SECTION 23
INSPECTION AND TEST[1]

E. F. "Bud" Gookins

[1] In the Fourth Edition, the material on inspection and test was prepared by Joseph J. Zeccardi.

INTRODUCTION

Inspection and testing activities always involve the evaluation of a characteristic as it relates to a specific requirement. The requirement can be in the form of a standard, a drawing, a written instruction, a visual aid, or any other means of conveying the characteristic specification.

Inspection and testing functions can be done automatically, manually, or both in a sequential manner. The evaluation process consists of the following steps applied to each characteristic (Juran 1945, p. 23):

1. Interpretation of the specification
2. Measurement of the quality of the characteristic
3. Comparing 1 with 2
4. Judging conformance
5. Processing of conforming items
6. Disposition of nonconforming items
7. Recording of data obtained

These steps apply to both product and service items.

The inspection and testing evaluation can be determined by using the intrinsic senses of the human being (i.e., smell, taste, sight, hearing, and touch), or it can be made using a nonvariable gage, a nonvariable electronic or laser instrument, a nonvariable chemical or physical testing device, or any other method in which a decision is made based on simply an "accept" or "reject" determination. Such inspection is commonly referred to as *attribute inspection.*

The inspection and testing evaluation that is determined by using any measurement device, be it mechanical, electronic, laser, chemical, or any other method that will display data generated by physically measuring the characteristic, in which a decision is made based on actual value readout, is commonly referred to as *variable inspection.*

The primary purpose of inspection and testing is to determine whether products or services conform to specification. This purpose is often called *acceptance inspection* and *acceptance testing.* The components of the inspection and testing function can be broken down further into subclassifications. The most salient are listed in Table 23.1.

People engaged full time in inspection work commonly carry the title of *inspectors* but often are recognized as product appraisers, product auditors, and product verifiers for organizations that produce a manufactured item. For organizations engaged in nonmanufactured goods or services, the

TABLE 23.1 Subclassifications for Manufacturing Products and Service

Type	Prime function
Receiving (or incoming) inspection and testing	To ensure that incoming product is not used or processed until it has been inspected or tested and found to be conforming to specified requirements
In-process inspection and testing	To ensure that in-process product is not moved forward until it has been inspected or tested and found to be conforming to specified requirements
Final inspection and testing	To ensure that finished component and/or product is not dispatched until all the activities have been satisfactorily completed
Layout inspection and functional testing	To ensure that all customer engineering material and performance standards have been appraised prior to production
Shipping inspection and testing	To ensure that all shipped products are conforming to specified requirements
Qualification inspection and testing	To judge the service capability of the product and the possible extreme applications of the product
Dock inspection and auditing	To ensure that the product and its testing (product-related packaging, identification, information, etc.) are released to the customer conforming to all requirements
Service inspection (health inspector, environmental inspector, etc.)	To ensure that all specified requirements are met and to evaluate and measure nonconformancies found in the system
Nonproduction inspection and testing	To evaluate a specific task requested by quality assurance i.e., gage repeatability and reproducibility, process capability, 100% appraisal, etc.
Initial sample inspection request (ISIR)	To assure the customer that the first production run will be in conformance with all their designated characteristics and to submit those actual characteristic measurements and attributes to the customer for verification and approval
Production part approval process (PPAP)	A request by the customer indicating the level of inspection to conduct first production run checking all characteristics and indicating actual dimensions, gage repeat-ability and reproducibility studies, capability studies, material verification, or any other outside processing; e.g., heat-treating, plating, etc., and submit to the customer for approval (used primarily by the auto industry)

inspection work is identified pertaining to that particular function, e.g., safety inspectors, environmental inspectors, health inspectors, etc. People engaged in part-time inspection work are commonly referred to by the title of their major activities, e.g., machine operators, setup people, assemblers, welders, platers, foundry people, etc. The people engaged full time in the testing function commonly carry the title of *tester* but often are recognized as assaying technicians, laboratory technicians, chemists, metallurgists, etc.

