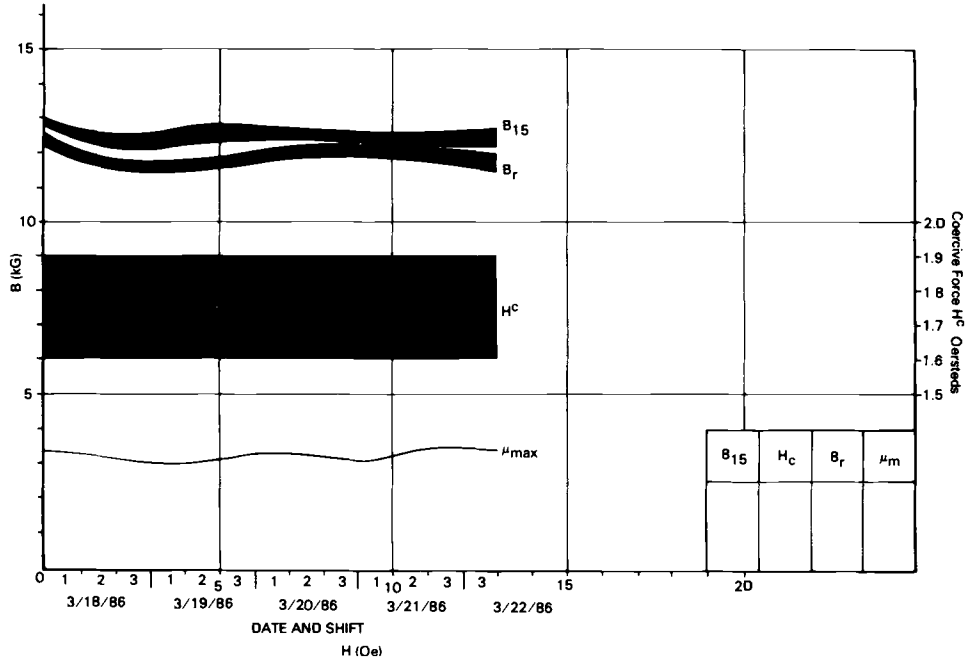


FIG. 5—*Test properties of 0.45% phosphorus iron from 1120°C production sinter.*

or in respect to the properties resulting from the laboratory sinter. In fact, the data are within the precison of ASTM Test Method A 596. At least five cycles equivalent to that described above have been conducted to sinter the above specimens. All properties have been found to be equivalent and within the precision of ASTM Test Method A 596. This example again demonstrates that the quality of the sinter controls the quality of the magnetic

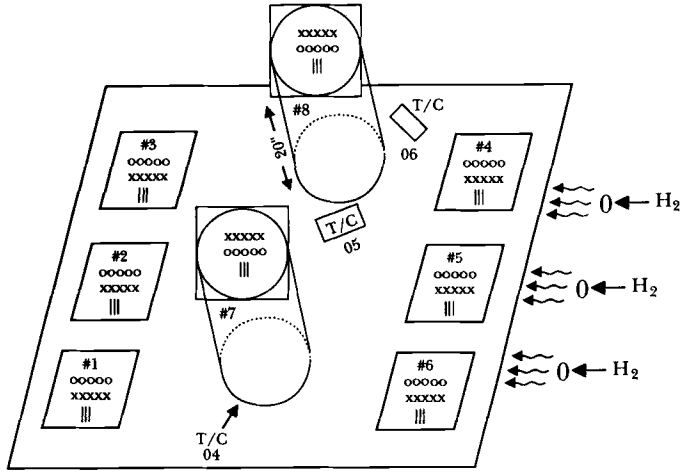


FIG. 6—*Location of sets of test specimens in production furnaces.*

TABLE 1—*Magnetic properties of 0.45% phosphorus iron specimens sintered in production vacuum furnace.*

Property	Site 1	Site 2	Site 3	Site 4	Site 5	Site 6	Site 7	Site 8	Lab
Magnetic Induction At An Applied Field of 15 Oe	13,300	13,600	13,600	13,700	13,700	13,600	13,600	13,600	13,500
Remanent Magnetization - gausses	12,800	12,900	12,800	13,000	13,000	13,000	13,000	12,900	12,800
Relative Maximum Permeability	5,900	6,000	6,000	6,200	5,800	6,000	5,700	5,800	5,500
Coercive Force - oersteds	1.2	1.2	1.2	1.2	1.2	1.2	1.3	1.2	1.2
Sintered Density - g/cm³	7.43	7.46	7.43	7.45	7.45	7.45	7.46	7.44	7.43
Average Grain Size - Microns	68	65	60	60	62	63	57	64	66
Chemical Analysis - %									
C	0.0006	0.0012	0.0010	0.0007	0.0008	0.0001	0.0030	0.0002	0.0020
N	0.0005	0.0004	0.0004	0.0004	0.0004	0.0003	0.0004	0.0006	0.0012
O	0.0512	0.0490	0.0507	0.0504	0.0506	0.0575	0.0543	0.0558	0.0354
S	0.0030	0.0029	0.0031	0.0028	0.0027	0.0029	0.0034	0.0029	0.0030

test data. The example suggests, again, that sintering test rings with production parts may represent the best method for quality control of functional sintered P/M magnetic parts.

Ferritic Stainless Steels

The following is the last example presented to demonstrate this method for quality control of magnetic parts. When the P/M task group charged with preparation of material specifications decides a specification is necessary, a round robin of the parts fabricators that produce magnetic parts of these materials is organized. Rings are compacted and sintered by the participating members under established protocol. Seven commercial parts fabricators participated in the ferritic stainless-steel round robin. The protocol established sintering the rings at 1260°C (2300°F) for 1 h in either hydrogen or vacuo. The results of the round-robin study are included as Table 2 [*8*]. Again, regardless of the participating company, the data are within the precision of ASTM Test Method A 596.

Summary

P/M parts that provide outstanding magnetic properties can be sintered. These properties can be achieved by many parts fabricators who have their sintering process under control. This control has been established in at least four round-robin programs that have led to publication of viable ASTM material specifications. These specifications should be specified by the purchaser of P/M magnetic parts to assure quality and performance.

To assure reliability of data, a satisfactory number of rings, as agreed upon between purchaser and supplier, should be sintered in conjunction with the P/M parts. Magnetic properties, as determined from the rings, should be supplied as a quality report of acceptance of the P/M parts. Permeameter tests should not be accepted, owing to losses of leakage

TABLE 2—*Magnetic properties of commercially sintered stainless steel powders.*

Property	Company	A410L	A434	B410L	C410L	C434
Magnetic Induction at an Applied Field of 15 oersteds	1	10,500	9,300	9,500	10,000	9,200
	2	10,900	9,600	10,200	10,500	10,000
	3	10,300	9,500	9,800	10,400	9,600
	4	11,300	10,200	10,700	11,300	10,100
	5	10,500	9,600	9,100	10,500	10,000
	6	10,900	10,200	10,400	10,800	10,300
	7	11,100	9,800	10,500	11,200	10,300
	Mean	10,800	9,700	10,000	10,700	9,900
	Std. Dev.	400	300	600	500	400
Remanent Magnetization gauss	1	8,300	7,200	7,400	7,900	7,100
	2	8,800	7,400	8,000	8,200	8,000
	3	8,300	7,600	7,800	8,400	7,600
	4	9,600	8,200	8,800	9,400	8,100
	5	9,200	7,800	7,500	9,200	8,400
	6	9,000	8,700	8,500	8,800	8,800
	7	9,200	8,900	8,700	9,500	8,700
	Mean	8,900	8,000	8,100	8,800	8,100
	Std. Dev.	500	600	600	600	600
Relative Maximum Permeability	1	1,300	1,200	1,000	1,100	1,100
	2	1,800	1,500	1,300	1,500	1,600
	3	2,000	1,700	1,300	1,700	1,500
	4	2,700	2,200	1,600	2,200	1,700
	5	2,500	1,700	1,400	2,200	1,800
	6	2,200	2,200	1,400	1,900	2,300
	7	2,200	1,700	1,500	2,100	2,200
	Mean	2,100	1,700	1,400	1,800	1,700
	Std. Dev.	500	400	200	400	400
Coercive Force oersteds	1	2.9	3.1	4.1	3.4	3.2
	2	2.5	2.2	3.2	2.6	2.4
	3	1.6	1.7	2.5	1.9	2.1
	4	1.4	1.6	2.6	1.9	2.0
	5	1.4	2.2	2.3	1.8	2.1
	6	2.0	1.7	2.9	2.0	1.6
	7	1.9	2.1	2.8	2.0	1.6
	Mean	2.0	2.1	2.9	2.2	2.1
	Std. Dev.	0.6	0.5	0.6	0.6	0.5

flux that are not easily measured from P/M parts of unfavorable geometry, necessitating need of a calculated demagnetizing field. If the magnetic properties are acceptable, then it may be assumed that the parts of equivalent density have been properly sintered and will provide equivalent magnetic properties.

References

[1] Moyer, Kenneth H., "Selection of Powders and Processing for P/M Soft Magnetic Components," *Advances In Powder Metallurgy*, 1990, Volume 2, Metal Powder Industries Federation, 105 College Road East, Princeton, New Jersey, 08540-6692, pp. 385–399.

[2] Moyer, K. H., et al., "The Effect of Porosity on the D-C Properties of Iron Compacts," 1980 TMS-AIME Fall Conference, Pittsburgh, Pennsylvania, October 5–9, 1980.
[3] Moyer, K. H., "Translation of Laboratory Sintering Practice to Production Processing of Phosphorus Iron Parts for Magnetic Applications," International Manufacturing Technology Show, Society of Manufacturing Engineers, Chicago, Illinois, September 5–12, 1990.
[4] Darken, Lawrence S. and Gurry, Robert W., *Physical Chemistry of Metals,* McGraw-Hill Book Co., Inc., 1953, pp. 373.
[5] Moyer, Kenneth H., McDermott, Michael J., and Ryan, James B., "Powder Metal (P/M) Parts for Magnetic Applications," Coil Winding/Electrical Manufacturing Expo '82', October 6–8, 1982, pp. 92–95.
[6] Moyer, Kenneth H. and Ryan, James B., "Extending the Usage of Phosphorus Iron P/M Parts for Magnetic Components," *Progress In Powder Metallurgy,* 1986, Volume 42, Metal Powder Industries Federation, 105 College Road East, Princeton, New Jersey, 08540, pp. 435–455.
[7] Moyer, Kenneth H. and Jones, William R., "Vacuum Sintering—A Way to Improve Ductility and Toughness of Structural P/M Parts at the Right Price," 1992 Powder Metallurgy World Congress, San Francisco, California, to be published by Metal Powder Industries Federation, 105 College Road East, Princeton, New Jersey, 08540.
[8] ASTM Report, "Report of Round Robin Project for Preparation of Ferritic Stainless Steel Magnetic Sintered Powdered Materials," ASTM, 1916 Race Street, Philadelphia, Pennsylvania, 19103-1187.

Edward G. Schilling[1]

THE TRANSITION FROM SAMPLING TO SPC

REFERENCE: Schilling, E. G., "The Transition from Sampling to SPC," Quality and Statistics: Total Quality Management, ASTM STP 1209, Milton J. Kowalewski, Jr., Ed., American Society for Testing and Materials, Philadelphia, 1994.

ABSTRACT: Modern quality control practice emphasizes the importance of process control in the creation of a quality product. When the process is in control it is possible to characterize the product forthcoming from the process, for the particular characteristic controlled, at the point in the process at which the control is instituted. When the process is out of control, the relationship between product and process is lost. It is then necessary to rely on acceptance sampling procedures to characterize the process. By combining the power of process control with the assurance of acceptance sampling, the process experimentation necessary for continual improvement can be undertaken with minimal risk to the consumer. An approach has been developed to transition from a process lacking control and/or capability to a controlled process with a C_{pk} value 1.33 indicating that the process is operating at an average level four sigma from the specification(s).

KEYWORDS: ABC Plan, Acceptance Sampling, Process Control, Capability, Mixed Plan, Sampling

Introduction

In a progressive atmosphere, new products, process and product modifications, as well as manufacturing difficulties can lead to periods of increased nonconforming material. It is at these times that statistical process control procedures are vital to obtain optimum process capability and as a vehicle for process improvement. And it is precisely at these times that acceptance control procedures can act as a bulwark to prevent defective material from passing through and to allow the process to continue to operate to achieve the full benefit of process control.

[1]Director and Professor, Center for Quality and Applied Statistics, Rochester Institute of Technology Rochester NY, 14623

As a process or product is introduced, little is known about potential assignable causes or, in fact, the particular characteristics of the process which require control. At that time, it is appropriate to do 100% inspection or screening while data are collected to allow for implementation of more economic procedures. After the start-up phase of production process development, acceptance sampling plans may be instituted to provide a degree of protection against an out of control process while at the same time collecting data for eventual implementation of process control. When dealing with a process, acceptance sampling should be viewed as an adjunct and precursor of process control, rather than as a substitute for it.

Sometimes acceptance sampling plans can play the role of a process control device. When this is done, emphasis is on feedback of information on the process rather than simple acceptance or rejection of lots. Eventually enough information has been gathered to allow implementation of control charts and other process control devices along with existing acceptance sampling plans. Finally a state of control is attained and verified. It is at this point that acceptance sampling of lots should be phased out in preference to expanded process control as control is achieved. In its turn, when a high degree of confidence in the process exists, control charts, too, should be gradually eliminated in favor of check inspection or no inspection at all.

PROCESS AND PRODUCT CONTROL

Acceptance sampling plans have been devised to be used in two distinct situations (i) Product Control (acceptance of unique lots); (ii) Process Control (acceptance of a process). The former applies to a specific lot for which a frame can be constructed and from which samples are drawn. The latter relates to a series of lots from an individual producer. The stream of lots is regarded as a process. Here the population is the conceptually infinite number of units that could have been produced had the process been left to operate as it was when the lot was produced. The frame consists of the lot itself from which a sample is drawn to characterize the process. So it is also with a control chart where samples of, say, five successive units are used to provide a snapshot of the population at a given point in time rather than to characterize the population (or populations) represented by the pieces actually produced. It tells us when the hypothetical population changes and, hence, when more than one population is involved. For prediction, and hence estimation of the capability of a process, is impossible without statistical control.

The difference between sampling a unique lot (Type A Sampling) and sampling a process (Type B Sampling) is clarified somewhat by W.E. Deming's distinction between two types of study:

- Enumerative Study--aim is better knowledge about material in a frame
- Analytic Study--aim is to obtain information by which to take action on a cause system which has provided material in the past and will produce material in the future

Clearly, the inspection of an individual lot involves an enumerative study. A plan set up to provide acceptance on a process, however, involves an analytic study since its purpose is action on the process in terms of

encouraging process control. Most advanced acceptance sampling procedures were devised to be used with the process this way, including for example: skip-lot, chain, cumulative results plans, AOQL plans, and MIL-STD-105E.

Process control studies are by nature analytic. The objective is to characterize the process at any point in time and not necessarily the product which is to be produced. Therefore, the sampling procedures are not necessarily those appropriate for random sampling from the population or lot of product produced over a given period. Rather, the samples are structured to give sure and definitive signals about the process. This is the essence of rational subgrouping as opposed to randomization. It explains why it is reasonable to take regular samples at specific intervals regardless of the quantity of product produced. It also indicates why successive units produced are sometimes taken as a sample for process control, rather than a random selection over time. Thus, to characterize the product produced in a given period it is essential to use acceptance sampling procedures, unless the process remained in-control.

RATIONAL SUBGROUPS

Use of the control chart to detect shifts in process centering or spread requires that the data be taken in so-called rational subgroups. These data sets should be set up and taken in such a way that variation within a subgroup reflects non-assignable random variation only, while any significant variation between subgroups reflects assignable causes. Experience has shown that the reason for an assignable cause can be found and will give insight into the shifts in process performance that are observed. Rational subgrouping must be done beforehand. Control charts are no better than the effort expended in setting them up. This requires technical knowledge about the process itself.

In contrast, random samples are used in acceptance sampling to characterize the population of product produced throughout a given production interval. Thus, an hour's production might be characterized by a rational subgroup of five successive units for a control chart or by a random sample of size five over all units produced that hour for acceptance sampling purposes. Only if the process is in control will the former be an effective substitute for the latter.

TRANSITION PROCEDURES

Many applications of SPC require a process in control with a capability well within specifications. The measure of the product/process relationship is the capability index C_{pk} which is the ratio to three sigma of the minimum distance of the average to the specification limit(s). An approach has been developed, called the ABC plan, to transition from a process lacking control and/or capability to a controlled process with a C_{pk} value 1.33, indicating that the process is operating at an average level four sigma from the specification(s). The road to attain such a C_{pk} may sometimes be long and hard. Accordingly, associated sampling plans are used to characterize the product while the process is in danger of lack of control. These plans use the AOQL principle, with screening of rejected lots, to assure that product outgoing from the inspection will not be worse than 0.3 percent nonconforming in the long run. This is roughly equivalent to C_{pk} =1. The sample size is reduced as control and capability are achieved until only the associated

control chart is used for both product and process acceptance after $C_{pk} = 1.33$ is finally attained. For processes already capable, this can be quickly achieved. Switching from stage to stage is through stringent requirements in terms of capability and control.

The ABC plan is a product acceptance procedure explicitly utilizing SPC to achieve high levels of process capability.

The ABC plan is subject to several constraints:

- AQL's are not utilized
- Acceptance number of zero (i.e., c=0)
- Simplicity

The plan progresses through three stages:

- Stage C - control being established
- Stage B - capability being established
- Stage A - capability being maintained

Stage C employs a simple single sampling plan, while Stage A utilizes only a control chart. Stage B uses a mixed variables - attributes procedure which is a double sampling procedure using a variables plan on the first stage and attributes on the second.

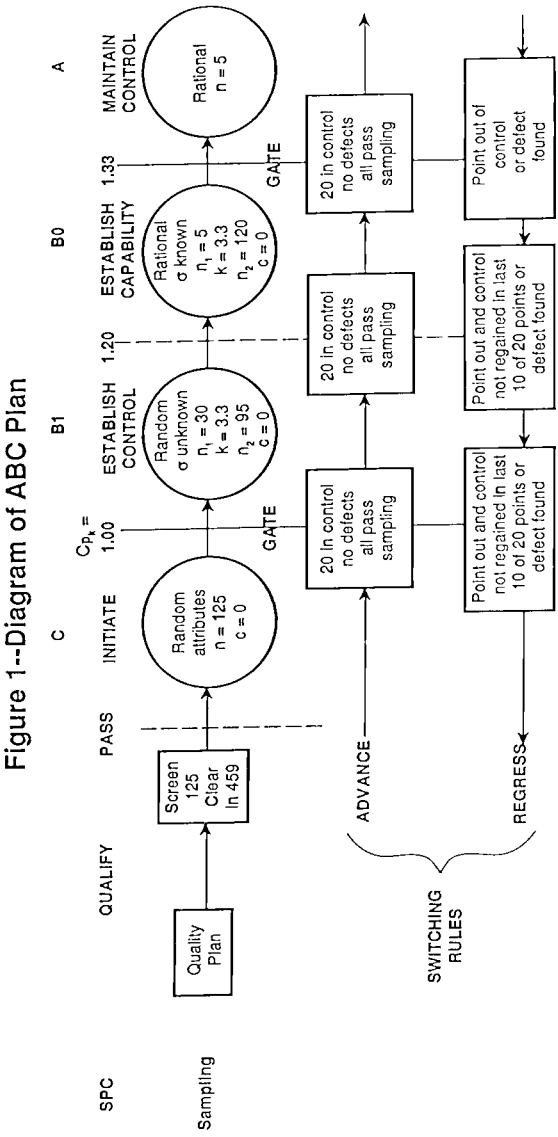
These categories allow suppliers to be classified by the user as a motivational device. The objective of the ABC plan, then, is to progress through these stages until a capability index of $C_{pk} = 1.33$ has been demonstrated and is being maintained.

Since there is no capability without control, control charts are central to the ABC plan. But since product performance cannot be predicted from control charts unless the associated process is in control, sampling plans were necessarily incorporated into the procedure to be used until such time as a capability of $C_{pk} = 1.33$ and the capacity for continuing control are demonstrated. When this is accomplished, sampling ceases. Thus, the supplier is given the latitude to develop and use a sound SPC program, while the customer is protected by the sampling procedure.

The approach is so structured that the user will learn about proper implementation of SPC in moving through the stages. The plan is intended as an educational vehicle as well as a process and product verification procedure. This aspect of the plan is deemed essential since SPC demands increasing levels of sophistication in progressing toward high levels of capability and control. It is through this knowledge that the suppliers will ultimately build quality into the product while containing costs. The educational aspect also serves to enhance other products and processes which are not directly under the ABC plan but which benefit from increased process and product understanding.

ABC PLAN

The ABC procedure was developed to act as a means to eventually attain and demonstrate statistical control and the desired capability, while at the same time encouraging ever increasing sophistication on the part of the suppliers. This procedure is shown in (Figure 1).



ACCEPTANCE

- Must screen 125 clear of defects
- Lotting - Screen all rejected lots
- SPC - Screen immediate and next production Interval when out of control

CONTROL

- Keep control chart for $\bar{X}$ and R throughout
- Can pass to stage A in 20 points on control chart If entering ABC with C_{P_k} = 1.33 already

The following prerequisites are assumed to be met before the plan is implemented.

1. Measurement with process average $\leq$ 1.0 percent
2. Steady stream of product
3. SPC Plan

The SPC Plan would contain a detailed description of the proposed procedure for developing and maintaining control at the capability desired. This is important from an educational perspective as well as providing a rationale for the procedures to be used with the specific process and product involved in developing control and capability.

Items to be addressed in the SPC Plan include

1. What constitutes a lot, sublot, or production interval
2. What constitutes in and out of control
3. How are control limits updated
4. How and when is C_{pk} updated and reevaluated
5. What corrective action procedures will be followed
6. Conditions for exemption from switching rules

The plan begins with a screening procedure which continues until a sequence of 125, of at most 459 units, is found conforming. Here, for simplicity, units out of specification are regarded as defective. This qualifies the supplier to start the plan. If such a sequence is not found, the process is not capable of meeting the ABC Plan and some other procedure should be used. Once qualified, the plan moves to stage C.

Stage C makes no assumption about the control or capability of the process involved. A standard sampling plan of n = 125, c = 0 is used. This plan has an AOQL of 0.3 percent, thus assuring a long term average nonconforming less than or equal to 0.3 percent if rejected lots are screened. This roughly equals the percent nonconforming for a process with capability $C_{pk} = 1.00$. During this stage efforts are started to establish control and demonstrate a capability of $C_{pk} = 1.00$. When this is accomplished with a process in control for at least 20 points, switch is made to stage B.

Stage B involves continued efforts to maintain control and to attain a capability index of $C_{pk} = 1.33$. At the beginning of this period, if C_{pk} is less than 1.20, the provisional sampling plan B1 is used with $n_1 = 30$, $k = 3.3$, $n_2 = 95$, $c = 0$. This is a so called mixed plan with unknown standard deviation. The latter is necessary since efforts to attain control at the capability desired may cause fluctuations in variability. Once a capability of $C_{pk} = 1.20$ has been attained, switch is made to the known standard deviation plan B0 with $n_1 = 5$, $k = 3.3$, $n_2 = 120$, $c = 0$. The B0 plan utilizes the control chart estimate of σ as a more accurate estimate of variability in light of a controlled process. Both Plans B1 and B0 have AOQL $\cong$ 0.3 percent. Note, that to reduce the ASN, the first sample is inspected for both variables and attributes before taking the second sample. Switch is made to Stage A when $C_{pk} = 1.33$ after 20 or more points in control have been attained with process capability at that level.

Stage A relies upon the control chart completely for acceptance of product. Switch is made back to the appropriate previous Stage B plan at any point in the procedure if the process goes out of control and control is not regained for 10 successive points in a maximum of 20 successive points. Note that

this should not immediately penalize the supplier since the Stage B0 mixed plan uses sample size 5 with the rational subgroups from the control chart. Switch to the previous stage is, of course, also made any time a defective unit is found. Similarly, switch is made to Stage C if the requirements of stage B are not met.

Operating and switching rules for the implementation of the ABC plan are as follows:

OPERATING RULES

- Sigma known requires range in control
- Screen all rejected lots
- Lotting of production intervals is allowed where random sampling is indicated i.e. stage C, B1
- At any time the customer may require the supplier to verify that the C_{pk} level is appropriate to the criteria being used.

SWITCHING RULES

- A point outside control limits or one defect in a sample is considered evidence of lack of control
- Screen immediate and next production interval when an out of control condition is observed
- Return to previous stage if control not regained in last 10 of at most 20 samples, or if defect is observed in sample (must be in control with $C_{pk} \geq 1.33$ to stay on A)
- Evaluation of C_{pk} for switching requires the last 20 points in a row to be in control ($.997^{20} = .94$)
- Re-evaluate C_{pk} any time control limits are changed and move to appropriate stage accordingly.

By combining the power of process control with the assurance of acceptance sampling the ABC plan provides a means by which continual improvement can be undertaken.

TECHNICAL CONSIDERATIONS

Properties of the plans are shown in Figure 2. The unknown mixed plan was determined using the Wallis approximation to derive the unknown standard deviation plan having similar properties to the known standard deviation plan: $n_1 = 5$, $k = 3.3$. The relationship, to develop the unknown standard deviation sample size, n_1^* is

$$n_1^* = \left(1 + \frac{k^2}{2}\right) n_1$$

and the same acceptance constant k = 3.3 is used. Properties of the unknown standard deviation plan were developed using Monte Carlo techniques.

Figure 2--Properties of ABC Sampling Plans

	C n = 125 c = 0		B1 Unknown sigma n_1 = 30, k = 3.3 n_2 = 95, c = 0			B0 Known sigma n_1 = 5, k =3.3 n_2 = 120, c = 0		
p	**Pa**	**AOQ**	**Pa**	**AOQ**	**ASN**	**Pa**	**AOQ**	**ASN**
.000317	.961	.00003	.999	.00003	36.0	1.000	.00003	12.0
.0005	.939	.0005	.964	.0005	74.1	.967	.0005	65.6
.001	.882	.0009	.914	.0009	89.0	.918	.0009	85.9
.00135	.845	.0011	.879	.0012	94.3	.881	.0011	93.9
.005	.534	.0027	.563	.0028	105.3	.558	.0028	115.6
.008	.366	.0029	.383	.0031	102.1	.380	.0030	117.4
.01	.285	.0029	.300	.0030	98.6	.295	.0029	117.3
.02	.080	.0016	.080	.0016	81.3	.082	.0016	113.1
.05	.002	.0001	.002	.0001	50.0	.002	.0001	97.8

Note that while the AOQL of the plans in stages A, B0 and B1 are maintained, the average sample number, ASN, decreases in a rational way in moving through the stages. This is shown below.

Cpk	**p**	**Stage**	**Typical p**	**ASN**
< 1	> .00135	C	.01	125
1-1.2	.00135-.0002	B1	.0005	74.0
1.2-1.33	.0002-.00003	B0	.00003	12.0
> 1.33	< .00003	A	.00001	5.0

Stage B utilizes a mixed variables attributes plan as originated by H.F. Dodge [1] and extended by Schilling and Dodge [2] for known standard deviation. Mixed plans for unknown standard deviation are discussed in Gregory and Resnikoff [3] and Schilling [4]. Process control aspects of these procedures will be found in Ott and Schilling [5]. More detail on the ABC Plan is given in Liuzza, Roediger and Schilling [6].

CONCLUSION

The ABC plan provides a vehicle to attain control of a process at an acceptable level of capability. It illustrates the synergy which can be achieved between process control and acceptance sampling when these methods are properly applied. The plan presented here is a prototype, subject to further development as experience is gained in application of the procedure.

ACKNOWLEDGMENTS

The development of the ABC Plan was done in cooperation with Carmen Liuzza and Paul Roediger of USAAMCCOM whose contributions were central to the development, statistical thinking and practical nature of the result. Their guidance, assistance, and intellectual support cannot be understated. The research on the ABC Plan was supported by the U.S. Army USAAMCCOM, AMSMC-QAH-P Bldg 62 Picatinny Arsenal, N.J. 07806-5000. A more extensive discussion of the technical aspects of implementation of the ABC Plan is given in Liuzza, Roediger and Schilling [6].

REFERENCES

[1] Dodge, H. F. (1932), Statistical Control in Sampling Inspection, *American Machinist*, October: pp. 1085-1088; November: pp. 1129-1131.

[2] Schilling, E. G., and H. F. Dodge (1969), Procedures and Tables for Evaluating Dependent Mixed Acceptance Sampling Plans, *Technometrics*, 11 (2): pp. 341-372.

[3] Gregory, G., and G. J. Resnikoff (1955), *Some Notes on Mixed Variables and Attributes Sampling Plans*, Technical Report No. 10, Applied Mathematics and Statistics Laboratory, Stanford University, Stanford, California.

[4] Schilling, E. G. (1982), *Acceptance Sampling in Quality Control*, Marcel Dekker Inc., New York, pp. 347-365.

[5] Ott, E. R. and E. G. Schilling (1990), *Process Quality Control*, McGraw-Hill New York, pp. 171-220.

[6] Luizza, C.J., P. A. Roedigar and E. G. Schilling (1992), "ABC: A Sampling Plan to Attain Process Control and Capability", Second Annual Statistical Process Control Conference Proceedings, The American Defense Preparedness Association, Statistical Process Control Division, Hunt Valley, MD, March 3, 1992, pp 19-28.

Milton J. Kowalewski, Jr.,[1] and Charmaine R. Farrar[2]

OVERVIEW OF TECHNICAL AND SOCIAL SYSTEMS IN ORGANIZATIONAL CHANGE

REFERENCE: Kowalewski, Milton J.Jr., Farrar, Charmaine R. " **Overview of Technical and Social Systems In Organizational Change** ", Quality and Statistics: Total Quality Management, ASTM STP 1209, Milton J. Kowalewski, Jr., Ed., American Society for Testing and Materials, Philadelphia, 1994.

ABSTRACT: What light does Behavioral Science and Organization Development shed on Total Quality Management (TQM) in Earth's global environment ? The social and technical systems of Europe have been joined together in the European Community 1992 (EC '92) coalition. As part of EC '92 the International Organization of Standardization (ISO) has provided the technical standardization system for inter-country approval and use. Other non-European countries have joined the ISO movement including the United States of America. Change has been a result of this movement. Most of it has been on technical, quality, and management systems. The next century will require and reveal changes in behavioral thinking and organizational structure. TQM is at the frontier of this change. A case-study of change resulting from TQM in government is presented in this paper.

KEYWORDS: total quality management, organizational change, European Community, International Organization of Standardization, management systems, behavioral science, organization development, case-study, Vision 2000, structural frame, human resource frame, political frame, symbolic frame, learning organization, critical thinking, atunement, empowerment, process, entrepreneurial, cultural assessment, continuous improvement, facilitation, quality action teams.

INTRODUCTION

In 1968 Eric Trist, a well-known Social Psychologist, described TQM before it was an abbreviation. He concluded that [TQM] is " ... a joint optimization of the technical and social systems as a goal of organizational change ". [1] The September 20, 1992 vote by the French populace on approval of the Maastricht Treaty provisions for a common currency in Europe by 1999 barely passed. This indicates that the French people and other nations in Europe may not be ready to relinquish control of their financial future to the European Monetary Unit of exchange. This trend conflicts with the total quality goals of the European Community (EC) for rapid implementation of the EC single market arrangement, targeted for full operation in 1992. [2]

[1] Quality Auditor, EG&G Rocky Flats, Inc., P.O. Box 464 Bldg. 850, Golden, CO 80402. This paper is prepared under the auspices of the U.S. Department of Energy contract with EG&G, Rocky Flats, Inc., # DE-AC34-90-RF62349.

[2] Civil Engineer, Federal Highway Administration, 555 Zang Street, Denver, CO 80225.

Vision 2000 [2]

The report of the Ad Hoc Task Force of the International Organization for Standardization (ISO) Technical Committee ISO/TC 176 on Quality Assurance, which has become known as "Vision 2000", describes a strategic plan for ISO 9000 series standards architecture, numbering, and implementation. Globalization has become a present reality in the few years since the ISO 9000 series was published in 1987. Today, all but the smallest or most local commercial and industrial enterprises are finding that their principal market-place competitors include companies headquartered in other countries. Four generic product categories are affected by the ISO 9000 series: hardware, software, processed materials, and services. In addition, four strategic goals for the ISO 9000 standards, and their related ISO 10 000 series standards have been developed by ISO/TC 176. They are: universal acceptance, current compatibility, forward compatibility, and forward flexibility. Just as organizations will have to learn to implement ISO 9000 quality management and quality assurance skills, they will be required to keep their focus on market need, harmonization and incorporation of other European Standards (EN 45000 and ISO/IEC Guides), and continuous improvement, as portrayed by The Partnering Pyramid of Sematech[3] (Figure 1). [3]

HISTORY OF ORGANIZATIONAL DESIGN

The word organization derives from the Greek *organon*, meaning tool or instrument. It is thus hardly surprising that the concept of organization is usually loaded with mechanical or instrumental significance. [4] Is [organizational] management something that was only discovered within the last 50 years or so ? Clearly not. In Biblical times, Joseph in Egypt proved himself to be an excellent forecaster and an astute manager of scarce resources. The Roman Empire, at its height, had assumed responsibility for managing a good part of the known world. Organization was characterized as dominant when over 2,300,000 blocks of stone, each weighing two and one-half tons, were arranged to build the Great Pyramid at Giza. [5] Classical management theory had the basic thrust of management as a process of planning, organization, command, coordination, and control. The principles of classical management theory were: unity of command, scalar chain, span of control, staff and line, initiative, division of work, authority and responsibility, centralization of authority, discipline, subordination of individual interest to general interest, equity, stability of tenure of personnel, and espirit de corps. [6] The Wharton School inaugurated the first undergraduate school of business in 1881. The growth of big business under Andrew Carnegie in 1873, John D. Rockefeller in 1879, Henry Ford in 1907, and W.C. Durant (General Motors) in 1908 revolutionized the structure of business organizations from the owner-managed enterprise to the publicly held corporation in which ownership was separated from management. [7] This distancing of owner from manager, and subsequently manager from worker, has created the chasms of collaboration, cooperation, and celebration that David Severson says, "....distinguishes world-class companies, large and small, from their competitive, and often competitively counterproductive, industry associates [when present]." [8]

The War Years

The 1900s have been occupied by wars fought across continents and oceans. These activities fostered a very militaristic mechanical context for organizations, where a task was defined, and resources were allocated

---

[3] Used with permission of SEMATECH, a consortium dedicated to improving the competitive position of U.S. semiconductor manufacturers and to developing standards. Designed by Jeff Riddle.

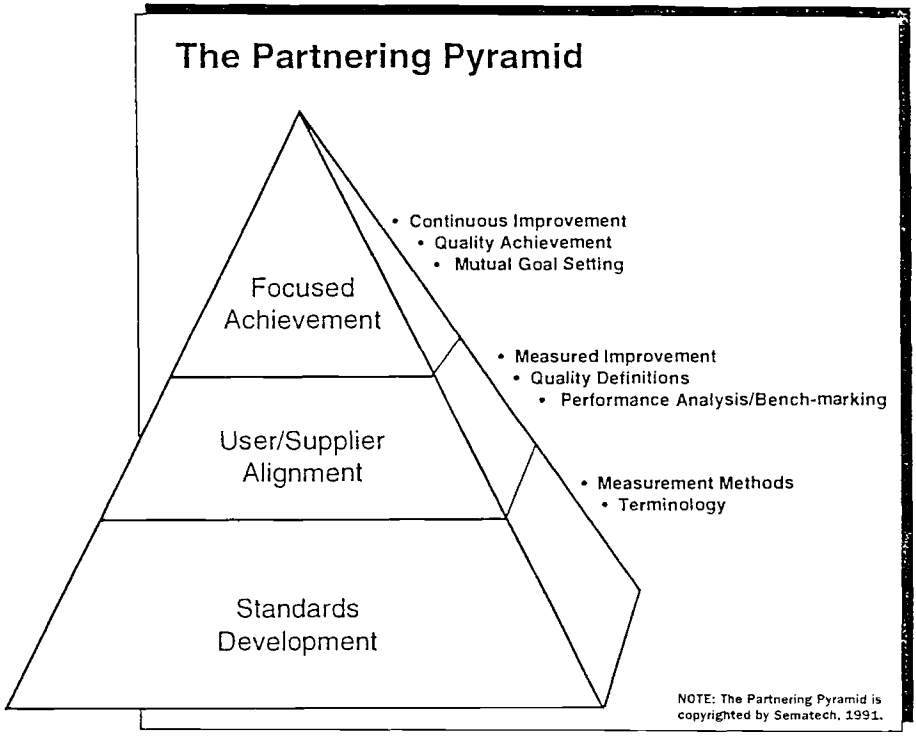


FIG. 1--Sematech Partnering Pyramid

to meet the critical demands of a " war-machine ". Individual needs were sacrificed for the good of the country, and for organizations supporting the war efforts. In the 1950s the emphasis on professional athletics impacted organizational context. A transition to team-based organizational management occurred. This was influenced greatly by the work of earlier and contemporary " Quality Gurus "; Walter A. Shewhart, George Edwards, Martin Brumbaugh, Eugene Grant, Harold F. Dodge, W. Edwards Deming, Joseph Juran, Armand E. Feigenbaum, Kaoru Ishikawa, Genechi Taguchi, and others. The demands of communications placed upon Western Electric, and weapon materials being manufactured for the U.S. Government brought some of the above individuals together to establish tools of quality for the statistical measurement and graphing of variation in pieces, parts, and components. These activities fostered teamwork, quality circles, value engineering(a team-based problem-solving technique based on function started by Lawrence Miles at General Electric in the 1950s), and the sciences of quality control and quality assurance. It was through these individuals and activities that Eric Trist's observations of joint optimization of technical and social systems became a recognized reality of organizational change.

MODERN ORGANIZATION DESIGN/TQM

Organization can be viewed from many frames or metaphors. One can use the metaphor of sport teams. Your organization may be organized and function like a baseball team with nine vital elements, each with their area of responsibility and authority, similar to a structured machine-like company. Another analogy can be made with a basketball team where only five elements function to accomplish the goal of winning. However, this team is more individualistic and creative, similar to an entrepreneurial organization of extroversive, sensing, feeling, perceptive Meyers/Briggs [9] personality types. Lastly, not to overlook the football enthusiasts, the metaphor of 11 elements closely arranged to protect key individuals in the organization, and to sacrifice limb and life reeks of the bureaucratic militaristic context.

Bolman and Deal [10] discuss four frames that organizations can be referenced to. They are: the structural , the human resource, the political, and the symbolic.

The Structural Frame

McDonalds, which made the Big Mac a household word, is an enormously successful organization that dominates the fast-food business. Another well structured entity is Harvard University. Both of the above, although very different in style and geographic concentration, share the structural characteristics of organizations with goals, boundaries, levels of authority, communications systems, coordinating mechanisms, and distinctive procedures. This form was first recognized by Fredrich W. Taylor in 1911, the father of time-and-motion studies, and in 1947 by German sociologist Max Weber. Weber's major dimensions of bureaucracy are: (1) a fixed division of labor, (2) a hierarchy of offices, (3) a set of rules governing performance, (4) separation of personal from official property and rights, (5) technical qualifications for selecting personnel(not family ties or friendship), and (6) employment as a long-term career. [11]

The Human Resource Frame

The human resource frame draws on a body of research and theory built around the following assumptions:

1. Organizations exist to serve human needs(rather than the reverse)
2. Organizations and people need each other.(Organizations need ideas, energy, and talent; people need careers, salaries, and work opportunities.)
3. When the fit between the individual and the organization is poor, one or both will suffer: individuals will be exploited, or will seek to exploit the organizations, or both.
4. A good fit between individual and organization benefits both: human beings find meaningful and satisfying work, and organizations get the human talent and energy they need. [12]

The Political Frame

At sunrise on January 28, 1986, it was clear but very cold at Cape Canaveral-more like New Hampshire, where Christa McAuliffe was a high school teacher, than Florida. At 11:38 A.M., space shuttle *Challenger* lifted off. One minute later, there was a massive explosion in the booster rockets. Neither the structural frame nor the human resource frame of organization focuses on the kinds of political issues that led to the *Challenger* disaster. The following 5 propositions summarize the political perspective.

1. Organizations are coalitions composed of varied individuals and interest groups (for example, hierarchical levels, departments, professional groups, gender, and ethnic subgroups)
2. There are enduring differences among individuals and groups in their values, preferences, beliefs, information, and perceptions of reality. Such differences change slowly, if at all.
3. Most of the important decisions in organizations involve the allocation of resources: they are decisions about who gets what.
4. Because of scarce resources and enduring differences, conflict is central to organizational dynamics, and power is the most important resource.
5. Organizational goals and decisions emerge from bargaining negotiations, and jockeying for position among members of different coalitions. [13]

The Symbolic Frame

This frame represents a world that departs significantly from traditional canons of organizational theories: rationality, certainty, and linearity. It is based on the following unconventional assumptions about the nature of organizations and human behavior:

1. What is most important about any event is not what happened, but what it means.
2. Events and meanings are loosely coupled: the same events can have very different meanings for different people because of differences in schema that they use to interpret their experience.
3. Many of the most significant events and processes in organizations are ambiguous or uncertain-it is often difficult or impossible to know what happened, why it happened, or what will happen next.
4. The greater the ambiguity and uncertainty, the harder it is to use rational approaches to analysis, problem solving, and decision making.
5. Faced with uncertainty and ambiguity, human beings create symbols to resolve confusion,increase predictability, and provide direction. (Events themselves may remain illogical, random, fluid, and meaningless, but human symbols make them seem otherwise.)
6. Many organizational events and processes are important more for what they express than for what they produce: they are secular myths, rituals, ceremonies, and sagas that help people find meanings and order in their experience.

The symbolic frame is especially helpful in understanding the dynamics of legislatures, public agencies, educational organizations, religious orders, and even health care organizations. But the recognition that symbols are also important in corporations has spread rapidly in recent years. [14]

Total Quality Management

American corporations are at a critical watershed because they face transforming economic and social environment which has emerged since the 1960s. This new context for corporate America makes past responses less effective; it changes the tasks for management at all levels and encourages the search for better ways to involve the entire work force in innovative problem solving. Three new sets of skills are required to manage effectively in such integrative, innovation-stimulating environments, where individual " corporate entrepreneurs " work through participative teams to produce small changes that can later add up to big ones. First are " power skills "-skills in persuading others to invest information, support, and resources in new initiatives by an " entrepreneur." Second is the ability to manage the problems associated with the greater use of teams and employee participation. And third is an understanding of how change is designed and constructed in an

organization-how the microchanges introduced by individual innovators relate to macrochanges or strategic reorientation. [15] A model that has been purported to accomplish the above transformation is Total Quality Management (TQM).

Principles of TQM--James Saylor [16] provides an acrostic of principles of TQM that describe the TQM model. It is as follows:

Provide a TQM environment
Reward and recognize
Involve everyone in everything
Nurture supplier partnerships
Create a continuous improvement process
Include quality as an element of design
Provide training and education
Lead a long-term improvement for zero defects
Encourage cooperation and teams
Satisfy internal and external customers

Habits to foster--Stephen Covey [17], who admonishes us to be proactive, begin with the end in mind, put 1st things 1st, think win-win, seek 1st to understand...then to be understood, and synergize; says, " The problem with TQM in the U.S. is that it only emphasizes structure and systems. What is holding quality back is lack of trust, communications, and patience. " He goes on to say, " Quality improvement is more akin to farming-which is dependent on principles or laws external to it." He contends that leadership-a major component of TQM operates at four levels; personal, interpersonal, managerial, and organizationally. Dr. Covey also points out that the key principle of trust-made up of character and competence-is based on trustworthiness. Trust is the fruit of the act of trustworthiness. [18]

Excellence--Tom Peters, an internationally recognized expert in management and leadership, spotlights " champion " organizations that apply competitive free-enterprise " people and customer-centered " business and management principles, successfully and productively in " Excellence In The Public Sector ", a video filmed by Enterprise Media, Inc. in 1989. The organizations spotlighted are composed of people who work from the heart, on the front line, as change agents both meeting and exceeding customer expectations, and succeeding at risk-taking.

Rules of the Road--Dr. Joseph Juran [19] has identified behavioral " rules of the road " that address the important human side of quality issues needed for successful social and technical systems integration. They are: (1) provide participation to the recipients, (2) avoid surprises, (3) provide time for accommodation, (4) start small and keep it fluid, (5) create a favorable social climate, (6) weave the change into existing culture, (7) provide a *quid pro quo*-something for something, (8) respond positively, (9) work with the informal leadership of culture, (10) treat people with dignity, and (11) keep it constructive-no blame, emphasize benefits and solutions.

THE SYSTEMS APPROACH

" Learning disabilities are tragic in children, but they are fatal in organizations. Because of them, few corporations live even half as long as a person-most die before they reach the age of forty." [20] Peter Senge identifies five core disciplines for building the learning organization. They are: (1) shared vision, (2) mental models, (3) team learning, (4) personal mastery, and (5) systems thinking. [21] The fifth and most important, systems thinking is a conceptual framework, a body of knowledge and tools that has been developed over the past 50 years, to make the full pattern of change clear, and to help us see how to merge social and technical systems effectively. [22] " People enter business as bright,

well-educated, high-energy people, full of energy and desire to make a difference, " says Hanover's O'Brien. " By the time they are 30, a few are on the "fast track" and the rest put in their time to do what matters to them on the weekend. " [23]

Critical Thinking

Leaders within organizations need to encourage and guide their associates to challenge the old assumptions, and after thoughtful consideration and analysis, establish new objectives that involve risk-taking and stretch goals. It is through this mental rigor and process that the "over the hill" gang will maintain the energy and desire to make a difference.

Teamwork

Team learning is vital because teams, not individuals, are the fundamental learning unit of modern organizations. [24] Success is influenced by the actions of everyone else in the system. In order for one to succeed others must succeed as well. [25]

ORGANIZATIONAL CHANGE

Around the year 500 B.C. the Greek philosopher Heraclitus noted that " you cannot step twice into the same river, for other waters are continually flowing on." He was one of the first Western philosophers to address the idea that the universe is in a constant state of flux embodying characteristics of both permanence and change. [26] There are five conditions that prompt change: (1) purpose, (2) alignment, (3) linkage, (4) implementation tools, and (5) anticipation. Purpose is usually expressed in a vision or mission statement. Vertical alignment within an organization or multinational corporation enhances communications and reporting. Linkage ties together horizontally positioned groups that interface to accomplish specific tasks. However, the concept of *atunement* described by Harrison [27] combines vertical and linked subgroups to operate and perform like a well-practiced orchestra. Implementation tools of flow charts, control charts, cause and effect diagrams, histograms, check sheets, Pareto diagrams, and other statistical techniques help to identify opportunities for change. All of these above reasons for change are to no avail unless change is approached with anticipation and excitement.

How To Manage Change

The following elements have important implications for how we understand and intervene in the social and technical processes of organizational evolution.

Power--Empowering ourselves comes from acting on our enlightened self-interest. Politics is the pursuit of self-interest, and positive politics is the pursuit of enlightened self-interest, comprised of the following: meaning, contribution and service, integrity, positive impact on others, and mastery. [28] Establishing personal power through the application of these self-interests, and negotiating and bargaining for scarce resources to be positioned for corporate power fosters the transfer of empowerment to others.

Process--Perception of what is necessary to implement change goes deeper than penetration of the cerebral cortex. Observation of others requires listening, hearing, watching, and feeling their concerns, intents, and cries for help. Mental response to the tides of change only acts on the tip of the iceberg of an issue. To get to the heart of the problem the process of change needs to be monitored and interpreted.

Information--Beyond routine ubiquitous technical data lies the realm of behavioral reality. The use of psychological instruments, open-dialogue, and rites of passage reveal aeons of information about the hierarchical needs, motivation, and preferences of others.

Communication--Unless information flows freely and unrestricted from one to another, there will be gaps. These can be overcome through common courtesies of respect, patience, timeliness, gratitude, and admitting error when it exists.

Consistency--Being predictable and dependable provides the stability needed in times of "high seas". Provide input, process, and reports that are expected, understandable, and correct.

Temporary Structures (Teams)--TQM does not require permanent team members or eternal teams. It does require forming teams of individuals with a vision of what change is, training them in group process, and letting them operate at their level of competence.

Pace--The rate at which change occurs within and outside of an organization needs to be responsive to the environment in which it lives. Moving too slow in an entrepreneurial organization is stifling to ideas and success. Conversely, moving too rapidly in a structured organization defeats the allies and kills the initiative of the change agents.

CASE-STUDY OF TQM IMPLEMENTATION IN THE FEDERAL LANDS HIGHWAY OFFICE

The public sector, or government as it is most commonly referred to, is a prime example of an organization that needs change. Competition from private enterprise is forcing the public sector to alter their old traditional habits of paper shuffling, rules and regulations to a system that searches for the best way to produce their work. In addition to its competition with the private sector, the public sector no longer has the luxury of abundant resources to perform ever increasing workloads.

Government personnel ceilings are being slashed to keep within the budget politicians have promised to their public. Mangers are left with the perplexing question of how to produce quality products with less resources. David Osborne and Ted Gaebler, in their book Reinventing the Government - How The Entrepreneurial Spirit is Transforming the Public Sector, write about case studies of how the government is changing. The authors' definition of a public sector that has the "entrepreneurial spirit" is one that "uses resources in new ways to maximize productivity and effectiveness."[29] The Federal Lands Highway Office (FLHO) of the Federal Highway Administration (FHWA) has adopted the principles of TQM to become one of those "entrepreneurial" public sectors.

An Introduction to the Federal Lands Highway Office

The FLHO performs engineering related services on highways that are on or serve Federal lands. These Federal lands cover one-third of our land area in the United States and provide access to outdoor recreation and places of scenic, historic and cultural interest. The FLHO's major customer is the public whom they service through such agencies as the National Park Service, Forest Service, and Bureau of Indian Affairs.

The FLHO is divided into three field Divisions and a Headquarters office located in Washington, D.C. There are approximately 615 employees within the combined offices. This is a small proportion of the some 3,500 employees that work for the FHWA.

Reasons for adopting TQM

From 1981 to 1982, the FLHO experienced a greater than 200 percent increase in their planning, design, and construction responsibilities due to the start of the Federal Lands Highway Program. However, from 1982 to 1985, the number of employees in their organization decreased by 15 percent. The Intermodal Surface Transportation Efficiency Act of 1991 will double their workload over the next 6 years. Their constantly increasing workload has not been paralleled with an equal growth in personnel ceiling.

In 1986, an Executive Order published a Government-wide productivity improvement program "to improve quality, timeliness, and efficiency of services provided by the Federal Government. Then in 1988 the Office of Management and Budget issued Circular A-132 with the objectives "to make continuous, incremental improvements, and implement quality and productivity management practices in executive departments and agencies." These two documents and a need to gain control of their workload prompted key executives and staff members to meet and discuss a plan of action In October of 1988, an interim coordination group came to the consensus that the management system of the FLHO should be directed toward the philosophies of TQM.

The First Phases of Implementation

The organization began their TQM implementation by developing a new project scheduling system. They realized their old system, that was based on "date guessing", created a climate for reactive or crisis type decision making. A new system was needed to help proactively plan for their increasing program of projects.

To develop the new project scheduling system, teams with members from all the disciplines within FLHO were formed to develop the input into the system. After a series of meetings, the Management Unit Teams (MUTS) developed a "critical path" network or flowchart of the project development activities. In addition, the teams estimated the number of resource hours and durations necessary to complete each of the activities. Developing the system helped management and employees to identify their customers, suppliers, and processes - and most importantly to understand how their entire product is produced.

Implementation Across the Organization

Creating the new project scheduling system set the stage for an organizational wide implementation of TQM. Managers searched for more knowledge about the philosophy by reading literature, watching videos, and networking with other organizations who were in the process of implementing TQM. In addition, forty of their top managers attended TQM management training at the Federal Quality Institute (FQI).

To integrate the "new philosophy" into the organization, the managers adopted the FQI's triangle model of TQM. The three sides of the triangle represent the principles of TQM - achieving customer satisfaction, making continuous improvements, and giving everyone responsibility. Internal building blocks to the structure support the three principles: effective and renewed communication, a long-term commitment, a focus on the customer, rewards and recognition, top management support and direction, commitment to training, employee involvement, and reliance on standards and measures.

The next step was to develop a quality organization structure that would create a successful transition. The structure consists of the following:

1. Executive Quality Council composed of top management from the three Divisions and Headquarters
2. Quality Coordination Team composed of TQM Coordinators from each of the Divisions and a Quality Coordinator from Headquarters
3. Multi-Division Quality Action Teams (QAT) formed to solve problems or improve processes that will affect the entire FLHO
4. Division Quality Councils composed of key personnel within each Division and the TQM Coordinator
5. Multi-Branch QAT teams formed to solve problems or improve processes that will affect an entire Division
6. Branch QAT teams formed to solve problems or improve processes that will affect an entire Branch within a Division

To create a climate of change, the organization developed a vision, outlined their mission, and provided a set of guiding principles:

1. The vision of FLHO: To be recognized as the best highway engineering organization through people dedicated to excellence.
2. The mission of FLHO: To develop and carry out the Federal Lands Highway Program through:

a. Program administration
b. Highway and transportation engineering services
c. Training and development of engineers
d. Development and dissemination of technology

3. The guiding principles of FLHO:

a. *Customer satisfaction*
b. Teamwork
c. Employee commitment
d. Communication
e. Excellence
f. Integrity

After the climate was set, a series of all-employee orientations were held to introduce the new organizational management philosophy. During the sessions, management explained their objectives for implementing TQM. They also fostered the vision, mission, and guiding principles and presented an overview of the quality organization structure. In addition, they introduced the employees to some tools such as brainstorming to take back with them to begin analyzing some of their problems.

Since the orientations, FLHO has developed a strategic plan to ensure the new philosophy is incorporated into their everyday operations. Strategies and actions include the three principles of TQM - customer satisfaction, employee involvement, and constant improvement. The strategic plan of FLHO also embodies the mission, vision, goals, and objectives of the Federal Highway Administration's Vision 2000 planning process.

The transformation of the organization is working and will continue to work because of top management support. They have provided each of the divisions with a budget, adequate personnel, and the necessary time to attend training and meetings. Management has shown their sincerity to make the change.

Strategies for maintaining the momentum of change

The FLHO's strategies for incorporating and maintaining the momentum of change in its organization include: employee involvement, rewards and recognition, development of customer/supplier relationships, training, communication, measurement, and continuous improvement.

Employee involvement--An underlying principle of TQM is that employees closest to the process know the best way to improve that process. Involving employees, therefore, is critical to the success of adopting the philosophy of TQM. Several of the ways the FLHO involves its employees are through the use of teams and a suggestion program.

Approximately 39 percent of our employees are on or have participated on a team. Their accomplishments are publicized in newsletters, posted on bulletin boards, and through group recognitions. Many of these teams cross the three Divisions and their Branches, while other teams are made up of employees within particular Sections.

Our Computer Aided Design and Drafting (CADD) Multi-Division Quality Action Team is an example of one of a currently operating teams. The team has helped the organization transition successfully from a semi-automated computer design and manual drafting environment to CADD. Team members monitor the new system, share information, and solve problems in meetings held throughout the year. The teamwork, especially across divisions, has improved communication and consistency in operations.

The Bridge to Improvement Suggestion Program, a modified version of a program used by the Coast Guard, is another method utilized by the organization to involve its employees. Suggestors may choose to enter his or her suggestion through either the formal or "short-cut" options of the program. The intention of the formal option is to solicit ideas that are of "potentially" significant value. The "short-cut" option is for "smaller, day-to-day" improvement ideas that will help to drive the organization forward in tiny increments. Over a four month period the organization received approximately 60 suggestions.

Rewards and recognition--To encourage employee involvement and participation, FLHO has developed a system of rewards and recognition. However, many of the rewards are still based on individual achievement. The organization recognizes this and is creating a system that will encourage the recognition of team achievements. The FHWA has also established a series of awards that recognizes the accomplishments of both individuals and teams.

Two of the divisions in FLHO have developed peer award systems. The Eastern Division developed the "Eddie" that is an award given by peers to those who prove individual excellence. This program is open to non-supervisory employees. Employees are awarded with a $25 cash award and a certificate. The Western Division has established a similar program called "Cheers for Peers."

Development of customer/supplier relationships--FLHO has incorporated strategies and actions into its operations to develop relationships with its external customers and suppliers. Customer satisfaction, another principle of TQM, will help the organization fulfill its vision and mission. To ensure customer satisfaction, good relationships with suppliers must also be created so that the product will satisfy the customer.

Channels of communication have been opened through meetings and surveys to learn how to satisfy customers. From the inception of a highway project to its final construction, employees stay in continuous contact with their customers throughout the process to ensure the final product meets their expectations. In addition, a survey has been developed to determine how well the customer's requirements have been met after a project has been constructed. Customers are acknowledging their appreciation of FLHO's desire to improve communication and its processes.

Partnering is a concept that FLHO is including in its construction contracts to create a good working relationship with its suppliers.

Through workshops and meetings conducted by a facilitator, the contractors and project managers develop a written agreement that encourages trust, open communication, and cooperation. Partnering agreements will help the organization produce high quality projects that are completed on time and within budget.

Training--Ongoing training is a critical element to the success of embodying the principles of TQM into an organization. Employees and managers should be trained in quality concepts, teamwork, and interpersonal communication skills. In addition, managers need training to develop both management and leadership skills. Employees should also be provided with information and the opportunity to further develop their chosen career path.

The all-employee orientations introduced the concepts of TQM to the organization. Since the orientations, employees have been trained in quality concepts, teamwork, facilitation, problem solving, and statistical process control. Management went through a series of workshops to receive training in various skills such as coaching and participatory management. Currently, the organization is developing a formalized training plan that will ensure all employees receive an appropriate amount of training.

For the past two years the FLHO has organized and presented an annual quality and planning workshop. Key personnel and employees from the three Divisions and Headquarters have attended the workshops. Items that were included on the agendas were strategic planning, measurement, and effective organizational structural development. The greatest benefit has been the transfer of information and the gathering of employees across the organization to discuss and resolve issues and concerns.

Communication--Organizations implementing TQM must ensure there is communication about the results, progress, actions, goals and future plans of the organization. FLHO uses a variety of methods to inform its employees about what is being done throughout the organization to incorporate TQM into its operations.

All-employee meetings are held during the year to recognize accomplishments of individuals and teams, report results, review ongoing efforts, and to discuss strategic planning efforts. In addition, the annual FLHO Quality and Planning workshops have been successful in helping employees to share information across the organization.

The FLHO also uses the FLHO Quality Newsletter as another medium to convey and share quality related information. The first edition was published approximately two years ago. Articles are written by employees about team accomplishments, information received at conferences and workshops, and additional TQM related material. The newsletter is also shared with other organizations interested in the efforts of the FLHO.

Measurement--An organization implementing TQM, must continuously assess its progress toward incorporating quality into its operations. TQM is not a program, it is a continuous improvement process. Employees should be surveyed to ensure there is a perception of improvement. In addition, the TQM implementation process should be measured against a set of criteria such as the U.S.Malcolm Baldrige National Quality Award criteria to determine the progress of the organization.

The FLHO uses a cultural assessment to survey its employees about their perceptions of how the organization is doing as a whole. The 70-question survey is a modified version of a survey developed by the Department of Defense. Over the last three years since the survey has been distributed there has been a 60 to 70 percent return rate with 90 percent of the areas surveyed showing increased scores, that indicate an overall increase in teamwork and employee involvement, participatory management, and employee

awareness of the importance of quality.

To assess its progress toward implementing quality into the organization, the FLHO uses the Quality Improvement Prototype (QIP) award criteria. The award follows a format similar to that of the Malcolm Baldridge Award. After an annual assessment using the criteria, an action plan is developed to improve the deficient areas. The Organization's goal is to apply for the QIP when they have reached the upper percentages of each of the criteria.

Continuous Improvement--Another fundamental concept of TQM is continuous improvement. Organizations need to continually review the processes they use to produce their products. They need to locate activities that do not add value to the processes or that are sources of rework within the processes. The FLHO is currently acquiring a consultant to develop teams to assess their processes.

The project management procedure will ensure customer requirements are met, the processes are brought under control, and measurement systems are initiated. A stakeholder session is held at the beginning of the procedure to assess the process, identify the process owner, charter the team, and clarify expectations. A project management team will then meet for five sessions during which they will define the process boundaries, investigate customer requirements, identify lack of control areas, and develop measurement points. After a period of data collection, the team will review the measurement trends and decide on an improvement action plan.

Concluding remarks

FLHO began implementing TQM in 1988 and is approaching the location in its journey where TQM will become institutionalized. It has not been an easy journey. There have been barriers such as resistance to change, the inability to involve everyone in the effort, and a belief that "this is only management's job." However, the FLHO is dealing with the barriers by helping employees to understand how they will benefit from this effort in the future. The FLHO realizes that incorporating TQM into its operations does "not end the journey", but that it provides a framework for a high quality organization.

REFLECTIONS

Based on the aforementioned status of the European Community coalescence, the activity in international standardization development and use, the movement away from more traditional organizatonal frames, the practiced principles of TQM, the learning organization, and increased acceptance of the characteristics of change, I offer the following reflections.

Optimization of Technical and Social Systems

Since the advent of the Industrial Revolutions in Europe and America, technical inventions, developments, and emphasis have tended to separate the entrepreneurial spirit of innovation from the more mechanical business of production and quality control. This chasm has worked against any motivation for trying new ideas, unless they have come from the "aristocracy". Eric Trist recognized that the gap between people and machines, and the divisions of labor, needed to be closed for organizational health and survival. With the new developments in total quality and management, teamwork, empowerment, commitment by all levels within the organization, and recognition of "change agents" as change influencers and directors, the walls of structured bureaucratic organizations have been damaged by human resource models, and more symbolic metaphors. However, since scarce resources and "rightsizing" of

organizations have significantly impacted the modern organization, it has become even more crucial and necessary to integrate technical and social systems for organizational adaptation, both nationally and internationally.

TRENDS OF THE 90s

" Today's business world bears scant resemblance to the boom economy that once prevailed in American industry. Companies that loathe defeat are redesigning their business processes [and people interfaces] from the inside out ". [30] Dr. Michael Hammer, president of Hammer and Company, says that, " Instead of limited supply, we now have limited demand. As a result, the issues for the foreseeable future are...the three Cs: customers[who]have the upper hand, intense competition, and the phenomenon of change-where whatever works today is going to be irrelevant tomorrow ". [31] How do we prepare for this reality ? Dr. Hammer suggests "re-engineering" which includes the following steps:

1. Task integration and compression of the process, where one person is given responsibility for an entire process whenever possible.
2. Decision making as part of the job, where many decisions are made by the person doing the work.
3. A hybrid of centralized and decentralized approaches, where common pools of information, accessible to all far-flung operations of a company, combine economies of scale and hands-on expertise with customers.
4. Concurrent performance of tasks, where work is done collaboratively and concurrently with communications networks, shared data bases, and teleconferencing.
5. Capture information once, at the source, where it is entered into an online data base once, through connected, integrated systems, and used over and over again.

However, Dr. Hammer alerts us to the fact that reengineering is not for the timid, it requires executive leadership and real vision. [32]

Virtual Reality

Another trend for the 1990s and beyond is virtual reality where users can really get "into" their work. " Virtual reality creates the illusion of being immersed in an artificial world or of being present somewhere else in the physical world. VR, as it's known to researchers, is the convergence of 3-D display technologies, computer graphics and head mounted computer displays...." [33] It is being pursued at research facilities for applications in radiology, architecture and design, telerobotic control, medical imaging and scientific visualization, and the entertainment industry. [34]

Computer Network/Data Highway

" Corporations, research labs, universities, and medical centers [ASTM, the American Society for Quality Control, and others] will interface through a national data highway transmitting visual and audio image thousands of times faster than today's fastest networks. These synergistic links between myriad scientists, scholars, government officials, and businesspeople should catalyze an information explosion profoundly transforming the way we live. Such a supernet could allow anyone on the data highway to harness up the power of supercomputers and, at least, provide users with calculations for complex applications such as climate modeling, stock-market analyses, cosmological research, [ASTM precision and bias statements], and medical diagnoses and treatment ". [35]

Standards Around the World [36]

Standardization around the world, influenced or not by the ISO movement, is occurring at a breakneck speed. The Singapore Institute of Standards and Industrial Research (SISIR), the Standards Institution of Israel, the Hungarian Office for Standardization, the Lithuanian State Standardization Office, the Ethiopian Authority for Standardization (ESA), the Standards Association of Zimbabwe (SAZ), and the Bureau of Indian Standards (BIS) are; becoming more industrialized, helping local industry grow, drafting new standardization policies, assisting manufacturing in increasing quality management, increasing productivity, reducing waste, providing information and training, and assisting in the transfer of technology.

CONCLUSION

" The "T" in total quality management (TQM) focuses on using ISO 9000 [and 10000] series standards and tools of quality, Deming's 14 Points, Juran's 10 steps, Crosby's 14 tenets...,[and Victor E. Kanes simple statistical tools for defect prevention [37]]. The "Q" relates to quality councils established to form vision and mission statements, teams, awards, celebrations, and recognition. The "M" focuses management on strategic thinking, critical analysis, management style, organization development, action research, and change agents ". [38]

" ... When the real basis of competitive advantage is people, a shared vision on the part of people forms a timeless competitive barrier. An internalized vision, versus a product focused vision, of service results in a long-term competitive advantage. The results of an organization driven by vision are characterized by the following ": [39]

1. Trust and mutual respect
2. A sense of ownership/partnership
3. Leadership vs. management
4. Tolerance of differences
5. The organization as a verb
6. People vs. position
7. Innovation and risk-taking

Those who understand the inferences of behavior, the context of the organization, and derive power from the technical and social systems they lead, as opposed to the hierarchical position they occupy in the firm's structure, will effectively lead their followers into the 1990s and beyond.

REFERENCES

[1] Trist, E.L., The Professional Facilitation of Planned Change in Organizations, Reviews, Abstracts, Working Groups. XVI International Congress of Applied Psychology, Amsterdam:Swets and Zeitlinger, 1968, pp. 111-120. In McLennan, R. Managing Organizational Change Englewood Cliffs, NJ:Prentice-Hall, 1989, p. 42.

[2] Marquardt, D., Chove', J, Jensen, K.E., Petrick, K., Pyle, J., and Strahle, D., " Vision 2000: The Strategy for the ISO 9000 Series Standards in the 90's, " Quality Progress, Vol.24, No. 5, May 1991, pp 25-31.

[3] Riddle, J., and Washburn, N., " Accelerated Microelectronic Contamination Standards, " ASTM Standardization News, Vol.19, No.11, November 1991, pp 37-39.

[4] Morgan, G., Images of Organization, Sage Publications, Newbury Park, California, 1986, p 343.

[5] Ibid, p 275.

[6] Ibid, p 25-26.

[7] AMA Management Handbook, Second Edition, William K. Fallon, Ed., AMACOM, New York, 1983, pp 1-4 and 1-5.

[8] Severson, D., " Collaboration, Cooperation, and Celebration, " Quality Progress, Vol.25, No.9, September 1992, p 63.

[9] Meyers, I.B., Gifts Differing, Consulting Psychologists Press, Inc., Palo Alto, CA, 1980.

[10] Bolman, L.G., and Deal, T.E., Reframing Organizations, Jossey-Bass Publishers, San Francisco, 1991.

[11] Ibid, pp 43-47.

[12] Ibid, p 121.

[13] Ibid, pp 183-186.

[14] Ibid, pp 244-245.

[15] Kanter, R.M., The Change Masters: Innovation For Productivity In The American Corporation, Simon and Schuster, New York, 1983, pp 35-36.

[16] Saylor, J.A., The TQM Field Manual, McGraw Hill, New York, 1992.

[17] Covey, S. , The Seven Habits of Highly Effective People, Simon and Schuster, New York, 1987.

[18] Comments made by Dr. Stephen Covey at an address to Regis University in Denver, Colorado in October 1992.

[19] Juran, J.E., Juran On Leadership For Quality, The Free Press, New York, 1989, p 318.

[20] Senge, P.M., The Fifth Discipline-The Art & Practice of The Learning Organization, Doubleday Currency, New York, 1990, front cover flap.

[21] Ibid, p 12.

[22] Ibid, p 7.

[23] Ibid, p 7-8.

[24] Ibid, p 10.

[25] Ibid, p 50.

[26] Morgan, G. (1986), p 233.

[27] Harrison, R. " Leadership and Strategy for a New Age ", 1982. In Adams, J.D., Transforming Work, Roger Miles River Press, Alexandria, VA. 1984, pp 101-102.

[28] Block, P., The Empowered Manager: Positive Political Skills at Work, Jossey-Bass Publishers, San Francisco, 1987, pp 86-92, 105.

[29] Gaebler, T. and Osborne, D., Reinventing Government-How The Entrepreneurial Spirit is Transforming the Public Sector, Addison-Wesley Publishing Company, Reading, MA, 1992, p xix.

[30] Hammer, M., " Making the Quantum Leap ", Beyond Computing, International Business Machines Corporation, Armonk, NY, March/April 1992, p 10.

[31] Ibid, p 12.

[32] Ibid, p 14.

[33] Rheingold, H., " How Real Is Virtual Reality ? ", Beyond Computing, International Business Machines Corporation, Armonk, NY, March/April 1992, p 26.

[34] Ibid, p 27-29.

[35] Marsa, L., " An Interview with Robert E. Kahn: On cyberjockeying the electronic grids of the twenty-first century ", OMNI, OMNI Publications International Ltd., New York, Vol. 15, No.3, December 1992, p 83.

[36] Schindler, B., Editor-in-Chief, " Standards Around the World ", ASTM Standardization News, Philadelphia, Vol. 20, No. 12, December 1992, p 1.

[37] Kowalewski, M. J., a book review of Defect Prevention System by Victor E. Kane, " ASTM Journal of Testing and Evaluation ", Philadelphia, Vol. 19, No.1, January 1991.

[38] Van Nuland, Y., " The New Common Language for 12 Countries ", Quality Progress, Milwaukee, Vol.23, No. 6, June 1990, p 40.

[39] Shaw, J., The Service Focus, Dow Jones Irwin, Homewood, IL, 1990, pp 198-202.

Larry H. Bernstein, M.D.[1]

REALIZATION OF THE PROJECTED IMPACT OF A CHEMISTRY WORKFLOW MANAGEMENT SYSTEM AT BRIDGEPORT HOSPITAL

REFERENCE: Bernstein, L. H., **"Realization of the Projected Impact of a Chemistry Workflow Management System at Bridgeport Hospital,"** Quality and Statistics: Total Quality Management, ASTM STP 1209, Milton J. Kowalewski, Jr., Ed., American Society for Testing and Materials, Philadelphia, 1994.

ABSTRACT: Any approach to improving the quality of laboratory operations by data management has to be concerned with its effect on patient test management. For the laboratory to be effective, instruments have to communicate with computers that improve the speed and efficiency of the work in a real-time distributed-processing environment. Chemistry services are highly information-intensive in terms of quantity, variety and rate of generation.

We used a systematic process with defined and measurable outcome expectations to restructure automated chemistry. Our goal was to improve dissemination of medical information. We had spectacular success, allowing effective reassignment of staff to support our service base. This study uses laboratory and hospital trends for staffing, work volume, workload statistics, and carefully designed time studies to analyze the crossovers in replacing chemistry analyzers and modifying a computer interface to realize automatic results verification with callup of results outliers for review.

In 1987 work in automated chemistry was mainly done on two of five analyzers supported by six technologists on three shifts. We considered replacement of these five with two instruments and constructed a model comparing the measured characteristics of the existing system with its replacement with a projected savings of $1.2 million over seven years. Measurements were made on an intermediate, and on the final installed system, and on the contribution of technical and nonlabor expenditures to the expected savings. The expected savings are adjusted to $1.5 million over seven years based on a 30% increase in testing. 22% of the time eliminated is in test generation and 30% is accounted for by interface changes with an effect on information transfer, editing, and rechecking the results generated.

KEYWORDS: Technology replacement, workflow management, laboratory productivity, cost impact analysis.

* * *

Bridgeport Hospital's full-service laboratory serves the fourth-largest metropolitan, voluntary, not-for-profit hospital in Connecticut. The lab offers the most comprehensive services in the southern part of the state. In an increasingly difficult operating environment, the hospital had to downsize in staffed beds in the mid-1980's and reduce

[1]Section Chief, Chemistry, Department of Pathology, Bridgeport Hospital, Bridgeport, CT 06610

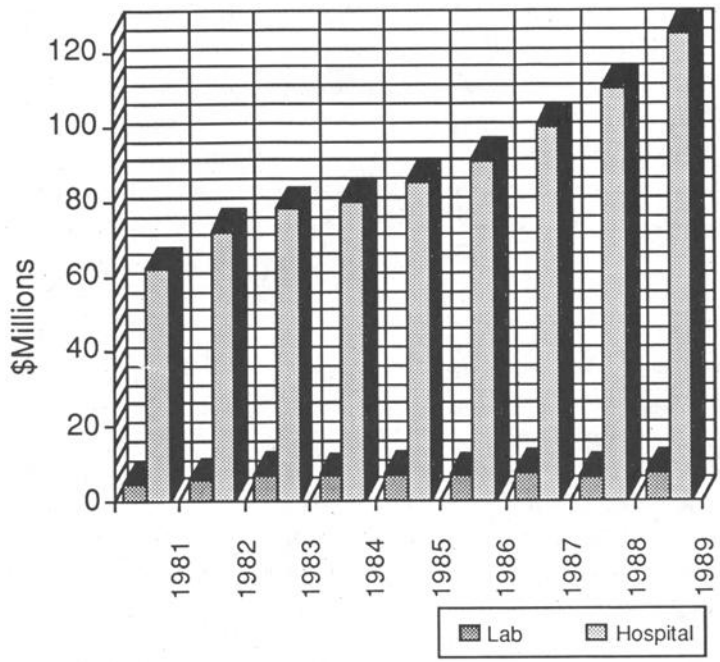
length of stay per admission. These were a response to Medicare's "diagnosis-related groups" payment system, which provides reimbursement on a mostly per-stay basis, rather than per-diem.

A rising median age of the hospital population and the increasing burden of uninsured patients--with critical medical conditions associated with drugs, trauma, and societal disruptions--added to the financial risk incurred by the hospital in the mid-1980's. These costs could be compensated partly by pass-throughs for medical education programs and partly by increasing outpatient revenues, but other parts of the hospital had a role to play, too.

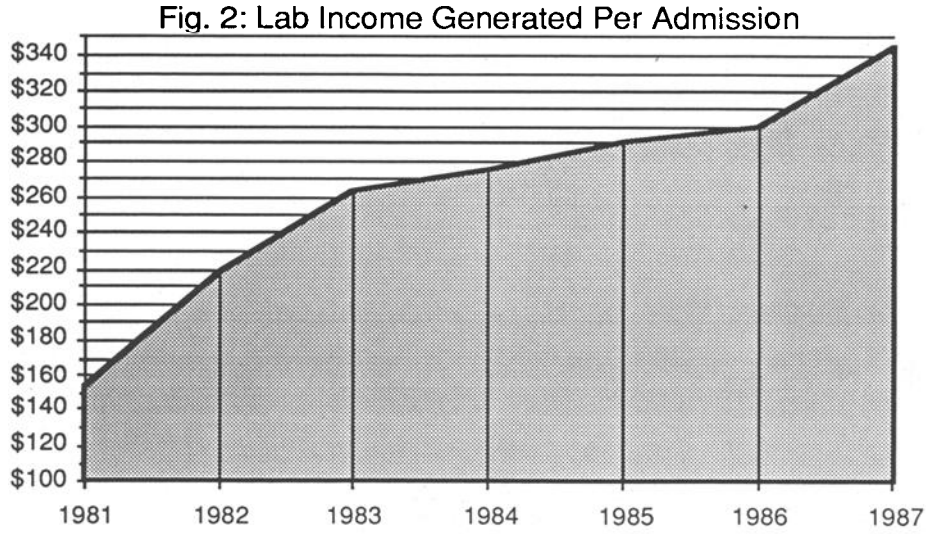
Lab Competitiveness

The hospital's laboratory, as a profit center, contributed substantially to maintaining the competitiveness of the hospital over this period. Overall hospital costs rose faster than lab costs (Figure 1), particularly in the years 1986-1989.

Fig. 1: Lab & Hospital Annual Expenses



In the lab, cost-per-test was maintained at or below $10 for labor; consumables and reagents; and acquisition costs throughout the decade. Laboratory costs or revenues per admission rose (Figure 2) due to increased intensity of laboratory use in this shrinking financial environment, selecting for the sickest patients, with a requirement for more testing and shorter reporting turnaround times.



The laboratory, especially its clinical chemistry component, managed to keep pace by repeatedly reorganizing its services and equipment to meet current and anticipated requirements. In 1984 to 1986, the lab changed. Its limited profiling instrument and five other instruments (staffed by eight technologists on three shifts) were replaced.

New Equipment

The new equipment were: a high-throughput, 8700™ [2] analyzer operated on one shift; an ASTRA™ Mini-Ideal[3] operated on the day-shift and on the evening/night shifts for reporting within an hour of specimen receipt; and an Ektachem™ [4] 400 routine, discrete, non-STAT, non-profiler, operated on all shifts. An "aca"™ [5] and Cobas™ Bio[6] were utilized on all shifts for a limited number of tests.

The Roche Cobas Bio--the instrument relied on for cardiac enzymes / isoenzymes for evaluating myocardial infarction--required nearly a half-staff equivalent to support its work. Its replacement was planned in selecting the next instrument system, which would allow on-line, continuous rather than multiple batch testing to support the cardiology

---

[2]Boehringer Mannheim Diagnostics, Inc., Indianapolis, IN.
[3]Beckman Instruments, Brea, CA.
[4]Eastman Kodak Co., Rochester, NY.
[5]DuPont Clinical Systems Division, Wilmington, DE.
[6]Roche Diagnostics, Inc., Nutley, NJ.

service. These required a total of seven technologists to handle most of the routine chemistry work. The chemistry volume in 1987 was 1.2 million reportable tests, of which 900,000 were handled by two instruments (BMD 8700 and Kodak Ektachem 400). Even though there was no dramatic change in staffing, the laboratory kept pace with a much higher number of test requests and increased variety of services. We improved the generation and reporting of chemistry profiles and urgent test requests; expanded the load capacity of its capital base; created a previously nonexistent protein laboratory that evaluated dysproteinemias; supported a perinatal service that was expanding substantially by 1986; and supported an active cardiovascular program.

Information System Bottleneck

When the hospital's first laboratory information system (LIS) was installed in 1985, there was a significant bottleneck. The interfaces between the LIS and the instruments were not capable of matching patient identification with the specimen. Aligning the patient with the test results required substantial technologist intervention [See references 1, 2 last page]. There was also a problem in handling timed-sequential requisitions with longitudinal review and reporting such as cardiac isoenzymes.

We did a time study to examine the effect of a computer interface on degradation of test turnaround time and technologist productivity. In comparing the analysis on the Beckman Astra with and without the interface using 16 specimens for glucose, urea nitrogen, and electrolytes, we found a tenfold delay in turnaround time due to keystrokes (28 vs 290 seconds). The effect on the high-output profiler (BMD 8700) was as great, and that on the Kodak Ektachem 400 was least. Specimen identification numbers had to be entered; editing and release of results were cumbersome [1, 2]. It was necessary to reorganize the workflow to accomodate the LIS.

Specimens had to be split and saved if they were received late. Only the most urgent part of the chemistry panel was run; the remaining specimen was processed the next morning.

In fact, splitting and saving accounted for half of the specimens received. Reruns required 0.3 staff equivalents (FTEs) to support the process and generated 12% nonbillable test results. Even though a principal instrument was near its depreciation schedule and had to be replaced, the replacement had to be considered in the context of a different need than existed at the time of its installation.

This study takes into account a review of services, staffing, specimen, reagent and consumable volumes, and cost of equipment through purchase or lease in the 1980's.

"Mass" vs. "Lean" Production

It became apparent that test turnaround times are a byproduct of workflow management. We had to find a method to transition from a "mass production" to a "lean production" process in order to eliminate this substantial reprocessing of specimens and the more-than-8-hour delay in completion of test requisitions.

There are three phases in the analytical process: preanalytical, technical processing, and postanalytical. Improving the efficiency of each requires the effective application of computer technology. Extensive use of bar-coding is necessary to improve the first (pre-analytical) phase--the steps taken to order tests, collect specimens at bedside, and accession the work. Extensive use of windows technology and artificial intelligence on-line is required to support the last (post-analytical) phase, which mainly involves the review and release of accurate results. Our concentration in this discussion is largely, but not solely, with the middle phase--the purely technical analysis itself. In fact, we achieved the greatest improvement post-analytical.

Key Quality Issues

The following sequence was delineated for improvement in the technical process: log-in of specimens at workstations; set-up time; review of out-of-range flags; sending of results; start-up and re-runs; and maintenance. Key quality issues pertained to two areas: concurrent bidirectional interfacing of the instrument to the information system with reduced technologist intervention; and test validation and release applying an autoverification algorithm.

This autoverification is the on-line release of test results within assigned ranges. Results exceeding those limits have to be reviewed prior to release or rerun. Recalibration and non-billable costs had to be reduced to less than 5% and turnaround times for 85% of all testing had to be in less than three hours without degradation of rapid test capabilities (20-minute response in critical cases).

Reorganization, again

In 1988, we reorganized the chemistry laboratory again. We sought to reduce nonbillable work and improve operating efficiency in this cycle of equipment replacement.

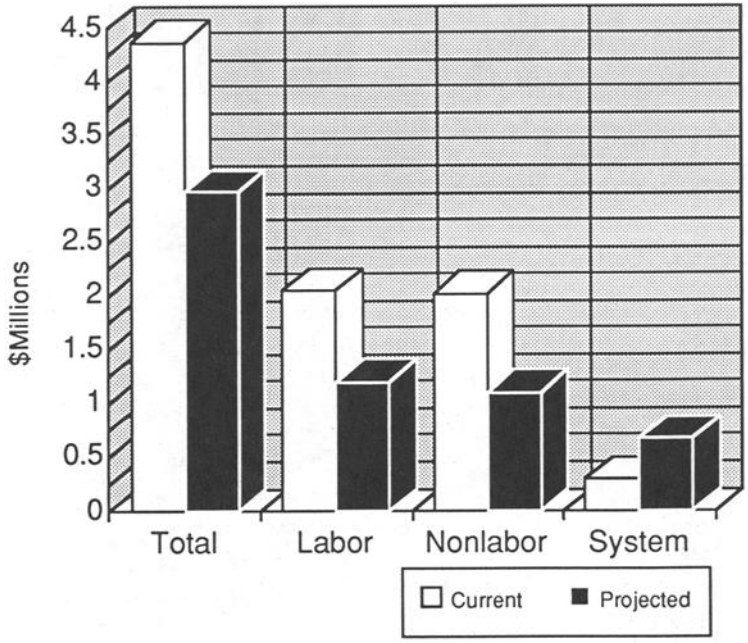
We created a model based on the components of test volumes for each instrument. This included costs for labor, nonlabor (reagent and materials), and system support based on comparing the existing and replacement options.

We projected the cost of a replacement system using alternative configurations based on the elimination of the workflow bottlenecks described [1, 2].

We estimated a monthly savings of $5,700 (34%). Over five years, this was projected to be $345,000 (34%) in labor including a 25% reduction in reagent expenditures.

Figure 3 compares the seven-year labor, nonlabor, and system costs ($ million) for maintaining the existing vs the projected replacement system.

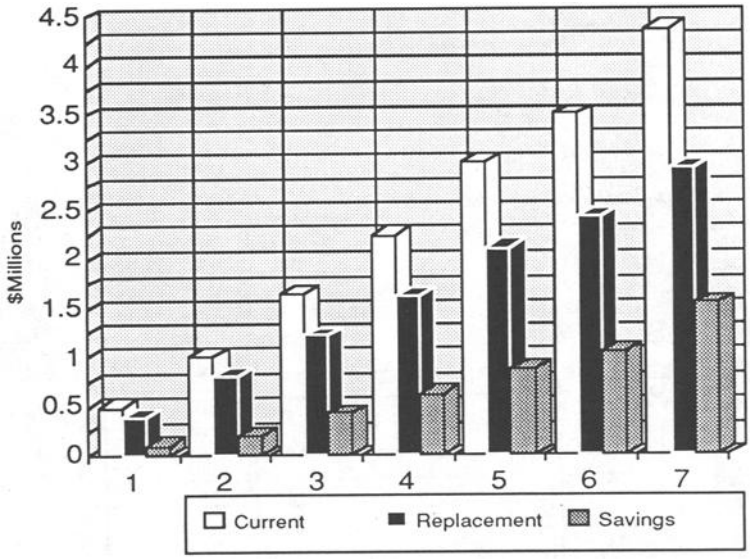
Fig. 3: Seven Year Cost Of Two Clinical Systems



Actually, keeping the existing system was not an alternative because it was not possible to continue to expand the demand for service and maintain the existing system. Nevertheless, it illustrates the cost benefit of the projected technology replacement.

Figure 4 is the cumulative seven-year expense comparison and savings for the existing vs the projected replacement system. This expense includes the costs of the two systems being compared.

Fig. 4: Seven Year Cumulative Savings



In 1989 we carried out parallel evaluation of a replacement for the high-throughput chemistry analyzer. We developed criteria which, if met, would permit a two-instrument system with the desired characteristics.

Initial Attempt Failed

An initial attempt at replacement failed. After a seven-month evaluation of an instrument installed on a trial basis, we were unable to approach the turnaround times necessary to provide adequate routine/rapid testing support. The instrument, expected to produce 900 tests an hour and to require minimum recalibration and rerunning of specimens, had a throughput of only 500 an hour and had an excessive rerun/recalibration rate. The system had a computer interface that allowed transmission of results to the laboratory information system and

automatically tied specimens to the patient identification. This had been an insurmountable problem with the system it was intended to replace. But the new system could not screen for results that fail defined criteria. We were also unable to eliminate the splitting and saving of specimens on late shifts, so the bottleneck in the morning processing of tests continued.

The trial system was rejected and an alternative system was reconsidered after the trial supplier was unable to respond to the problems identified.

In considering the new replacement system, we looked specifically at the interface requirements and reviewed turnaround time, projected labor, reagent and system requirements. For instance, it was essential to be able to do testing at a rate of at least 700 tests hourly and to be able to continuously load specimens, even testing out of order, and have acceptable turnaround times. In order to eliminate the individual call-up and review of 85% normal results on screens, we recruited a third party to provide and implement a unique information system interface allowing for a rules-based selection of results electronically screened online.

New Alternative

We again carried out a comparison between the system that was ultimately rejected and what became our current system, the Kodak Ektachem 700XR. Features that were critical for the operational comparison were a random-access capability for handling out-of-order tests without degrading performance and direct sampling without use of pour-off cups. The amount of repeat testing was used as a measure of reagent quality affecting the outcome. Editing and release of results on screens was also measured.

We carried out such a comparison of the performance of two systems and ran 40 specimens and 410 tests. Trials were run both prior to and after installation of the interface. The most important and difficult objective of the analysis was to realize the projected improvement in the final installation. It was difficult not only because of the complicated transition, but also because with the installation of the needed interface this change could not be adequately measured. Review of workload and productivity reports was used to assess this effect. The times measured were both the technical (that which required hands on FTE support to run the system) and non-interventional, mostly in: calibration, pour-off of specimens, editing and repeating tests, and

sending results. In reviewing the significant time savings, 74% was in calibration, editing and repeat testing (62% of this was in nontechnical labor); 22% was in running the instrument; and 31% was a consequence of the interface.

There is no reason that these components should total 100%. They themselves are functional categories obtained by adding essential steps defined in the processing sequence. The analysis illustrates that 25-40% of a technologist's time is actually required either for running the instrument or for processing specimens. More than half the remaining processing time is not necessarily controlled by the technologist. This time is technology-driven. Nevertheless, inefficiencies in system performance have a labor-cost penalty that isn't a part of traditional cost-benefit analysis in a capital purchase decision.

The unique interface requirements were a major issue because of the planned reduction from a three- to the two-instrument configuration already described, and the associated reduced labor as well as lower capital requirements.

TABLE 1: OPERATING SYSTEM INSTRUMENT TIME COMPARISONS*

ANALYZE	**Technical**	**Nontechnical**	**Total**
System	83.6	266	349.6
Replacement	19.0	171	190.0
Saved	**64.6**	**95**	**159.6**
INTERFACE	**Technical**	**NonTechnical**	**Total**
System	129.2	140.6	269.8
Replacement	19.0	0	19.0
Saved	**110.2**	**140.6**	**250.8**

*(Minutes Per 24 Hours)

Table 1 compares the operating characteristics of the two systems in minutes per 24 hours and the savings from running the analyzer with the interface. The dramatic improvement in system performance is explained by the following: reduced hands-on requirement to operate, random-access capability, elimination of nonbillable tests, and creation of an electronic editing function.

Productivity Gained

Technical and nontechnical labor per 24 hours to support the new configuration was 3.55 hours compared with 15.6 hours. 62% of the productivity was gained in elimination of nontechnical time. This is equivalent to 4,441 hours per year. The productivity gain is 15.8% of the workload measured by 1992 productivity reports showing 28,144 required hours for the chemistry section.

This means that it would have required an additional three full-time technologists to support the work at the preinstallation level of support. Part of the productivity gain has been incorporated into 1.5 staff equivalent of labor support for- an operating room laboratory, a same-day surgery test site, and near-the-patient glucose monitors. We also increased the existing support for the perinatology program, while adding a service for nutrition support monitoring. These were developed without plans for additional staffing.

TABLE 2: PROCESSING STEP IMPROVEMENTS

Step	Improvement
Specimen ID	Bar Code
Specimen Validation	Check Against Previous Value
Edit/Review	Autoverify
Logging/Transcription	Eliminated
Repeats/Recalibration	Reduced
Routine/STAT TAT	Reduced
Specimen Splitting	Reduced

Table 2 lists the changes in processing steps that were made by replacing the automated chemistry system.

Compared with Peer Hospitals

The results of the reliance on planned technological replacement for Bridgeport Hospital have not gone unnoticed. The Arthur Andersen Consulting Group compared Bridgeport Hospital with peer hospitals of similar size and complexity in a 1990 study. They found that Bridgeport Hospital's 160% greater workloaded hours and 62.6% fewer laboratory

staff was associated with a 2.55-fold greater work/staff ratio (shown in Table 3).

TABLE 3: HOSPITAL LAB LABOR COST COMPARISON: BRIDGEPORT VS. SIMILAR CONN. FACILITIES

Units	Bridgeport Hospital	Group A Hospitals	Difference
CAP Hours	162,539	101,741	60,800
Lab FTEs	120	191.7	71.7
Units/FTE	1,354	530	824
FTE Cost (@ $15.86/Hr.)	$3,960,000	$6,340,000	$2,380,000
Output (@ $15.86/Hr.)	$2,580,000	$1,610,000	$960,000
Output vs. Expense	65.1%	25.5%	37.62%
Expense Per CAP Hour	$24.36	$62.28	$37.92

The Arthur Andersen study did not examine the relationship of staffing to the method of correcting nonstandard workloaded hours.

A 1991 study of the Massachusetts Hospital Association Advanced Management Systems included Bridgeport Hospital among the most productive with an average worked/paid ratio of 87.5% compared with an expected 87.8% for this group (Best of the Best). The Arthur Andersen findings were actually comparable because at an efficiency of 65% for CAP workloaded hours, the correction for total workload hours would be 90% in 1990.

It is of some interest that these studies are comparisons of all laboratory performance. However, despite a high productivity that exists throughout the laboratory, that achievable in microbiology, blood bank, and hematology is considerably more restricted than for chemistry. The hospitals compared have similar analyzer systems. A major difference in configuration is the use of an intermediate computer between the LIS and the chemistry analyzers.

Analyzer, Interface Synergy

An enormous productivity gain has been realized as a result of the analyzer system with a distinct synergy derived from the interface. On the basis of the splitting and saving of specimens, one can estimate a half-staff position reduction. In addition, the projected savings of $1.2 million over seven years could not take into account an

unanticipated 30% increase in testing. When corrected for volume, savings would be $1.5 million.

The increased volume associated with utilization is not necessarily a reflection of wasted resources. There was a better awareness of the improved turnaround time with the completed conversion. It became apparent that more complete test orders would provide more information sooner and would reduce repeat phlebotomy on the same patients.

However, the volume added could not have been handled with the previous equipment configuration, and not simply because of an estimated staff saving per analyzer and another for the interface. The configuration is changed so that there is a substantial increase in the loading capacity so that perhaps only half its capacity is being utilized without any negative impact on services. This would allow for substantial growth in testing, including the ability to bring in outside non-hospital revenues, without additional staffing or capital investment.

Most Significant: "Autoverify" Function

The most significant factor in accelerating this workflow enhancement is the "autoverify" function. The autoverify function has assigned criteria for accepting data that falls within medically defined limits and extracts outliers for review by the technologist eliminating the one-by-one reiterative process.

This could account for another 0.5 staff equivalent gain in using the system. The data showed the importance of this function in reducing technologist interaction in nontechnical work. Only 30% of the effort is now required in reviewing data prior to release and in retesting. Based on these findings, we recommend that lab managers consider this system function.

Conclusion

This study shows how we benefitted by using each replacement of technology to examine the mix of skilled technologists and automated systems to provide more, and more complex, services in a community teaching hospital. We have had a philosophy of optimizing for both the volume and the rate of data generation in automated chemistry.

The only way to effect the necessary change was to examine the rate-limiting steps in the process. The amount of technologist intervention in entering tests for analysis, in validating correct

specimen identity, and in results review were costly impediments to an efficient service which we have minimized.

The comparison with other institutions has some value, even allowing for some differences in accounting for staffing requirements, for evaluating the competitive position of the laboratory. However, the basis for performance has to be predicated on an objective and measurable assessment of expected vs actual outcome, not on peer comparison. The greatest obstacle to achieving this discipline is the need for adequate collection of data about the operational environment. This is the relevant information that must be available for activity-based cost accounting decisions [3].

These findings are part of a continuing ongoing process of improvement in my laboratory. I previously reported that the introduction of an LIS can introduce a rate-limiting step into the laboratory workload processing sequence. There will be a mismatch of performance vs expectations if there are incompatibilities between the system design and implementation requirements--especially in this transaction-intensive environment [4]. An analysis of the workload processing sequence makes it possible to match performance against expectations allowing for adjustments at any point in the process [4, 5].

Acknowledgements: Special recognition is made to Ronald Brouillette, Associated Press, Inc. whose financial model was validated by this experience, and to Terence McArdle, Eastman Kodak Company, in manuscript editing.

References:

[1] Bernstein, L.H., Sachs, B., Stapleton, V., Gorton, J., Lardas, O., "Implementing a laboratory information management system and verifying its performance", Informatics in Pathol 1986; 1:224-233.

[2] Bernstein, L.H., "An LIS is not all pluses", MLO 1986; 18 (11):75-80.

[3] Kaplan, R.S., "Management accounting for advanced technological environments. Science 1989; 245:819-823.

[4] Bernstein, L.H., "The contribution of laboratory information systems to quality assurance", Amer Clin Prod Rev 1987; 6:10-15.

[5] Bernstein, L.H., "Predicting the costs of laboratory testing", Pathologist 1985; 39:19-21.

Kwok W. Yeung,[1] and Gail Taylor[2]

IMPLEMENTATION OF A QUALITY ASSURANCE SYSTEM IN CLOTHING ESTABLISHMENTS

REFERENCE: Yeung, K. W., and Taylor, G., "**Implementation of a Quality Assurance System in Clothing Establishments,**" Quality and Statistics: Total Quality Management, ASTM STP 1209, Milton J. Kowalewski, Jr., Ed., American Society for Testing and Materials, Philadelphia, 1994.

ABSTRACT: The clothing industry is the most important manufacturing industry in Hong Kong. In an effort to cope with recent changes and developments, there are signs that the clothing establishments are becoming more quality conscious. A recent survey however reveals that these establishments are still not paying enough attention to this important aspect of their business and a proactive, innovative approach to quality improvement is less common than responsiveness to market forces on a needs basis. Based on a case study of implementing the ISO 9000 quality assurance standard in a clothing establishment, areas of interest and potential problems have been identified which are unique to the clothing industry, and solutions have been suggested to how these could be overcome.

KEYWORDS: clothing industry, quality improvement, quality assurance system, ISO 9000, BS 5750

1. INTRODUCTION

Hong Kong is the second largest exporter of clothing in the world, ranking after Italy in trade significance. Domestic exports worth HK$ 55,229 million were exported in the first 9 months of 1992, representing a 3% increase over the same period in 1991. Clothing is the Territory's biggest export earner, and the single most important industry sector for employment. When re-exports are combined with clothing exports, Hong Kong is the leading supplier of garments to international markets. Value of re-exports for the first 9 months of 1992 equalled HK$ 56,956 million, an increase of 26% over the same period in 1991.

1.1 Quality Awareness in Hong Kong

For reasons such as safety and serviceability, quality of clothing products is sometimes considered by consumers of lesser importance than for products such as home appliances and electronic goods; as a consequence, initiatives for registration of clothing factories through local and international bodies have been slow to arise. Nevertheless, the situation is improving: between February 1990 and November 1992, the Hong Kong Quality Assurance Agency recorded 32 out of a total of 720 ISO 9000 enquiries from the textile and clothing sector, and a number of these firms are pursuing registration.

[1] Head, Institute of Textiles and Clothing, Hong Kong Polytechnic, Kowloon, Hong Kong

[2] Principal Lecturer, Institute of Textiles and Clothing, Hong Kong Polytechnic, Kowloon, Hong Kong

One reason why Hong Kong clothing industrialists are paying more attention to the quality of products is competition. Other manufacturing bases are increasingly eroding Hong Kong's share of the international market, in addition to which, outward processing arrangements designed to combat such competition necessitate more formalised, standardised means of 'control', e.g., for production carried out in subsidiary companies.

A related, second reason is quota: the quantity of exports to most of Hong Kong's markets is fixed and, since manufacturers may not increase business in terms of items produced, they must secure more profits by moving up-market.

Quick Response is a third, important reason for increased emphasis on quality systems. In traditional manufacturing the time frames for planning, manufacturing and distributing products permitted greater leniency in such matters as rework and the communication of changed requirements; current market conditions, however dictate that all personnel involved in the textile-apparel activities chain be far more efficient.

1.2 Industry-specific Inhibitors

Organisations such as the Hong Kong Productivity Council, Hong Kong Quality Assurance Agency, and private consultancy firms have recently accelerated their efforts to increase awareness of quality in Hong Kong, and a course was recently organised for the registration of local assessors, potentially simplifying the situation for ISO 9000 certification. Nevertheless, no garment companies in Hong Kong have applied for such certification to date.

Many papers [1-9] have been published in the study and analysis of the implementation of the ISO 9000/BS 5750 quality system standard in the manufacturing industry. There is however only a small number written specifically for textile and clothing firms, or alternatively to highlight accreditation and its benefits on a "case basis".

Problems that are commonly found in small companies relating to quality systems of the manufacturing sector can be summarised as follows:

1. Lack of documented work instructions and procedure--The functional, usage, fit and packaging requirements require translation into work instructions and procedures, for example in the production site, which may give rise to ambiguities and increase levels of variability to a point where they become unacceptable. Confusion over customers' emphasis on differing quality characteristics further complicates the documentation process.

2. Only the minimum of records are kept--Though it is typical for larger firms to retain records of every stage in production, smaller companies lack the resources and specialist expertise to generate, store, retrieve or modify data for methods improvement and other purposes. Such a weakness inevitably impacts upon productivity levels as well as quality standards.

3. Company policy objectives and strategies for quality are not formulated--The initiative taken to devise mission statements, policies and strategies for the company which form the basis or foundation for more specific objectives and action is, for the most part, not followed by small companies. Quality, therefore remains confined to the quality department, attended to by a quality expert, and quality is neither a matter for senior management nor 'everybody', contrary to recommendations for good practice.

4. Lack of written job responsibilities for personnel--Within the quality function in particular, and the firm in general, there are overlaps between activities which create what Juran and Gyrna have referred to as "a potential jurisdictional question". The clarification of job specifics is another area where duplication, and hence inefficiency - if not disputes - arise.

5. Inadequate planning of resources--The ratio of people, equipment, materials and space takes careful consideration for optimal productivity in manufacturing.

6. Absence of certain key activities: supplier assessment, personnel training, internal audits--In addition, small firms are disinclined to undertake in-house training, preferring instead to hire workers who have learnt their skills in larger factories, in advance of joining the firm, and this results in inconsistencies of working standards. The trust factor comes into play once more where auditing is concerned, a first reaction being 'what is wrong', and a second being to question validity where no 'criticism' is rightfully levelled.

In addition, the following problems are faced when companies assess the feasibility of ISO 9000/BS 5750 implementation:

1. Communication from the top down is often considered difficult to achieve. The wrong culture can kill the introduction of a quality system.

2. Difficulty in selecting the right part of the standard because guidelines are vague, sometimes even non-existent.

3. Lack of essential quality assurance activities and perceptions. To meet the standard, new procedures have to be developed, established and implemented.

4. Lack of experience of the activities outlined in the standard and in procedures and documentation makes writing the requisite procedures difficult and time consuming. This will lead to complaints from management.

5. Difficulties which arise from employees' failure to follow the written procedure.

6. Difficulty faced in understanding the requirements of the standard, its parts and clauses, and how to adapt it to a particular situation.

7. Paperwork is increased and the system has a tendency to measure the efficiency of paper work procedure rather than improvement of activities.

Other papers specific only to the textile and clothing industry, are concerned with the general quality concepts [10-12], the examination of the practice of quality management in the textile industry [13] or the preparation for application of the BS 5750 system in the clothing industry [14-16]. Little analysis has been done regarding problems faced during the actual implementation of the system in the clothing industry.

The present paper therefore aims first to look into the present situation of policy for quality of the clothing industry in Hong Kong and second, based on a case study, to identify areas of interest and potential problems which are unique to the clothing industry when implementing a quality assurance system such as the ISO 9000.

2. SURVEY TO DETERMINE POLICY FOR QUALITY IN HONG KONG CLOTHING FIRMS

In order to gauge the current situation vis a vis quality in Hong Kong, a questionnaire survey was carried out, and the quantitative data derived was analysed in the light of comments from experts interviewed.

In 1989, clothing firms in Hong Kong were similarly surveyed by means of a mailed questionnaire; at that time, data was compiled on the basis of 216 responses, and a similar mailing list (edited to remove addresses of companies which had ceased to trade) was used for comparison with the questionnaire survey conducted in 1992.

From the 1989 mailing list of 1,000 manufacturers of both knit and woven goods as well as trading establishments having garments as their main product, approximately 900 addresses were derived after editing. A questionnaire, similar to that of 1989 (but with an additional question addressing the matter of ISO 9000) was sent out, and 133 responses were analysed using the Statistical Package for Social Sciences (SPSS). A profile of the respondents is shown in Tables 1.1 to 1.4.

TABLE 1.1--Number of Persons Employed in Company

	1989	1992
under 50	14%	31%
50 - 100	40%	28%
100 - 500	40%	34%
500 - 1,000	4%	3%
1,000 - 2,000	1%	1%
2,000 and over	1%	3%
Total Respondents	215	133

TABLE 1.2--Number of Pieces Produced per Month

	1989	1992
under 1,000	11%	7%
10,000 - 50,000	46%	48%
50,000 - 100,000	25%	22%
100,000 - 200,000	8%	12%
200,000 and over	10%	11%
Total Respondents	194	131

TABLE 1.3--Number of Years in Business

	1989	1992
under 5	14%	8%
5	23%	22%
5 - 10	33%	22%
10 - 15	16%	15%
20 and over	14%	33%
Total Respondents	216	133

2.1 Existence of a Quality Control (QC) Department (Tables 2 & 3)

69% of respondents had autonomous quality control departments, a figure slightly lower than was found in the earlier survey. The majority of respondents (48%) employed less than 5 people for quality procedures, yet a minority (5%) employed more than 50 persons in their companies. Those respondents having no quality control condition gave, for the majority of cases (81%) the reason that quality was the concern of every department in the firm (table not shown).

TABLE 2--Quality Control QC Department

	1989	1992
Companies having QC Dept.	72%	69%
Companies with no QC Dept.	28%	31%
Total Respondents	216	132

TABLE 3--Number of Persons in QC Department

	1989	1992
under 5	44%	48%
5-10	32%	26%
10 - 20	14%	13%
20 - 50	5%	8%
over 50	5%	5%
Total Respondents	156	92

2.2. Material Control (Tables 4 & 5)

The control of incoming fabric was common amongst companies surveyed, 88% having responded that they carried out inspection of some form - a figure comparable with that of the previous survey. Of those who carried out inspection, 60% conducted this by visual means, and the remaining 40% used "scientific" methods, i.e., physical and chemical tests. On the other hand, only 32% of firms surveyed had in-house testing facilities, 68% using the services of local, commercial testing houses such as SGS and Labtest for fabric and garment evaluation.

Of the 133 respondents, 28% used American Standards, 25% European Standards, 10% Japanese Standards and 20% International Standards for testing purposes (more than one option available). By far the majority relied on standards specified by the buyer, this figure being 53%. As was the case for the 1989 survey, many respondents indicated use of two or more methods of testing, which might indicate a need for the establishment of in-house standards to avoid possible confusion. Since the earlier survey, a few of the more proactive, large-scale manufacturing enterprises have been working towards such a goal.

TABLE 4--Control of Quality of Incoming Fabric

	1989	1992
Visual inspection and simple measurements only	48%	53%
Visual inspection as well as physical and chemical tests	38%	35%
Company does not control quality of incoming fabric	14%	12%
Total Respondents	200	132

TABLE 5--Use of Standards for Garment and Fabric Testing

	1989	1992
American Standards	42%	28%
European Standards	35%	25%
International Standards	15%	20%
Japanese Standards	8%	10%
Methods prescribed by the buyer	42%	53%
Total Respondents	216	133

(More than one choice permitted)

2.3 Defective Garments (Table 6)

Respondents were asked to quantify levels of defects and, though less than 4% was typical, a number experienced much higher levels. Comments from respondents as to why their fault rates were 'unacceptably' high indicated that printed and other, specially finished piece goods were the cause. In the 1989 survey, it was noted that garments finished by garment dye/garment wash processes - or emerised silk piece goods - were the reason for higher fault percentages.

TABLE 6--Percentage of Defective Garments

	1989	1992
2% or less	35%	41%
2% - 4%	46%	49%
4% - 10%	15%	10%
10 % or more	4%	-
Total Respondents	213	133

2.4 Quality Emphasis (Table 7)

Asked to define how quality emphasis was promoted by their companies, respondents to the 1992, as for the 1989 survey indicated that they placed emphasis on quality considerations in every department (66% in 1992), yet they seemed disinclined to release staff for seminars on quality aspects of clothing production (14%), or to organise in-house quality meetings (41%). These latter figures, nevertheless showed some improvement over the situation in 1989. Most disappointing was the response to the option "we carefully document all procedures of relevance to quality": only 24% responded positively. However, it is

quite encouraging to note that there is also no direct correlation between the size of organisations and their emphasis on quality. Some small organisations in fact are putting much more effort in the improvement of quality than their larger counterparts.

TABLE 7--Quality Emphasis

	1989	1992
Company holds regular in-house quality meetings for staff	41%	41%
Company emphasises quality considerations in every department	63%	66%
Sends staff to seminars on quality	9%	14%
Regularly reviews quality procedures	31%	48%
Carefully document all procedures	N/A	24%
Total Respondents	204	133

(more than one choice permitted)

2.5 Specifications for Communication of Standards Required in Manufacturing (Table 8)

To gauge one aspect of "quality of communications", firms surveyed were asked about illustrated, as opposed to standard "form" or "note" specifications used in the company. 77% responded positively, a slight decrease compared to results from 1989, when the comparable figure was 83%.

2.6. Use of Attachments (Table 8)

Since work aids such as edge margins, folders and hemmers can assist sewing operatives to maintain consistent quality of workmanship and measurements, respondents from the manufacturing sector were asked to state whether they made or bought attachments. 71% responded positively, an improvement over 1989, when the figure was 53%.

2.7 Maintenance Policy (Table 8)

Regular maintenance of equipment - cleaning, sharpening, aligning, and replacing worn parts - is similarly considered essential to the quality system of a firm. 80% of respondents were "rigorously" maintaining their equipment with a view to fault prevention, this figure being similar to that for 1989.

TABLE 8--Specifications, Work Aids and Maintenance Policy

	1989	1992
Simple, well-illustrated factory standards	83%	77%
Application of work aids	53%	71%
Rigorously observed maintenance policy	78%	80%
Total Respondents	177	130

(Percentages for respondents indicating "yes".)

2.8 Quality Audit (Table 9)

62% of respondents used a quality audit, a figure very similar to that obtained in the earlier survey. Lack of formal monitoring of quality conformance and progress in implementing systems within the remaining companies may be said to be a cause for concern. It is also observed that only a few organisations that employed less than 100 persons have initiated a quality audit system.

TABLE 9--Quality Audit

	1989	1992
Quality Audit adopted	61%	62%
No quality audit	39%	38%
Total Respondents	211	131

2.9 Improvement in Quality Standard over a 12-month Period (Table 10)

The majority of respondents noted that they had seen a marked improvement in quality of goods produced in the last year (92%); this was a significantly better result than that of 1989, when the figure was 69%.

TABLE 10--Recent Improvement of Standards

	1989	1992
Standards have improved in the last 12 months	69%	92%
No Improvement	31%	8%
Total Respondents	212	129

2.10 Intention to Register for ISO 9000

Respondents were predictably negative about ISO 9000, as is borne out by comments and statistics on registration enquiries from interviewees. Only 14 organisations intended to seek certification (no comparison is possible with the 1989 survey). It is interesting to note that only 2 out of these 14 organisations were among the larger firms (employing more than 500 persons) with the majority (6) coming from organisations of less than 50 employees. Some commonalities are observed among these organisations in that they are all quality conscious, judged by their low % of defective garments produced, common use of quality audit, strong emphasis of quality management techniques and commitment to the improvement of quality.

2.11 Intention to Improve Quality Standards in the Future

Most respondents (over 90%) indicated their intention to improve standards in the future; a number of those giving negative response added the comment that quality was always a priority, therefore the question was irrelevant. This figure compared favourably with the one obtained in 1989 - yet appears contradictory to some of the preceding data.

2.12 Conclusion

As was the case in 1989, some tightening up of procedures and within garment enterprises is apparently required if they are to actively pursue a policy of upgrading and assessing the quality of merchandise. Many of the organisations surveyed were not even aware of the ISO 9000 standard. It may be the case that expectations for 1997 (the year when Hong Kong is to revert its sovereignty back to China) still encourage short term planning - this meets buyers' immediate needs in terms of flexibility, yet precludes significant investment of resources in quality systems. A proactive, innovative approach to quality improvement is therefore less common than responsiveness to market forces on a needs basis.

3. IMPLEMENTATION OF A QUALITY ASSURANCE SYSTEM IN CLOTHING ESTABLISHMENTS

This study was based on a recent experience of the authors in the implementation of a quality assurance system in a clothing factory in China (solely owned by a Hong Kong company with a long tradition of garment manufacturing) specialising in the manufacturing of high quality jeans and casual trousers. The factory concerned - a completely new establishment - was a small one employing just over 250 persons and was managed principally by personnel from the parent company in Hong Kong. The quality system was also 'borrowed' from the parent company but was modified with the aim of satisfying all requirements of the ISO 9002 accreditation scheme. Reasons for selecting the ISO 9002 for the clothing industry have been described elsewhere [14-16].

As indicated in the results of the survey of garment firms, there was a deficiency of documented quality procedure and work instructions in the existing quality system. There were also minimal quality records, written job responsibilities for staff and a complete lack of internal audits, training and supplier assessment. In order to upgrade the quality system, numerous measures had been carried out to correct the deficiency. In the process, some observations relating to the Chinese experience were recorded:

1. In the preparation of the quality manual, standard operation procedures and work instructions, undue problems arose because of the reluctance of staff participating in the work. Their concern was mainly with the issue of confidentiality; as the cultural background of Chinese is less open, supervisors of all levels tend to be more autocratic, and therefore there was an unwillingness to contribute to the documentation of procedures - especially to the work instructions.

2. Difficulties arose in persuading the workforce of the importance of quality, as products were sold in the past irrespective of their quality.

3. Quality of staff was also a major problem. Although the workforce is plentiful in China, yet they are generally poorly educated and their understanding of modern management concepts is inadequate. This problem may not be too serious for the operatives but was detrimental to the success of the project in the case of staff of management level. This shortcoming, however, was partially overcome by personnel's willingness to learn, the rather low mobility of the workforce and the higher degree of staff loyalty to the company. In the end, all personnel in senior management, and some in the middle management levels were recruited from Hong Kong.

4. The implementation of an internal quality audit met with a surprisingly good response. This could be attributed to the general mentality of Chinese, in that they are used to being scrutinised by others without questioning the authority of those in charge.

5. Insignificant attention is paid at the outset of planning to this matter, especially where cost differentials between, for example host and offshore country of manufacture are considerable.

Based on the same study in China, other observations and problems considered to be unique for the clothing industry had been revealed and these are discussed in the following:

1. Top management is generally unsure of the benefits (in terms of quality) of implementing the ISO 9000 system, as clothing industrialists always consider that quality depends on individual workmanship, rather than the system approach.

2. The industry is still predominately controlled by demands from buyers, and as long as the requirement of any quality system is not made mandatory, organisations will resist it.

3. Clothing products are most varied and style dependent. With the recent trend of small order size and short delivery time, the continuous updating of operation procedures documentation is very difficult, if not impossible. In the present study, a modular Quality Plan approach was adopted. A series of standard operations for various components were developed and, by selecting the most appropriate operation and incorporated into a Quality Plan for each style, the problem was largely overcome.

4. As workmanship is still considered to be the most crucial factor to guarantee quality in clothing production, a good process control mechanism is to be rigorously enforced. It is unfortunate that the ISO 9000 standard does not place emphasis on the process requirement [17]; this may not be a crucial determining factor for a highly automated factory but in clothing production, which is still largely people-dependent, process control becomes much more important. It is also because of this reason that it is vitally important that all operations must follow the quality system precisely and not be swayed by personalities or human relationships.

5. Whereas the importance of vertical communication is generally recognised, it is worthwhile mentioning that horizontal communication among operatives could have a decisive effect on the final quality of the clothing product. It could be argued that the Modular Production System, where small groups of operatives are responsible for their own performance, would facilitate the production of higher quality products to a greater extent than the more popular Progressive Bundle System.

6. In the final analysis, there is no fast and rigid rule governing the determination of quality standards for clothing products. Much depends upon the subjective views of the people who are inspecting the goods. Discrepancies and often conflicts arise when test and inspection results are under consideration by different parties of interest.

7. Supplier assessment of accessory products must be carefully conducted as the importance of these products are sometimes overlooked by the clothing manufacturers. Clothing companies typically purchase from suppliers in small lots, and the number of individual suppliers may be considerable. The supplier-manufacturer relationship may work in the favour of the former, in bargaining situations, when the size of the firm is small.

CONCLUSION

In the present study, a survey carried out for the purpose of determining attitudes towards, and policies for quality provided data to demonstrate current practice in Hong Kong. Further, case-based information indicated experience in implementing a system in China for ISO 9000 accreditation. It has been apparent that, in its efforts to satisfy consumer demands, as well as meet stringent import requirements for compliance purposes and market expansion, China has taken more readily to ISO 9000 than Hong Kong itself. The Hong Kong industry, on the other hand has demonstrated ability in delivering quality goods - current issues of trans-shipment and country-of-origin may, in time determine greater acceptance of formalised systems such as that promoted to China.

In the authors' opinion, ultimately the adage "get it right first time" has much to do with fail-safe communication. Though this requires a broad interpretation, formal reporting of quality matters, detailed briefs and specifications place certain responsibilities on the procurement function, therefore personnel responsible for purchasing textile and clothing merchandise require similar implementation of systems to those in a position to manufacture or supply it. Further to communication, the time factor is of particular importance to quality of supply, and in this respect, the time-sensitivity of manufactured products from the textile and clothing sector necessitates increased adoption of electronic data interchange. Finally, cultural factors specific to the firm, as well as general in the sense of manufacturing site or base, play a considerable role in system adoption. The importance of workforce training cannot therefore be stressed too strongly.

Buyers actively in South East Asia do not, currently insist upon ISO certification before placing orders for clothing merchandise, and there is undoubtedly a trade off between price and expertise for both supplier and purchaser. The factory which has achieved ISO/BS 'grading' and has all aspects of a quality system in place will undoubtedly command a price advantage and attract 'better quality' customers, such factories, although in very small numbers, exist in Thailand, Malaysia and other countries in the region, yet an interim solution to certification is required until the necessary critical mass is created, and purchasing from ISO 9000 registered companies becomes the norm. Further studies are therefore warranted to investigate the attitude of buyers towards the requirement of and reasons for factories in other countries to certify for such a standard. Findings from such studies would be most useful as a future directive for the clothing industry of Hong Kong.

REFERENCES

[1] Gould, D. T., "Implementing a Quality System in a UK Manufacturing Company," International Journal of Quality and Reliability Management, Vol. 8, No. 3, 1991 pp. 38-43

[2] Whittington, D., "Some Attitudes to BS 5750: A Study," International Journal of Quality and Reliability Management, Vol. 6, No. 3, 1989, pp. 54-58

[3] Jennings, G. M., "ISO 9001/9002 - Use, Misuse and Abuse," Quality Forum, Vol. 18, No. 1, March 1992, pp. 33-35

[4] Hersan, C. H., "A Critical Analysis of ISO 9001," Quality Forum, Vol. 16, No. 2, June 1990, pp. 61-66

[5] White, K., "Quality in the Market Place," Knitting International, Vol. 95, No. 1135, 1988, pp. 76-77

[6] Duke-Williams, D., "Management Requirements, Communications and Training," Knitting International, 4-part series, Vol. 95, 1988, No. 1132, pp. 96-97; No. 1134, pp. 106-108; No. 1136, pp. 74-75; and No. 1138, pp. 144-145

[7] Long, A. A., Dale, B. G. and Younger, A., "A Study of BS 5750 Aspirations in Small Companies," Quality and Reliability Engineering International, Vol. 7, 1991, pp. 27-33

[8] Barton, R. E., "From Inspection through Assurance to Profit," Journal for Asia on Textile and Apparel, December 1990, pp. 33-44

[9] Rayner, P. and Porter, L. J., "BS 5750/ISO 9000 - The Experience of Small and Medium-sized Firms," International Journal of Quality and Reliability Management, Vol. 8, No. 6, 1991, pp. 16-28

[10] Abed, M. H. and Dale, B. G., "An Attempt to Identify Quality-related Costs in Textile Manufacturing," Quality Assurance, Vol. 13, No. 2, June 1987, pp. 41-45

[11] Taylor, G. and Yeung, K. W., "Quality in Hong Kong's Clothing Industry," Textile Asia, Sept. 1989, pp. 206-219

[12] Ford, J. E., "Textiles and Clothing in the Single Market," Quality Forum, Vol. 16, No. 3, Sept. 1990, pp. 152-154

[13] Allen, N. and Oakland, J. S., "Quality Assurance in the Textile Industry," International Journal of Quality and Reliability Management, 3-part series, Vol. 5, No. 5, 1988, pp. 25-37; Vol. 8, No. 1, 1991, pp. 22-30; and Vol. 8, No. 2, 1991, pp. 9-20

[14] Chuter, A. J., "Quality Assurance: Part I - the Competitive Edge," Apparel International, May 1989, pp. 31,33

[15] Chuter, A. J., "Quality Assurance: Part II - Preparing for Registration," Apparel International, June 1989, pp. 35,36

[16] Chuter, A. J., "Quality Assurance: Part III - Preparing the Submission," Apparel International, August 1989, pp. 28,30

[17] Gaskin, M. R., "The Future of Quality Management Systems," Quality Forum, Vol. 17, No. 2, June 1991, pp. 83-86

Alex T.C. Lau [1]

APPLICATION OF STATISTICAL QUALITY ASSURANCE TECHNIQUES TO CONTINUOUS PROCESS ANALYZER SYSTEMS FOR ON-LINE PRODUCT QUALITY CERTIFICATION

REFERENCE: Lau, A. T. C., "Application of Statistical Quality Assurance Techniques to Continuous Process Analyzer Systems for On-Line Product Quality Certification," Quality and Statistics: Total Quality Management, ASTM STP 1209, Milton J. Kowalewski, Jr., American Society for Testing and Materials, Philadelphia, 1994.

ABSTRACT: In the petroleum refining industry, use of Continuous Process Analyzer Systems to perform product quality measurement functions has been common practice for over at least two decades. Designed to operate unattended, these systems provide analytical information ranging from physical properties such as viscosity, density to total and trace chemical composition, on samples continuously extracted from the main process streams. While this information has primarily been used for process monitoring, control, and real time optimization, its usage for on-line product quality certification has been limited. In spite of the capability of these systems to equal or outperform their labour-intensive laboratory counterparts both in precision and accuracy, most organizations are reluctant to use them for product certification purposes. This is primarily due to the lack of a structured, objective, data-driven approach taken by both internal and external users (customers) of these systems to measure, compare, and demonstrate their capabilities. With the increasing cost of crude oil, and more stringent product specifications demanded by customers and environmental legislation, the business incentive capturable by using a more capable measurement system for product quality certification has become increasingly significant.

This paper begins with an overview discussion of the measurement process, its uses, and performance metrics. The main body deals with the application of basic statistical quality assurance concepts, principles and practices to identify the appropriate (statistical) model, quantify, compare, and continuously demonstrate in-control status of measurement systems. Control charts referenced include X-bar, I, R, MR2, and EWMA. Statistical tools referenced include t, F tests of significance, normal probability plot, ANOVA. This paper concludes with examples of how these tools and practices are successfully applied to the on-line certification of gasoline vapour pressure quality using a continuous process analyzer system.

Detailed presentation on mathematical formulae and computational mechanics for statistical tools and control charts are beyond the intended scope of this paper. Readers interested in these details are referred to statistics and statistical quality assurance textbooks. Several are listed in the reference section.

[1]Fuels Blending and Quality Assurance Specialist, Imperial Oil, Products Division; 111 St. Clair Ave. W., Toronto, Ontario, Canada M5W 1K3

KEYWORDS: measurement process, continuous process analyzer system, specification, product quality certification, statistical quality assurance, precision, bias, true value, point estimate, interval estimate, statistical control, control charts

---

MEASUREMENT PROCESS AND ITS USES IN THE REFINING INDUSTRY

The measurement process is a sub-process integral to all manufacturing processes. Like any process, it has inputs and outputs. The inputs are usually samples taken at various stages of the manufacturing process, the outputs are usually numerical values. Use of these numerical values, in general, fall under three categories:

> manufacture process control: to test the hypothesis that the manufacturing process performance meets predetermined criteria (e.g.: in "statistical control").

> product quality certification: to test the hypothesis that the batch of product from which the sample tested is taken meets predetermined criteria (e.g.: specifications).

> measurement process self-monitoring: to test the hypothesis that the measurement process itself is "in statistical control".

In the petroleum refining industry, two types of measurement processes are commonly encountered. One is the traditional laboratory-based process where a small sample is extracted from the main process streams and analyzed off-line. The other is a field-deployed, fully automated instrumentation system designed to provide on-line analytical information on samples continuously extracted from the main process streams. This is commonly known as a continuous process analyzer system.

Manufacturing process control is usually implemented continuously and automatically, necessitated by the frequent changes in raw material qualities (different crude oils) and production strategies for optimum yields of different types of products. Control action are calculated from algorithms developed from classical process control theory. Control adjustment frequency and magnitude are tailored specifically to the processes gain and dynamics. Continuous process analyzer systems are an integral part of this control strategy as they provide the necessary information for control action calculations and measure of control effectiveness. The SPC (Statistical Process Control) philosophy of "intervene only when an out-of-control signal is detected", commonly deployed in the discrete components manufacturing open-loop processes, is not practiced due to underlying process model differences.

Product quality certification, on the other hand, are typically carried out using the laboratory-based (off-line) measurement process, where small quantities of samples are taken from finished-product tanks and analyzed to ascertain specification conformance status. It is quite common to find this practice followed even when similar analytical information is available from continuous process analyzer system(s) already installed on the process streams discharging into these product tanks. In almost all cases, these continuous process analyzer systems are technically equivalent or better than the laboratory-based measurement processes. This duplication of product quality measurement efforts can be eliminated (or minimized) by first collecting performance data for both processes through a properly designed and conducted evaluation study. After appropriate statistical analyses of the data, the better measurement process is then selected as the preferred process for product quality certification.

Measurement process self-monitoring schemes are usually employed to demonstrate its "statistical control" status to users. This is achieved by regular control charting of results obtained on aliquots prepared from selected batches of production material over time. The collection of activites to achieve the aforementioned objective are referred to as a measure process quality assurance program.

In the sections to follow, the main topics discussed are:

> Use of measurement process outputs and statistical inference techniques to ascertain product quality conformance to specifications

> Application of statistical techniques to evaluate and select the better measurement process between the two described above for product quality certification;

> Design and implementation of measurement process quality assurance program to continuously demonstrate "in control" status over time

References to statistical tools and control charts are made with intent to show the applications of these tools. Detailed presentation on mathematical formulae and computational mechanics are beyond the intended scope of this paper. Readers interested in these details are referred to textbooks on statistics and statistical quality assurance techniques. Several are listed in the reference section.

USE OF MEASUREMENT PROCESS OUTPUTS AND STATISTICAL INFERENCES TO ASCERTAIN PRODUCT QUALITY CONFORMANCE TO SPECIFICATIONS

Product quality specifications are usually established between supplier and customer to quantitatively define the desired value and limits of acceptance for the specific product qualities. For the refining industry, integral to quality specifications are associated measurement methodologies for determination of specification conformance. ASTM test methods are usually referenced for this purpose. These test methods, in reality, are description of measurement processes. Due to the natural variation of many factors inherent to these processes (test methods), repeat application of the same process to the same or different samples taken from the same batch of homogenous product will usually yield different numbers. Unlike the well established acceptance sampling plans practiced for decades in the discrete component manufacturing industry, most contractual documents in the refining industry fall short of explicitly defining the procedure for how to apply the measurement process to arrive at a numerical value for determination of specification conformance. In the absence of contractual agreement to the contrary, an objective and scientific approach is to assume that the specification requirements apply to the *true* value for the quality attribute.

For most of these ASTM test methods, materials with industry-accepted reference values, AND, are compositionally representative of the products these methods are intended to measure, do not exist. Under these circumstances, the true value is only defined in terms of the test method, and hence represents a theoretical concept with the following definition:

> "For practical purposes, the value towards which the average of single results obtained by *n* laboratories tends, as *n* tends towards infinity; consequently, such a true value is associated with the particular method of test." [1]

In a production environment, the quest for true quality value for a batch of product is very seldom practiced since it is neither

practical nor economically feasible. Typically, the measurement process is applied to one or two samples taken from a batch of product. These results are then interpreted to be "exact" in nature by most producers. Products with results close to or at specification limits will often be released without additional testing. While this practice minimizes producers' risk at the expense of increasing consumers' risk, it is not statistically correct if the true value approach for specification interpretation is used. Due to the inherent natural variation of the measurement process, these results should be treated only as individual *point-estimates* of the true value.

An objective and more scientific approach is to use statistical inferential technique, in combination with the results, to estimate the *probability* of product conformance to specification. While most suppliers would like to claim, and customers would like to specify this probability to be 100%, this is not attainable by classical statistical definition. These point-estimates should be used in conjunction with knowledge of the *precision* (defined later) of the measurement process, to construct an interval which may contain the true value with a quantifiable probability. This is referred to as an *interval estimate* of the true value. The width of this interval is generically referred to as the error of estimate.

Using the above approach, the supplier and customer can establish numerical values known as *acceptance limits* that the interval estimate *boundaries* should be compared to, for the purpose of product acceptance. Depending on the degree of assurance (criticality) required and associated economics, acceptance limits may be numerically equal to, better than, or worse than the specification limits [1].

Between the three factors listed below, for a given measurement process precision, specifying the values for any two factors automatically defines the value of the remaining one:

1) true value interval-estimate width
2) desired probability of capturing the true value
3) number of times the measurement process is to be replicated

Since it is always desirable to be able to capture the true value with a high probability, using a narrow interval and a small number of measurement process replication, given the choice of two measurement processes, the one with better precision is always preferred.

EVALUATION STUDY ON A LAB-BASED MEASUREMENT PROCESS AND ITS CORRESPONDING CONTINUOUS PROCESS ANALYZER SYSTEM

Use of Precision and Bias as Performance Metrics

Prior to evaluation of any process, performance metrics must be selected. For measurement processes, most commonly used metrics are precision and bias.

Precision--This is a generic term used to describe the variation aspect of the measurement process due to common causes. This variation comprises the cumulative effects of random variations from sources that are inherent to the process itself. Precision is usually measured under a predefined set of operating conditions. These conditions must be quoted with any precision figure to permit its proper interpretation.

Bias--This is a quantitative measure of the degree of agreement between actual versus expected outcome of the measurement process. For example, if the quality of a batch of material to be measured is boiling point, and the material is water, then, the *expected* value is 100 degree

C. at sea level. The difference between the latter and the theoretical *true* value attainable by infinite replication of the measurement process is a measure of its true *bias*. Bias can be viewed as a measure of the distance between the natural "centre" of the measurement process and an expected location of this centre, established through scientific research, theory, or by industry consensus.

Objectives of Evaluation Study

The main objectives of this study and associated rationale are listed below (listing order is not prioritized):

> Identify the underlying mathematical model that best describe each process behaviour pattern under study, and estimate its parameters.

Rationale: Proper model identification permits statistical inferential statements to be made on the "in control" status of the process based upon limited observations.

> Quantify the true bias of each measurement system: i.e.: ascertain if each process has a natural centre that is statistically different from some expected value.

Rationale: Any bias discovered, provided it is stable, predictable, and explainable scientifically, can and must be used in conjunction with the appropriate precision estimates to quantify the probability of capturing the true value by the interval estimates.

> Quantify relative bias between the two measurement systems: i.e.: ascertain if each process has a natural centre that is statistically different from each other.

Rationale: The relative bias can be used to relate observations from one measurement process to the other.

> Quantify total precision (variation) for each process, under typical operating conditions.

Rationale: Precision is the primary performance metric used to select the better process.

> Establish if the bias and precision are dependent on the level of quality being measured; if so, quantify this relationship.

Rationale: The relationship of bias and precision versus quality level has to be investigated and understood prior to use of any measurement process.

What Factors to Control and What to Vary

Similar to other processes, the factors that can affect the measurement process outputs can also be grouped into the five common categories (an approach used by a process improvement tool known as the "fishbone" diagram):

people, machine, material, method, environment.

In the design of this study, it is important to identify what factors should be "controlled", and what factors should be allowed to vary in order to achieve the study objectives.

Controlled factors--Typically, this type of study should be conducted on processes that measure similar quality attribute (e.g.: vapour pressure), at several different nominal values representative of

typical product quality levels under normal operating conditions. If multiple instruments are available for the lab process, the study should be limited to one instrument. Findings should be verified with a follow up study on a different instrument. Under this scenario, the "controlled" factors can be thought of as *machine*, *material*, and *method*.

Varied factors--For the lab-based process, both *people* and *environment* factors should be deliberately allowed to vary since these factors are inherently part of the process. It is also important to include the sample extraction procedure (from the manufacturing process streams or product tankage), performed by different operators on different days), as part of the study since the latter constitutes an integral part of the total measurement process.

For most continuous process analyzer systems, totally enclosed and automated sample extraction sub-systems are employed to deliver material representative of the manufacturing process streams continuously to the analytical instrumentation. As well, totally enclosed, fully-automated reference sample storage and injection facilities constitutes part of the standard system design feature. For these systems, only the *environment* factor can be meaningfully varied (different time and day) as the effect from the *people* factor is not expected to have any discernable effect on the sampling and injection tasks due to automation.

For analyzer systems that are not equipped with reference material storage and injection facilities (due to technical or practical reasons), selection of material to be studied will be limited to that being produced by the manufacturing process. Since there is no assurance that the exact material will be produced throughout the study, direct control of material factor is not possible. Hence, system precision without the material factor cannot be directly estimated. However, this can be mathematically inferred from the results of the study, if the corresponding lab-based process precision can be independently assessed.

For the remainder of this paper, analyzer system equipped with reference material storage and injection facilities is assumed.

How and How Much Data to Collect

Collection of data should commence only after ensuring all equipment is in good working condition and the instrumentation calibrations have been validated. During the data collection period, non-routine adjustments (including calibration) or equipment component changes should be avoided where possible, or duly noted if unavoidable.

Paired observations--Results should be obtained from both processes, under regular operation conditions, for the *same* sample. This pairing is required to minimize the material effect, and can be achieved as follows:

> Inject the reference material chosen for the study into the analyzer system.

> Within a short time after the injection, obtain a sample of the same reference material at a point located as close to the analytical portion of the analyzer system as practical, using the appropriate sampling procedure required for the lab-based measurement process.

> Obtain the analysis from the analyzer system, after it is ascertained that the system has come to equilibrium with the reference material.

> Obtain a result from the lab-based measurement process for the sample collected above within as short a time as practical.

The two results above (analyzer and lab) constitute one pair of observations for this study. A reasonable number of pairs of observations, at each nominal quality value studied, is between 15 to 25. This should provide meaningful estimates to the sought after parameters.

Data should be collected in an equally distributed fashion amongst as many operators as practical, and, spread out evenly over as long a period as practical. Replications within a short time in order to expedite data collection is not appropriate as this would not provide adequate variation of the *environment* factor (time is considered part of the environment factor), and consequently would lead to an underestimation of the total process variation over time. A general rule of thumb is to collect no more than 2 pairs per day, by different operators, separated by a minimum of eight hours.

DATA TREATMENT

Phase 1: Preliminary Assessment

Screen for suspicious results--Results from each process should first be visually screened for the existence of unusual values, such as those that could have been caused by transcription errors. The dot plot [2] is an excellent visual aid that can be used for this purpose. Results flagged as suspicious should be investigated. Discard of any data at this stage must be supported by evidence gathered from the investigation, and/or sound engineering judgement.

Screen for unusual patterns--The next step is to plot, on separate charts, results from each process, in chronological order. In SPC terminology, these are known as "run charts". The charts should be examined for non-random patterns such as continuous trending in either direction, unusual clustering, and cycles. Several of the non-random patterns described in control chart literature [3,4] can be used as guides at this step. Detection of any non-random pattern should trigger investigation for causes.

Assess process stability--If no obvious non-random patterns are detected from the run charts, Individual (I) and Moving Range of 2 (MR-2) control charts (described later) should be constructed. All control chart rules should be applied at this point to determine if both processes under study are under the influence of common causes variation only ("in statistical control"). Out-of-control points should be investigated in detail. For those associated with special causes, their exclusion from further data analysis is warranted only if the special causes can be deemed not to be part of the normal process.

If, after removal of appropriate results, there is a minimum of 15 observations remaining, continue with data analysis. Otherwise the entire study has to be repeated after the causes for unusable results are identified and suitably addressed.

Test normality assumption--From industry experience, the Normal model has been found to be applicable to most measurement process under "statistical control". It is also the model most control chart limits are based on. Normal probability plots should be constructed using results from each process to test for validity of this assumption. For small number of observations (15 to 25), this tool tends to provide better information regarding the 'goodness-of-fit' of the Normal model than the more commonly known histogram, which generally requires 40 or

more data points to be effective. As well, departure from normality detected by this plot may reveal abnormal conditions that are not discovered by the previous control chart screening steps.

Phase 2: Calculation of Statistics

If the Normal probability plot reveals no unusual pattern, proceed with detail calculations of statistics. The data should first be grouped into the following 3 variables: **analyzer, lab,** and **delta.** For the **delta** variable, the members are the signed difference between each pair of observations. The average (x-bar) and sample variance (s^2) statistic for each variable are then calculated.

Model parameter estimate--The Normal model can be fully described by two parameters, known in classical statistics terminology as the *mean* (μ) and *variance* (σ^2). The mean parameter can be most efficiently estimated by the sample average statistic. Although one type of estimate of the variance parameter is already available using the Range technique (from the MR(2) chart), at this point, a more efficient estimate should be obtained by calculating the s^2 (sample variance) statistic. Detail discussions on the pros and cons of these two techniques are beyond the intended scope of this paper. Interested readers can consult statistics textbooks or reference [3] for a detailed presentation on this topic.

Relative bias estimate--Results from the **delta** variable should be used for this purpose. A one-sample t-test should be performed using the sample average and variance statistics, to test if the data supports the hypothesis that there is no statistically significant bias between the two processes. In statistical terminology, this is a test of the *null* hypothesis: mean (μ) parameter of the variable **delta**=0, versus the *alternate* hypothesis of not equal to 0. The typical confidence level that this test is performed at is 95% (alpha=0.05).

If the test is significant, the sample average of **delta** variable can be used as a point-estimate of the relative bias between the 2 processes. Alternatively, an interval estimate of the relative bias can be constructed, using the sample average and the standard error (square root of the ratio sample variance/number of data points).

True bias estimate--If the material studied has an expected quality value, then, the one-sample t-test should be applied independently to the t-statistics for **analyzer** and **lab** variables to determine if a statistically significant true bias exists for each process, relative to an externally established reference point. The hypothesis under test is that the mean (μ) of each process is equal to the expected quality value, versus a not-equal hypothesis.

Precision estimate--The sample standard deviation of the **analyzer** variable represents an estimate of continuous process analyzer system precision.

The sample standard deviation of the **lab** variable represents an estimate of the total lab-based measurement process precision, *including* the sample extraction procedure.

The sample standard deviation of the **delta** variable represents precision estimate for this paired comparison process. In the long run, the expected value of this quantity, is the square root of the sum of the variances of the 2 measurement processes, using the additivity rule of variances.

Relationship between Precision and Quality Level

For studies conducted at 2 or more levels, a visual examination of the sample variance versus quality level scatter plot, for each process, is always advisable as this can provide clues as to whether an obvious relationship exists.

For 2 level studies, an F-test on the ratio of sample variances can be used to test the hypothesis that one variance is statistically different from the other. For 3 or more levels, other techniques such as Cochran's or Hawkin's test can be used to test for variance homogeniety across different levels.

Relationship between Bias and Quality Level

Similarly, visual examination of sample averages versus quality level scatter plots can provide clues as to whether an obvious relationship exists. To examine the relative bias of the 2 processes versus quality level relationship, sample averages of the **delta** variable, usually plotted versus the lab process, should be used. To examine absolute bias of each process versus quality levels, the *difference* between sample average and expected value at each level should be plotted.

For 2 level studies, a more rigorous approach is to use a two-sample t-test to test if the bias estimates at the 2 levels are statistically different. For 3 or more level studies, advanced statistical techniques such as one-way analysis of variance (ANOVA) can be used to test for statistically significant differences. Test for homogeniety of variance is usually part of the one-way ANOVA analysis.

Presentation of Bias and Precision Relationships with Level

If a relationship does exist, there are usually two approaches to present this information for every day use. The simple way is to simply state, in tabular format, the precision estimates at various levels of interest. The other is to describe this relationship mathematically, either expressed in percentage of the quality level, linear regression models, or higher order models from other curve-fitting methods. For practicality reasons, the simpler, tabular format is preferred.

SELECTION OF THE BETTER MEASUREMENT PROCESS

Pick the "Winner" based on Precision

Since the performance metric to be used for selection of the better measurement process is precision, an F-test of the ratio of sample variance statistics for the **analyzer** and **lab** variables should be performed, with the numerically larger value in the numerator. A significant outcome from this test is an indication that the data supports the hypothesis that there is a statistically significant difference between the precision (variances) of the 2 processes, with the process in the numerator being the one with the poorer precision (larger variance). In almost all cases, the precision of process analyzer systems are either significantly better, or no worse than the lab counterparts. Both scenarios support designating the analyzer system as the preferred process since, under the "no worse than" scenario, unnecessary work can be eliminated.

HOW TO DEAL WITH BIAS

Root causes for statistically significant biases, whether they are true or relative, should always be investigated for. If they are determined to be inherent to the process (e.g.: the analytical principles employed, or, operating conditions required by the instrumentation), and there is a need to compensate for the bias, the best way to deal with it is to *mathematically* compensate each outcome of the process by the appropriate amount. In most cases, the value used for compensation is the point-estimate of the bias. If a very high degree assurance of specification conformance is required [1], the worst case boundary value of the probability-specified interval estimate should be used. The mathematical approach is preferred over an instrumentation adjustment approach, as this would eliminate any variation associated with the adjustment procedure itself.

SITUATION MOST ENCOUNTERED IN THE REFINING INDUSTRY

In the refining industry, the situation encountered most frequently is one where the measurement process referenced in the specification is a lab-based process, while the manufacturing process is controlled using a process analyzer system that has a better precision **AND** a statistically significant bias relative to its lab-based counterpart. No reference material is available with composition similar to the product matrix and an industry-accepted reference value. Under this set of complex circumstances, many producers are reluctant to use the process analyzer system results *alone* to certify product quality. There is no technical reason why this cannot be done, provided a suitable mathematical compensation for the relative bias is applied to the analyzer results, AND, the relative bias is established with a high degree of precision (i.e.: large amount of paired-data at the appropriate levels).

Results and analyses from an evaluation study, conducted at a refinery, for a lab-based vapour pressure measurement process and its corresponding process analyzer system, is presented as Example 1.

EXAMPLE 1

Evaluation Study of Lab-based and Analyzer Measurement Processes for Vapour Pressure of Gasoline at nominal level of 10 units

Results from this initial study are displayed in Table 1.

Preliminary Assessment

From the dotplot of results (Fig. 1) no suspicious result was identified. Subsequent run chart and control chart assessment did not reveal any obvious patterns nor out-of-control points (Fig. 2a,b,c,d).

An investigation was triggered by the non-normality observed in the **analyzer** dataset (Fig. 3). The root cause was determined to be minor material non-homogeniety inside the reference storage cylinder caused by variation in the manufacturing process while the cylinder was being filled. Although this material variation affected both datasets, due to larger variability in the lab-based process, it was "masked" and not as prominent (Fig. 4). The reference cylinder fill and mix procedure was appropriately modified.

The above discovery led to the discard of all data from the initial study. A repeat study was conducted, with no abnormality encountered.

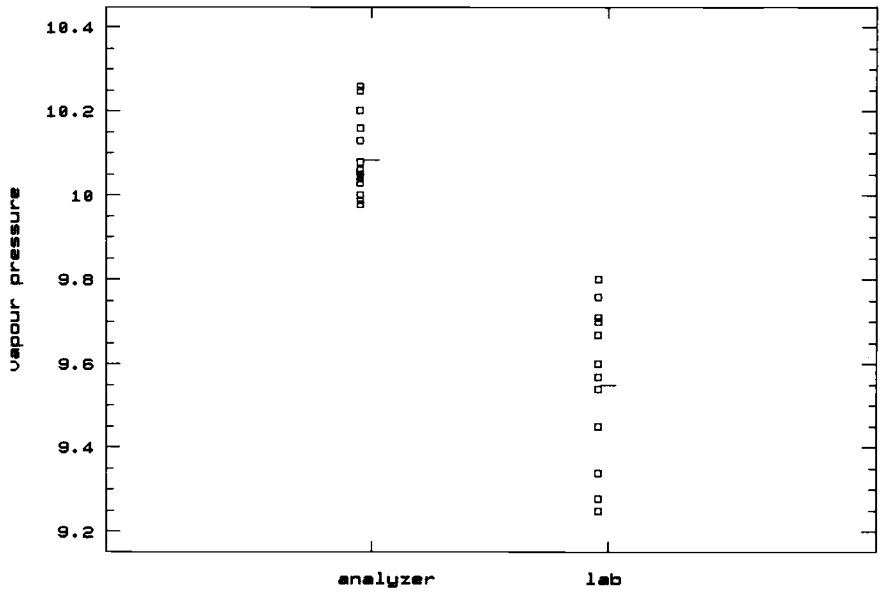


FIG. 1--Dot plot of results from initial study

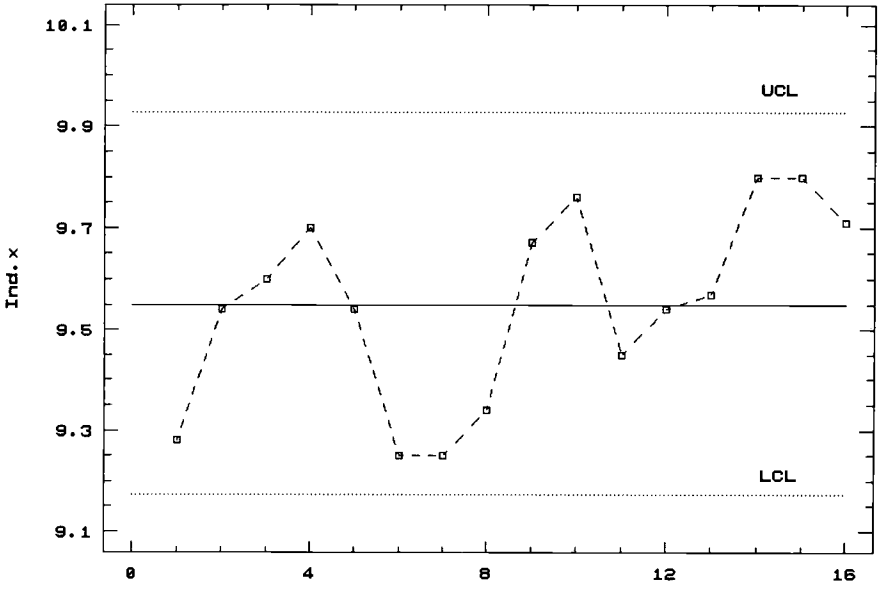


FIG. 2a--I chart for lab process

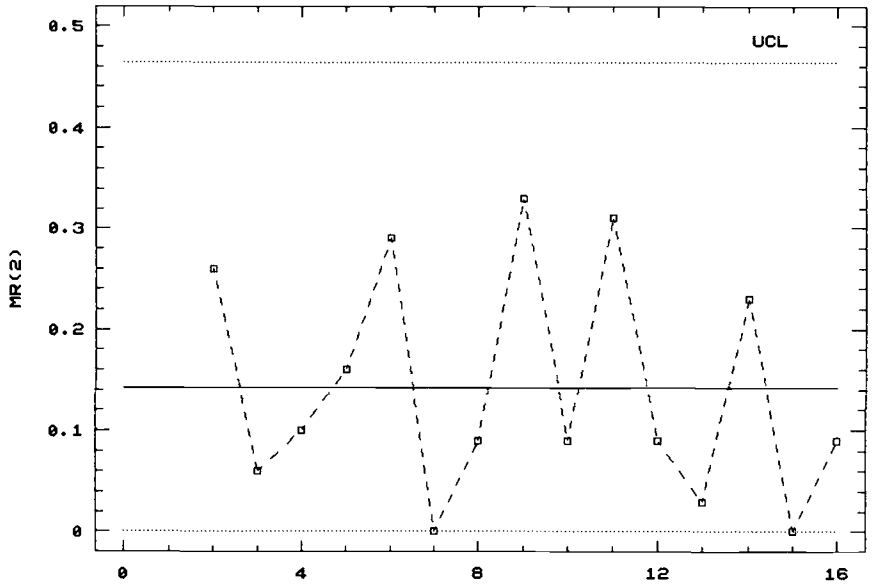


FIG. 2b--MR2 chart for lab process

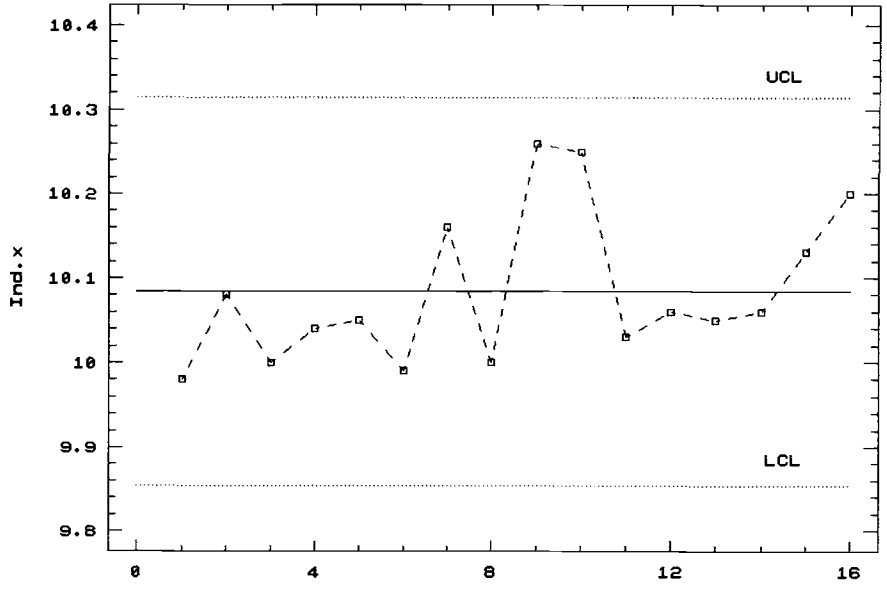


FIG. 2c--I chart for analyzer process

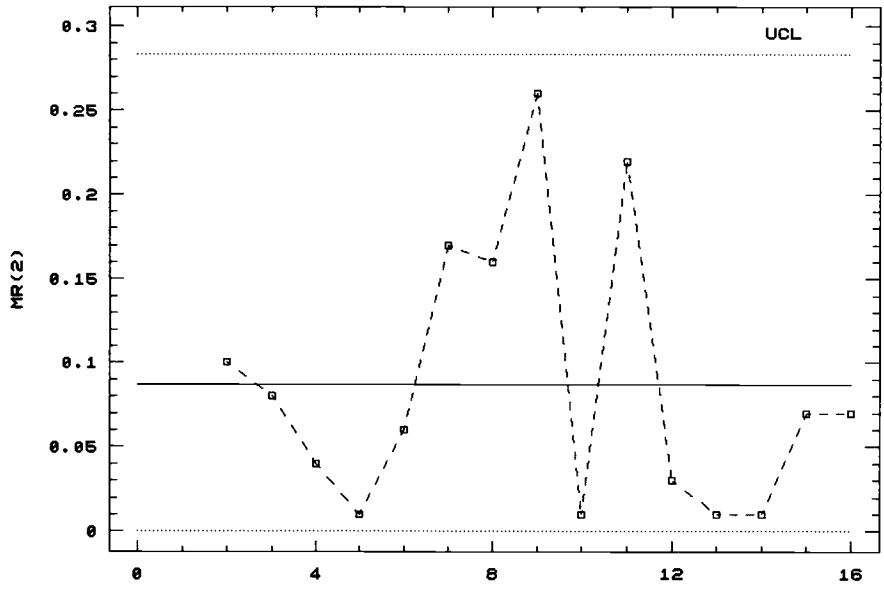


FIG. 2d--MR2 chart for analyzer process

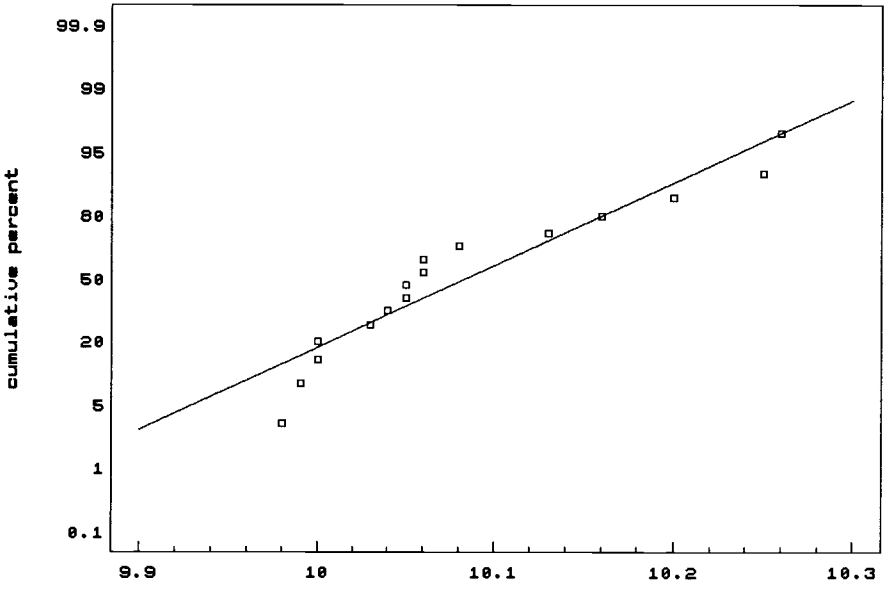


FIG. 3--Normal probability plot for analyzer process

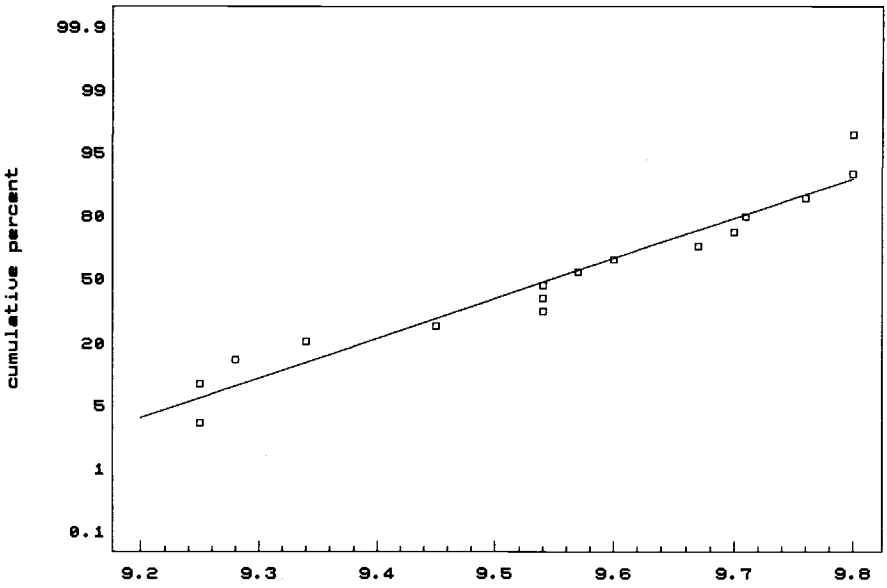


FIG. 4--Normal Probability plot for lab process

TABLE 1--Results of initial study

date	analyzer	lab	delta
may27	9.98	9.28	0.7
jun2	10.08	9.54	0.54
jun4	10	9.6	0.4
jun8	10.04	9.7	0.34
jun9	10.05	9.54	0.51
jun9	9.99	9.25	0.74
jun10	10.16	9.25	0.91
jun10	10	9.34	0.66
jun11	10.26	9.67	0.59
jun11	10.25	9.76	0.49
jun15	10.03	9.45	0.58
jun15	10.06	9.54	0.52
jun16	10.05	9.57	0.48
jun16	10.06	9.8	0.26
jun17	10.13	9.8	0.33
jun17	10.2	9.71	0.49

Calculation of Statistics (data from repeat study)

From the results of the repeat study (not shown), the following statistics were calculated:

Variable:	**analyzer**	**lab**	**delta**
Sample size	16	16	16
Average	10.1999	9.70268	0.497219
Variance	1.71342E-3	7.6771E-3	0.0110607
Standard deviation	0.0413935	0.0876191	0.10517
Standard error	0.0103484	0.0219048	0.0262925

Model Parameter Estimates

Estimates for process variances were extracted from the summary statistics as follows:

Analyzer process variance: 0.0017
Lab process variance: 0.0077

Since there is no expected value for the material, true biases are not applicable.

Relative Bias Estimate

Results from a one-sample t-test on variable **delta** are listed below:

Sample Statistics:	Number of Obs.	16
	Average	0.497219
	Variance	0.0110607
	Std. Deviation	0.10517
	Median	0.510295
Hypothesis Test for H0: Mean = 0		Computed t statistic = 18.9111
	vs Alt: NE	Sig. Level = 7.08966E-12
	at Alpha = 0.05	**so reject H0**.

Since the t-test was highly significant, it was concluded that there is a statistically significant bias between the lab-based measurement process and analyzer system, at the nominal level of 10 vapour pressure units. Point and 95% confidence interval estimates of the relative bias are: 0.497 and 0.441 to 0.553 respectively.

PICK THE WINNER

F-test on Ratio of Sample Variances (larger statistic used in numerator)

F ratio = 4.48 15 D.F. for both numerator and denominator
$F_{critical}$ = 2.40 @ alpha = .05
= 3.52 @ alpha = .01

Since the F-test is significant at both alpha levels, the alternate hypothesis that the variances are different was accepted, in spite of the limited number (16) of data points. It was concluded that the analyzer system is *more* precise. Based on the above results, the

process analyzer system is designated as the preferred process for product certification.

BIAS COMPENSATION

Given that the vapour pressure specification is a one-sided maximum specification, AND, conformance to maximum limit is a legal requirement, this quality is deemed critical [1]. Since the lab-based method is referenced in the specification, the minimum limit (0.441) of the interval estimate, rather than the point estimate, is used to compensate for the relative bias. This value (0.441) is subtracted from all results from the process analyzer system prior to comparison with a release value (lower than the maximum specification limit) to determine if product vapour pressure meets specification. This compensation value will be appropriately updated as the interval width narrows with increased number of data points.

MEASUREMENT PROCESS QUALITY ASSURANCE PROGRAM

In order to be able to use any measurement system with confidence, its in-control status needs to be continuously demonstrated. This is best achieved using *control chart* techniques to monitor the results from regular application of the measurement process to the same material. The latter is commonly referred to as *quality control* samples. The collection of quality control sample preparation procedures, frequency of sample analysis, control charts, decision rules, response procedures when process is deemed out-of-control, regular statistical analysis and management stewardship of data, constitute the essential elements of a *measurement process quality assurance program*. Both lab-based measurement processes and process analyzer systems can be monitored by programs of similar design. Program data should be periodically analyzed to evaluate process performance and validate the statistical model. Precision and bias estimates should be updated where appropriate. For practicality reasons, the sample extraction procedure for the lab-based measurement process is typically not part of the program design. Separate studies designed to monitor the latter should be regularly conducted.

Quality Control Material Selection and Treatment

Materials selected should be a stable, finished product extracted from the manufacturing process with values and composition similar to those regularly encountered by the measurement process. For lab-based processes, smaller aliquot portions are usually prepared from a larger volume of sample extracted from the manufacturing process at one time and mixed thoroughly to ensure homogeniety. These smaller portions are then stored in a controlled environment to preserve sample integrity over time. For continuous process analyzer systems, custom-designed sub-systems are usually employed to carry out the extraction, mixing, storage and delivery functions.

Frequency of Quality Control Sample Measurement

Depending on the criticality of the quality being measured by the process, the frequency of quality control sample measurements can range from once per day to twice per week. For lab-based processes, daily validations are quite common. Since most process analyzer systems are designed for unattended round-the-clock operation without shutdown, key system parameters are usually under self-diagnostics surveillance with auto-alarm capabilities. Operated in this mode, these systems can be viewed as steady-state continuous processes. With the current

technology, quality control validation frequency of 2 to 3 times per week is generally sufficient.

Control Charts

The most practical and technically applicable control charts are the Individual (I) and Moving Range of 2 (MR2) charts. The I chart should be enhanced with an EWMA (Exponentially Weighted Moving Average) overlay, at a lambda factor of 0.2. The values plotted on the I chart are results from applying the measurement process to the quality control sample *as if it was a regular production sample*. Special treatment of any kind to the quality control sample would seriously undermine the integrity of the control chart data in terms of providing a valid estimate of the long term measurement process precision.

The I chart, with the EWMA overlay, tracks total process location (bias) stability. This combination design is for the detection of both large and small shifts of the bias value simultaneously. The MR2 chart tracks the total process variation (precision) over multiple operators and time.

To "kick-start" these control charts, preliminary control limits from the evaluation study can be used. These limits should be periodically updated as new data become available. The update period should be based on a minimum no. of new data points. Update calculations should be preceded by an F-test on the ratio of sample variances for the new and previous data. If the test is not significant, a new estimate of the total process variance should be calculated based on a weighted average of the sample variances from both periods using the number of data points as weighting factor (this technique is known as statistically pooling of variances). A significant F-test should trigger an investigation for root causes.

The choice of lambda at 0.2, in addition to being the optimal value, also conveniently places the control limits (3-sigma) for the EWMA trend at the 1-sigma values for I chart.

It is not appropriate to perform replicate analyses in rapid succession and then apply the classical X-bar and Range (R) charts (commonly used in the discrete component manufacturing industry) to these results. These charts are designed for a process model that assumes the total variation can be estimated solely by the variation (range) within each subgroup. For measurement processes typically encountered in the refining industry, this is an incorrect model as the environment (time) and operator factors, both considered inherent to the process, can contribute significantly towards the total process variation. These contributions would not be included using the X-bar, R chart approach.

New Quality Control Material Certification

Due to inherent variation of the manufacturing process and different operation targets, minor or major differences may exist between batches of materials extracted at different times for quality control use. Since control limit calculations for the I chart require a centre value usually established by the process itself, a special procedure is required to ensure the centre line for a new batch of QC material is established with a process that is in statistical control. This procedure is generically referred to as the 'QC certification procedure', and can be designed as follows:

> Collect and prepare a new batch of QC material as the current supply is close to depletion.

> Start a new set of I, MR2 charts for the new material. Commence collection of data for the new material each time a validation is carried out using the current material. The new data is deemed valid if the process in-control status is established by the current QC material.

> Use the control limits for the current MR2 chart, at the appropriate nominal quality level, for the new MR2 chart.

> After a minimum of 5 in-control data points is collected on the new material, update the control limits of the new MR2 chart as appropriate, per the F-test and pooling described above.

> Calculate the centre line for the new I chart.

> Calculate the control limits for the new I chart using the new centre and new MR-BAR value.

With an effective QC material inventory management program, the QC certification and calculation/transfer of new control chart limits should be seamless.

Special Procedure for Relative Bias Monitoring

If a relative bias exists between the lab-based process and the analyzer system, AND, compensation of this bias is required for product certification purposes, regular control chart monitoring of this bias need to be an integral part of the measurement process quality assurance program. This can be most efficiently achieved in conjunction with the certification of a new QC material as follows:

> Construct I, EWMA [5] overlay, and MR control charts for the **delta** variable, using the data from the evaluation study. There should be 1 set of charts per nominal level of interest. The initial control limits are established using results from the evaluation study. Thereafter these limits are revised periodically based on accumulated data.

> Perform at least 2 "paired analysis" similar to those done in the evaluation study, for the new QC material. Plot the new delta readings.

> If the new MR values are in control, AND, all new delta points including the EWMA trend line fall within the 1 sigma zones of I chart, there is no evidence suggesting out-of-control status of both measurement systems.

> If one or more delta points fall within the 2 and 3 sigma zones, and/or the trend line exceeds the 1 sigma limit, and no apparent special cause is detected, collect several more pairs of data, AND, declare both systems are now under investigation.

> If any delta or MR points fall outside the respective control limits (3 sigma), declare both systems out of control and initiate troubleshoot procedure.

Expected Outcome when System is in "statistical-control"

If no assignable cause is present, then:

> No statistically significant change is expected for the bias values established in the evaluation study.

> Fluctuation is expected between I-chart centres for different QC batches. It represents material difference only.

> No statistically significant difference is expected for the analyzer MR chart control limits for different QC's at similar nominal levels.

Fig. 5 a, b illustrates how control charts were set up to monitor the relative bias, using results from the repeat evaluation study.

ESTABLISHING TRACEABILITY TO INDUSTRY-ACCEPTED REFERENCE MATERIAL

For measurement processes used in final product quality certification applications, traceability of performance to standard materials with similar composition to the product and industry-accepted values must be demonstrated. This can be achieved by periodic evaluation of such materials. For many measurement processes (test methods) referenced by the refining industry, such materials are not available. Hence the only viable means to establish this traceability is through regular participation in interlaboratory exchange programs. The I, EWMA, and MR2 control charts can be used to monitor an individual participant's performance in these exchanges. The variable charted should be the signed difference between the participant's result and a *recalculated* exchange average without the participant's result. From a practical point of view, provided the total number of participants is greater than 16, this recalculation is not necessary.

CONCLUSION

The use of statistical techniques for measurement process evaluation and validation provides a cost effective method of ensuring quality measurements. Use of continuous process analyzer systems for on-line product certification, coupled with an effective measurement process quality assurance program, can consistently contribute towards providing customers with a quality product at a lower cost. In order to achieve the end objective of using a proven measurement system with the least variability, there is NO shortcut to the associated data collection, accumulation, and data analysis tasks required.

REFERENCES

[1] ASTM D 3244, Standard Practice for Utilization of Test Data to Determine Conformance with Specifications, ASTM Standards on PRECISION and BIAS for VARIOUS APPLICATIONS, fourth edition, American Society for Testing and Materials, Philadelphia, PA, 1992.

[2] Caulcutt R. and Boddy R., Statistics for Analytical Chemists, Chapman and Hall, New York NY, 1989

[3] Montgomery, D.C., Introduction to Statistical Quality Control, second edition, Wiley, NY, 1992

[4] ASTM STP 15 D, Manual on presentation of data and Control Chart Analysis, Manual 7, American Society for Testing and Materials, Philadelphia, PA

[5] Hunter, J. S., "The Exponentially Weighted Moving Average", Journal of Quality Technology, Vol.18 No. 4, October 1986, pp.203-210.

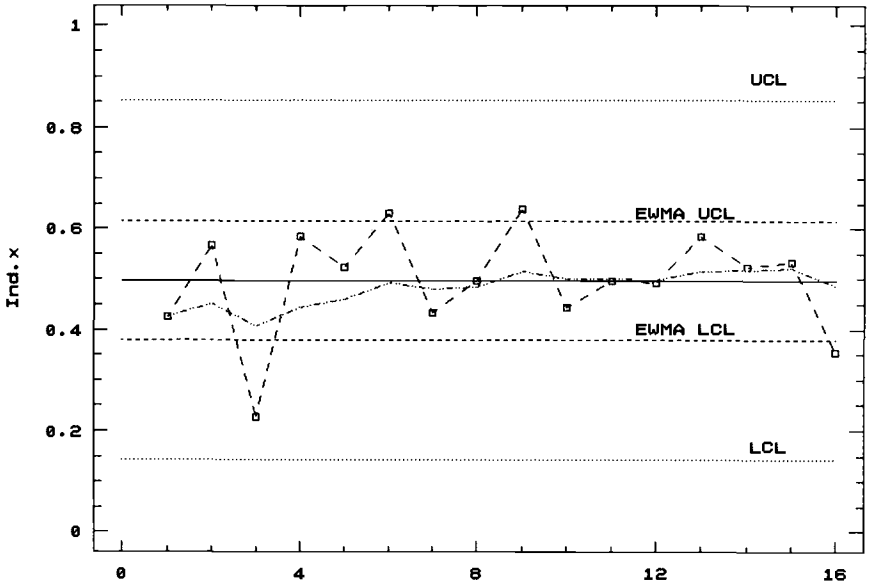


FIG. 5a--I, EWMA chart for relative bias

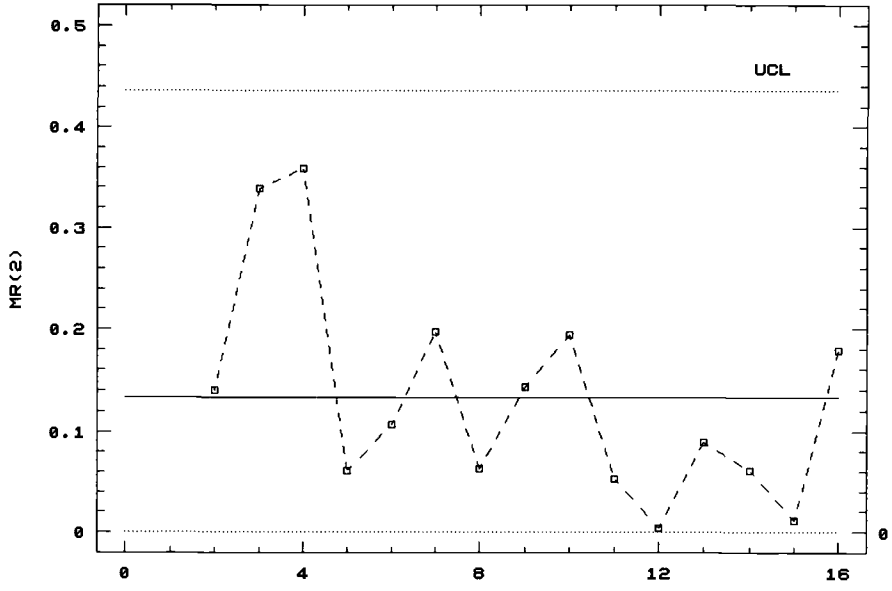


FIG. 5b--MR2 chart for relative bias

W. Virgil Ping[1]

Statistical Evaluation of Additive Effectiveness in Asphalt Paving Mixtures

REFERENCE: Ping, W. V., "Statistical Evaluation of Additive Effectiveness in Asphalt Paving Mixtures," Quality and Statistics: Total Quality Management, ASTM STP 1209, Milton J. Kowalewski, Jr., Ed., American Society for Testing and Materials, Philadelphia, 1994.

ABSTRACT: Asphalt pavements are susceptible to moisture damage. One of the procedures to eliminate or minimize the damage is to treat the asphalt paving mixtures with an antistripping agent such as hydrated lime or other commercially available antistripping additives. Statistical analysis is essential to evaluate the effectiveness of various treatments in asphalt mixtures. This paper summarizes a research study to evaluate the effectiveness of asphalt treatments utilizing statistical pairwise comparisons. The approach is to use the boiling test results to do the statistical comparison.

KEYWORDS: Asphalt mixtures, boiling test, antistripping additives, statistical comparison, Student-t test.

Moisture damage is a major problem for asphalt pavements constructed throughout much of the United States. As a result of research a number of tests and test procedures have been developed to evaluate the moisture damage potential of asphalt-aggregate mixtures. In addition, a number of procedures and recommendations to eliminate or minimize moisture damage have been formulated. One of these procedures involves treating the asphalt paving mixtures with an antistripping agent such as hydrated lime or other commercially available antistripping additives [1,2].

The research described in this paper was undertaken to evaluate the effectiveness of hydrated lime and other selected antistripping additives using the Texas boiling test which is a basic moisture susceptibility test method [3,4]. Statistical comparisons were utilized to evaluate the

[1]Assistant Professor, Department of Civil Engineering, FAMU/FSU College of Engineering, Tallahassee, Florida 32316.

effectiveness using the boil values. The experimental program and test results are discussed below.

EXPERIMENTAL PROGRAM

The objective of this study was to evaluate the effectiveness of hydrated lime and selected antistripping additives in protecting asphalt mixtures from moisture damage using statistical comparisons. To achieve the objective, the experimental program involved aggregates and asphalts from eight different highway districts in Texas [5], and thirteen commercially available antistripping additives and the hydrated lime. The Texas boiling test was performed on treated and untreated mixtures which were 1) mixed in the plant (plant mixtures), and 2) mixed in the laboratory (laboratory mixtures).

Materials

Plant mixtures--Loose samples of the hot asphalt mixtures utilized in actual field construction were obtained at the eight asphalt mixing plants. The types of aggregate and the source and amount of asphalt cement for the plant mixtures are summarized on Table 1. Two or more liquid antistripping additives and hydrated lime were used in each type of plant mixture with identical raw material sources (aggregates and asphalt cement). Fourteen different antistripping additives, including hydrated lime, were used in the eight plant mixtures. The selected additives and actual dosages are summarized on Table 2. The percentage of lime is by the total weight of dry aggregates, whereas the percentage of liquid additives is by the weight of asphalt cement.

Laboratory mixtures--The asphalt cements, aggregates, liquid antistripping additives and hydrated lime were obtained at the asphalt mixing plants. In the laboratory these materials were prepared and mixed using the laboratory mixing procedures in accordance with the mixture design established for the plant mixture. The laboratory additive dosage levels were essentially the same as those for the plant mixtures (Table 2). With the exception of the hydrated lime, the additives were blended with the preheated asphalt. The hydrated lime was placed on the aggregates in a slurry form for all of the lime treated laboratory mixtures.

Moisture Susceptibility Test

The Texas boiling test was used to evaluate the moisture susceptibility of the plant and laboratory mixtures. The boiling test [3,4] involved a visual determination of the extent of stripping of the asphalt from aggregate surfaces

TABLE 1--Summary of materials used in experimental program.

Location	Aggregates	Asphalt	Asphalt Content, %
Dist. 17 Hearne	.Processed gravel 55% .Washed sand 25% .Coarse sand 10% .Fine sand 10%	.AC-20 .Texas Gulf Refinery	4.9
Dist. 16 Odom	.Field sand 20% .Limestone Screenings 22% .Coarse Limestone 58%	.AC-20 .Gulf States Refinery	5.1
Dist. 13 Victoria	.Crushed .gravel 50% .Limestone 10% .Limestone screenings 20% .Field Sand 20%	.AC-20 .Texas Fuels & Asphalt Refinery	5.0
Dist. 6 Midland	.Rhyolite 56% .Screening 37% .Field Sand 7%	.AC-20 .American Petrofina Refinery	6.2
Dist. 25 Childress	.Coarse Aggr. 20% .Inter. Aggr. 34% .Screening 46%	.AC-20 .Diamond Shamrock Refinery	5.2
Dist. 1 Scherman	.Coarse sandstone 55% .Unwashed screenings 30% .Field sand 15%	.AC-20 .Total Petroleum Refinery	5.5
Dist. 19 DeBerry	.Coarse Aggregate 20% .Inter. Aggregate 40% .Screening 20% .Field sand 20%	.AC-20 .Lion Oil Refinery	5.6
Dist. 21 Cameron	.Coarse Aggregate 35% .Uncrushed aggregate 20% .Screening 25% .Field sand 20%	.AC-10 .Texas Fuel & Asphalt Coastal Refinery	5.2

TABLE 2--Type and dosage of antistripping additives used in experimental program.

Location	Test Sections	Additive Designation	Additive Dosage*, %
Dist. 17	.Control	0	0
	.Lime	1	1.5
Hearne,	.BA 2000	5	1.0
TX	.Perma-Tac	12	1.0
	.Control	0	0
Dist. 16	.Lime	1	1.0
	.Aquashield	3	0.5
Odom,	.Dow Anti-Strip	6	0.41
TX	.Pavebond LP	10	0.5
Dist. 13	.Control	0	0
	.Lime	1	2.0
Victoria	.BA 2000	5	1.0
TX	.Perma-Tac Plus	13	1.0
	.Control	0	0
Dist. 6	.Lime	1	1.0
	.Pavebond LP	10	1.0
Midland,	.Perma-Tac	12	1.0
TX	.Unichem	14	1.0
	.Control	0	0
Dist. 25	.Lime	1	1.0
	.Aquashield II	4	1.0
Childress,	.Fina-A	7	1.0
TX	.Perma-Tac	12	1.0
	.Unichem	14	1.0
	.Control	0	0
Dist. 1	.Lime	1	1.5
	.ARR-MAZ	2	0.75
	.Dow Anti-Strip	6	0.45
Sherman,	.Fina-A	7	1.0
TX	.Indulin AS-1	9	1.0
	.Pavebond Special	11	1.0
	.Perma-Tac Plus	13	1.0
	.Control	0	0
Dist. 19	.Lime	1	1.0
	.ARR-MAZ	2	1.0
	.Aquashield II	4	0.8
DeBerry,	.BA 2000	5	0.5
TX	.Perma-Tac	12	1.0
	.Control	0	0
District 21	.Lime	1	1.0
	.ARR-MAZ	2	1.0
	.Aquashield II	4	0.41
Cameron,	.Dow Anti-Strip	6	0.5
TX	.Fina-B	8	0.41
	.Pavebond LP	10	1.0
	.Perma-Tac	12	1.0

* The percentage of lime is by the total weight of dry aggregates; percentage of liquid additives is by the weight of asphalt cement.

after the mixture had been subjected to the action of boiling water for a specified time. To perform this test an asphalt mixture was prepared at 325 F (162.8 C) and boiled in distilled water for 10 minutes. After boiling, the mixture was allowed to cool, the water was drained, and the mixture was allowed to dry. The mix was examined the following day to estimate the degree of stripping present in the mixture. The stripping test results were reported as the percent of asphalt retained after boiling.

EXPERIMENTAL RESULTS

The treated and untreated, laboratory and plant mixtures were prepared and tested in the laboratory. The boiling test results are summarized in Table 3. The boil value is expressed as the percent of asphalt retained after boiling. The value is visually estimated by two independent operators according to the degree of stripping present in the mixture. Each presented boil value is an average of two visual estimates as described previously.

In addition, core samples were obtained from the field test sections after construction. The boiling test was conducted on the loose core mixtures. The procedure was the same as described before for the plant mixtures. The test results for the field cores are also reported herein (Table 3) so that a larger sample size can be used for the statistical evaluation.

STATISTICAL ANALYSIS

The statistical approach was utilized to evaluate the effectiveness of selected liquid additives and hydrated lime. The approach was to use the boiling test results to do statistical comparisons. A detailed description of the statistical analysis and the comparisons utilized in the evaluation are discussed in the following sections.

Pooled Variance for Boil Values

The sample variance is defined as the following [6,7]:

$$S^2 = \sum_{i=1}^{n} \frac{(x_i - \bar{x})^2}{(n-1)} \qquad (1)$$

where, s^2 = sample variance
X_i = sample values of variable X

TABLE 3--Summary of boiling test results.

District	Additive Name	Asphalt Retained After Boiling, %		
		Lab Mix	Plant Mix	Field Core
17	No Additive	50.0	52.5	50.0
	Lime	85.0	94.0	92.5
	BA 2000	92.5	92.5	92.5
	Perma-Tac	90.0	50.0	90.0
16	No Additive	77.5	82.5	75.0
	Lime	75.0	82.5	77.5
	Aquashield	77.5	85.0	82.5
	Dow	77.5	85.0	75.0
	Pavebond LP	77.5	85.0	72.5
13	No Additive	77.5	77.5	82.5
	Lime	96.5	96.5	90.0
	BA 2000	97.5	96.5	92.5
	Perma-Tac	96.5	95.0	92.5
6	No Additive	50.0	70.0	60.0
	Lime	72.5	72.5	60.0
	Pavebond LP	60.0	85.0	87.5
	Perma-Tac	65.0	80.0	75.0
	Unichem	67.5	85.0	90.0
25	No Additive	50.0	77.5	72.5
	Lime	85.0	87.5	85.0
	Aquashield II	96.5	77.5	72.5
	Fina-A	94.0	94.0	92.5
	Perma-Tac	90.0	92.5	87.5
	Unichem	94.0	87.5	77.5
1	No Additive	82.5	90.0	91.5
	Lime	92.5	92.5	95.0
	ARR-MAZ	90.0	97.5	95.0
	Dow	82.5	91.5	92.5
	Fina-A	92.5	95.0	97.5
	Indulin AS-1	92.5	96.5	95.0
	PVBD Special	92.5	95.0	97.5
	Perma-Tac Plus	92.5	95.0	95.0
19	No Additive	85.0	85.0	87.5
	Lime	94.0	90.0	85.0
	ARR-MAZ	92.5	90.0	90.0
	Aquashield II	92.5	94.0	90.0
	BA 2000	92.5	96.5	90.0
	Perma-Tac	92.5	90.0	90.0
21	No Additive	37.5	25.0	25.0
	Lime	81.0	37.5	45.0
	ARR-MAZ	55.0	57.5	50.0
	Aquashield II	77.5	67.5	67.5
	Dow	57.5	52.5	47.5
	Fina-B	80.0	75.0	65.0
	Pavebond LP	65.0	67.5	62.5
	Perma-Tac	55.0	61.0	60.0

$\bar{x}$ = sample mean of the n sample values

n = sample size

The pooled variance of two samples with equal size n can be obtained as the following:

$$s^2 = \frac{s_1^2 + s_2^2}{2} = \frac{(\sum x_1^2 + \sum x_2^2)}{2(n-1)} \quad (2)$$

where, s^2 = pooled sample variance of samples x_1 and x_2

s_1^2 = sample variances of x_1

s_2^2 = sample variances of x_2

x_1 = deviations from the sample mean (sample X_1)

x_2 = deviations from the sample mean (sample X_2)

n = sample size of x_1, x_2

With unequal sample sizes n_1 and n_2, the pooled variance of two samples are then:

$$s^2 = \frac{(\sum x_1^2 + \sum x_2^2)}{(n_1 + n_2 - 2)} \quad (3)$$

where, n_1 = sample size of x_1
n_2 = sample size of x_2

Following these equations (Eqs 1, 2, and 3), the pooled variances for the laboratory and plant mixtures of all eight projects were calculated and summarized in Table 4.

Standard Error Estimate of Boil Values

The standard error estimate of the boil values was computed using the pooled variances. The results are also summarized in Table 4.

Pairwise Comparison of Student-t Test Using Boil Values

The effectiveness of the various liquid antistripping additives and hydrated lime was estimated in terms of the boiling test values. The boil values of two additives were compared using the Student-t test. The Student-t test is

TABLE 4--Summary of pooled variance and standard error estimate for boiling test results.

District	Additive Name	Degree of Freedom	Pooled Variance	Standard Error Estimate
17	No Additive	8	14.94	3.86
	Lime	8	14.94	3.86
	BA 2000	8	14.94	3.86
	Perma-Tac	8	14.94	3.86
16	No Additive	10	12.50	3.54
	Lime	10	12.50	3.54
	Aquashield	10	12.50	3.54
	Dow	10	12.50	3.54
	Pavebond LP	10	12.50	3.54
13	No Additive	8	19.44	4.41
	Lime	8	19.44	4.41
	BA 2000	8	19.44	4.41
	Perma-Tac	8	19.44	4.41
6	No Additive	10	28.75	5.36
	Lime	10	28.75	5.36
	Pavebond LP	10	28.75	5.36
	Perma-Tac	10	28.75	5.36
	Unichem	10	28.75	5.36
25	No Additive	12	34.42	5.87
	Lime	12	34.42	5.87
	Aquashield II	12	34.42	5.87
	Fina-A	12	34.42	5.87
	Perma-Tac	12	34.42	5.87
	Unichem	12	34.42	5.87
1	No Additive	16	65.72	8.11
	Lime	16	65.72	8.11
	ARR-MAZ	16	65.72	8.11
	Dow	16	65.72	8.11
	Fina-A	16	65.72	8.11
	Indulin AS-1	16	65.72	8.11
	PVBD Special	16	65.72	8.11
	Perma-Tac Plus	16	65.72	8.11
19	No Additive	12	59.88	7.74
	Lime	12	59.88	7.74
	ARR-MAZ	12	59.88	7.74
	Aquashield	12	59.88	7.74
	BA 2000	12	59.88	7.74
	Perma-Tac	12	59.88	7.74
21	No Additive	16	33.38	5.78
	Lime	16	33.38	5.78
	ARR-MAZ	16	33.38	5.78
	Aquashield II	16	33.38	5.78
	Dow	16	33.38	5.78
	Fina-B	16	33.38	5.78
	Pavebond LP	16	33.38	5.78
	Perma-Tac	16	33.38	5.78

expressed as:

$$t = \frac{(r_1 - r_2)}{Std.Er.of(r_1 - r_2)} \qquad (4)$$

where, r_1 and r_2 = Boil values of two different additives

t = Student-t value with degrees of freedom (d.f.) as for the pooled variance of the additives

The standard error (S.E.) of the difference $(r_1 - r_2)$ can be obtained as the following:

$$[S.E.\ (r_1 - r_2)]^2 = [S.E.(r_1)]^2 + [S.E.(r_2)]^2 \qquad (5)$$

where, S.E. $(r_1 - r_2)$ = Standard error of the difference $(r_1 - r_2)$

S.E.(r_1) = Standard error of r_1
S.E.(r_2) = Standard error of r_2

The boil values of any two additives (or control) were compared using (Eq 4) with the standard error estimate of the difference in boil values between two additives (or control) obtained from (Eq 5). The procedure is illustrated in Table 5 using typical test results from the District 17 mixtures. The standard error of the difference $(r_1 - r_2)$ for each possible pair were calculated using (Eq 5) and presented as part (a) of the table. Then, the pairwise comparisons of Student-t test results were obtained using (Eq 4) and presented on part (b) of the table on a four by four matrix. A spread sheet program was used to perform the computation. Following this procedure, all of the eight projects was analyzed and the possible pairs were compared using the Student-t test. The detailed results of the pairwise comparisons can be found elsewhere [5].

Significance Level for Statistical Comparison

The α-level of 5% error rate was adopted to test the statistical significance between the difference of two boil values. An asterisk (*) was added to a critical value on (Table 5) to denote the significance at the 5% level.

However, objectives to the use of a single α-level in multiple comparisons have been raised. The probability that

TABLE 5--Typical tabulation of the pairwise comparison Student-t test procedure (District 17 mixtures)

Laboratory Mixture:
(a) Standard error of the difference between two additives.

Additive	ADD. 0	ADD. 1	ADD. 5	ADD. 12
Standard Error	3.86	3.86	3.86	3.86
ADD. 0	-	-	-	-
ADD. 1	5.46	-	-	-
ADD. 5	5.46	5.46	-	-
ADD. 12	5.46	5.46	5.46	-

(b) Pairwise Student-t test results between two additives.

Additive	ADD. 0	ADD. 1	ADD. 5	ADD. 12
Asphalt, %	50.0	85.0	92.5	90.0
ADD. 0	-	-	-	-
ADD. 1	-6.412*+	-	-	-
ADD. 5	-7.786*+	-1.374	-	-
ADD. 12	-7.328*+	-0.916	0.458	-

Plant Mixture:
(a) Standard error of the difference between two additives.

Additive	ADD. 0	ADD. 1	ADD. 5	ADD. 12
Standard Error	3.86	3.86	3.86	3.86
ADD. 0	-	-	-	-
ADD. 1	5.46	-	-	-
ADD. 5	5.46	5.46	-	-
ADD. 12	5.46	5.46	5.46	-

(b) Pairwise Student-t test results between two additives.

Additive	ADD. 0	ADD. 1	ADD. 5	ADD. 12
Asphalt, %	52.5	94.0	92.5	50.0
ADD. 0	-	-	-	-
ADD. 1	-7.602*+	-	-	-
ADD. 5	-7.328*+	0.275	-	-
ADD. 12	0.458	8.060*+	7.786*+	-

* Results significant at the 5% error (t*=2.306, d.f.=8).
\+ Significant according to the Bonferroni method with the 5% error rate (t+=3.482).

one of the multiple comparisons exceeds the 5% level is bound to happen. There are several methods available to protect against making erroneous decisions in multiple comparisons; however, one of the methods which is called Bonferroni's method [6] is best suited for this application. The Bonferroni method for multiple comparisons is to reset the α-level to α/N, where N is the number of comparisons, and use Student-t table entries for the significance level of α/N. For example, if the α-level of 5% error rate needs to be achieved, the t-value of $t_{0.05}/N$ from the t-table with appropriate degrees of freedom should be used to test the statistical significance for N multiple comparisons instead of using the t-value of $t_{0.05}$ from the t-table.

The Bonferroni method for multiple comparisons was used to test the significance between the difference of two boil values. The outcomes are also illustrated in Table 5 with a plus sign (+) to a critical value to denote the significance using the Bonferroni criteria.

EVALUATION OF EFFECTIVENESS

The effectiveness of the hydrated lime and the various antistripping additives was evaluated using the statistical comparisons based on the boil values. The α-level of 5 percent error rate and the Bonferroni method for multiple comparisons were used to test the statistical significance of the various additives. The Bonferroni method was more stringent than the α-level of 5 percent error rate; nevertheless, the two test criteria did not provide major differences with regard to the effectiveness of the various additives (or control). The effectiveness of the various additives (or control) at 5 percent significance level is summarized in Table 6 for the laboratory and plant mixtures based on the pairwise comparisons of the boil values.

Hydrated Lime

The hydrated lime was effective as compared to control for most of the laboratory mixtures as summarized in Table 7; however, the lime was not significantly better than control for the limestone material (District 16) and the crushed gravel aggregates (Districts 1 and 19). The reason was probably due to the high boil values of the control mixtures (Districts 16, 1, and 19); thus, the boil values of the lime treated mixtures were not significantly different from the control mixtures for the laboratory mixtures. As for the plant mixtures, in most cases the hydrated lime was not effective (Table 7). In addition to the materials which were not effective for the lime treated laboratory mixtures (Districts 16, 1, and 19), the lime was not effective on the plant mixtures of Districts 6, 25, and 21. For the plant mixture of District 6, the lime was probably not mixed well

TABLE 6--Summary of effectiveness of antistripping additives based on pairwise comparisons of boiling test results at 5% significance level

District	Effectiveness of Antistripping Additives or Control (In Decreasing Order of Effectiveness)	
	Lab	Plant
17	[BA 2000 Perma-Tac Lime] Control	[Lime BA 2000] [Perma-Tac Control]
16	[Aquashield Dow Pavebond LP] [Control Lime]	[Aquashield Dow Pavebond LP] [Lime Control]
13	[BA 2000 Perma-Tac Lime] Control	[Lime BA 2000] [Perma-Tac Control]
6	[Lime Unichem] [Perma-Tac Pavebond LP Control]	[Pavebond LP Unichem Perma-Tac Lime Control]
25	[Aquashield II Fina-A Unichem Perma-Tac Lime] Control	[Fina-A Perma-Tac Lime Unichem Aquashield II Control]
1	[Lime Fina-A Indulin AS-1 P.B. Special P.T. Plus ARR-MAZ Control Dow]	[ARR-MAZ Indulin AS-1 Fina-A P.B. Special P.T. Plus Lime Dow Control]
19	[Lime ARR-MAZ Aquashield II BA 2000 Perma-Tac Control]	[BA 2000 Aquashield II Lime ARR-MAZ Perma-Tac Control]
21	[Lime Fina-B Aquashield II Pavebond LP] [Dow ARR-MAZ Perma-Tac] Control	[Fina-B Aquashield II Pavebond LP Perma-Tac ARR-MAZ Dow] [Lime] Control

* The bracket indicates that the additives are not statistically different within the same bracket.

TABLE 7--Effectiveness* of antistripping additives as compared to control using boiling test results

Additives	Laboratory/Plant Mixture	SDHPT District No. 17	16	13	6	25	1	19	21
Hydrated Lime	Lab	E	N	E	E	E	N	N	E
	Plant	E	N	E	N	N	N	N	N
Aquashield	Lab		N						
	Plant		N						
Aquashield II	Lab					E		N	E
	Plant					N		N	E
ARR-MAZ	Lab						N	N	E
	Plant						N	N	E
BA 2000	Lab	E		E				N	
	Plant	E		E				N	
Dow	Lab		N				N		E
	Plant		N				N		E
Fina-A	Lab					E	N		
	Plant					N	N		
Fina-B	Lab								E
	Plant								E
Indulin AS-1	Lab						N		
	Plant						N		
Pavebond LP	Lab		N		N				E
	Plant		N		N				E
Pavebond Special	Lab						N		
	Plant						N		
Perma-Tac	Lab	E			N	E		N	E
	Plant	N			N	N		N	E
Perma Tac Plus	Lab			E			N		
	Plant			E			N		
Unichem	Lab				E	E			
	Plant				N	N			

* At 5 percent significance level.
E -Effective at 5% significance level as compared to control.
N -Not effective at 5% significance level as compared to control.

with the rhyolite aggregates in the plant. For the District 25 material, the plant mixture of control material had a fairly high boil value; therefore, the lime treated mixture was not significantly different from the control material. As for the District 21 material, the problem was associated either with the raw material itself or with the field application technique of the lime.

Liquid Antistripping Additives

Most liquid additives were effective in preventing moisture damage using the boil values for Districts 17, 13, 25 (except plant mixture), and 21 (all with gravel aggregates).

However, the liquid additives were not effective at all for Districts 16 (limestone), 6 (rhyolite), 1 (sandstone), and 19 (crushed gravel). The reason was possibly due to the high boil test values for control materials for Districts 16, 1, and 19. As for District 6, only the Unichem was effective on the laboratory mixture, whereas the other additives were not effective on the rhyolite material.

CONCLUSIONS

The conclusions based on the data and analyses from this study are summarized below.

1. Statistical comparisons were successfully used to evaluate the effectiveness of various treatments in asphalt mixtures.

2. The hydrated lime was effective on treating the moisture susceptible materials from stripping.

3. The liquid additives in general were not very effective on treating the limestone as well as the sandstone materials. However, most liquid additives were effective for the river gravels.

ACKNOWLEDGEMENTS

Funding for this research was provided by the Texas State Department of Transportation and the Federal Highway Administration. The study was conducted at the Center for Transportation Research of the University of Texas at Austin. The research reported here was part of a larger research project directed by Dr. Thomas W. Kennedy. Dr. David V. Hinkley with the UT-Austin reviewed the statistical analysis.

REFERENCES

[1] Lottman, R. P., "Laboratory Test Method for Predicting Moisture-Induced Damage to Asphalt Concrete," Transportation Research Record 843, 1982.

[2] Tunnicliff, D. G. and Root, R. E., "Antistripping Additives in Asphalt Concrete for Effectiveness of Antistripping Additives," Proceeding of the Association of Asphalt Paving Technologists, Vol. 53, 1983.

[3] Texas State Department of Highways and Public Transportation, "Manual of Testing Procedures," Bituminous Section, 1983.

[4] Kennedy, T. W., Roberts, F. L., and Lee, K. W., "Evaluation of Moisture Susceptibility of Asphalt Mixtures Using the Texas Boiling Test," Transportation Research Record 968, 1984.

[5] Ping, W. V. and Kennedy, T. W., "Evaluation of Stripping and Moisture Damage in Asphalt Pavements Treated with Lime and Antistripping Agents," Research Report 441-1, Center for Transportation Research, The University of Texas at Austin, October, 1990.

[6] Snedecor, G. W. and Cochran, W. G., "Statistical Methods," Seventh Edition, The Iowa State University Press, Ames, Iowa, 1980.

[7] Li, J. C. R., "Statistical Inference," Vol. I & II, Edwards Brothers, Inc., Ann Arbor, Michigan, 1966.

ISBN 0-8031-1872-4