Today, many manufacturing organizations have moved toward total automated inspection and testing or semiautomated inspection or testing augmented by productive operation appraisal. Within this movement of proactive product acceptance, the inspection and testing become integrated into the operational function and verified usually by a new breed of inspectors called *audit inspectors.*

PURPOSE OF INSPECTION AND TEST

The purpose of inspection and test is to determine the conformance of the product or service to the standard or specific requirements and to disposition the product or service based on the results of the evaluation. This determination involves three main decisions (Juran and Gryna 1980):

Conformance decision: To judge whether the product conforms to specification

Fitness-for-use decision: To decide whether nonconforming product is fit for use

Communication decision: To decide what to communicate to outsiders and insiders

The Conformance Decision. Except in small companies, the number of conformance decisions made per year is simply huge. There is no possibility for the supervisory body to become involved in the details of so many decisions. Hence the work is organized in such a way that the inspectors or production workers can make these decisions. To this end, they are trained to understand the products, the standards, and the instruments. Once trained, they are given the job of making the inspections and of judging conformance. (In many cases the delegation is to automated instruments.)

Associated with the conformance decision is the disposition of conforming product; the inspector is authorized to identify the product ("stamp it up") as an acceptable product. This identification then serves to inform the packers, shippers, etc. that the product should proceed to its next destination (further processing, storeroom, customer). Strictly speaking, this decision to "ship" is made not by the inspectors but by management. With some exceptions, a product that conforms to specification is also fit for use. Hence the company procedures (which are established by the managers) provide that conforming products should be shipped as a regular practice.

The Fitness-for-Use Decision. In the case of nonconforming product, a new question arises: Is this nonconforming product fit for use or unfit? In some cases the answer is obvious—the nonconformance is so severe as to make the product clearly unfit. Hence it is scrapped or, if economically repairable, brought to a state of conformance. However, in many cases the answer as to fitness for use is not obvious. In such cases, if enough is at stake, a study is made to determine fitness for use. This study involves securing inputs such as those shown in Table 23.2.

Once all the information has been collected and analyzed, the fitness-for-use decision can be made. If the amount at stake is small, this decision will be delegated to a staff specialist, to the quality manager, or to some continuing decision-making committee such as a material review board. If the amount at stake is large, the decision usually will be made by a team of upper managers.

The Communication Decision. Inspection and test serve two purposes: to make decisions on the products and to generate data that provide essential information for a wide variety of uses, such as those listed in Table 23.1. The conformance and fitness-for-use decisions likewise are a source of essential information, although some of this is not well communicated.

Data on nonconforming products are usually communicated to the producing departments to aid them in preventing a recurrence. In more elaborate data-collection systems there may be periodic summaries to identify "repeaters" or the "top 10," which then become the subject of special studies.

When nonconforming products are sent out as fit for use, there arises the need for two additional categories of communication:

1. *Communication to "outsiders":* They (usually customers) have a right and a need to know. All too often the manufacturing companies neglect to inform their customers when shipping nonconforming products. This may be as a result of bad experience; i.e., some customers will seize on such nonconformances to secure a price discount despite the fact that use of the product will not add to their costs. Usually, the neglect indicates a failure even to face the question of what to communicate. A major factor here is the design of the forms used to record the decisions. With rare excep-

TABLE 23.2 Inputs Required for Fitness-for-Use Decision

Input	Usual sources
Who will be the user?	Marketing
How will the nonconforming product be used?	Marketing, client
Are there risks to human safety or structural integrity?	Product research and design
What is the urgency?	Marketing, client
What are the company's and the users' economics?	All departments, client
What are the users' measures of fitness for use?	Market research, marketing, client

tions, these forms lack provisions that force those involved to make recommendations and decisions on (a) whether to inform the outsiders and (b) what to communicate to them.

2. *Communication to insiders:* When nonconforming goods are shipped as fit for use, the reasons are not always communicated to the inspectors and especially not to the production workers. The resulting vacuum of knowledge has been known to breed some bad practices. When the same type of nonconformance has been shipped several times, an inspector may conclude (in the absence of knowing why) that it is just a waste of time to report such nonconformances in the first place. Yet in some future case the special reasons (which were the basis of the decision to ship the nonconforming goods) may not be present. In like manner, a production worker may conclude that it is a waste of time to exert any effort to avoid some nonconformance that will be shipped anyway. Such reactions by well-meaning employees can be minimized if the company faces squarely the question: What shall we communicate to the insiders?

PREPRODUCTION EVALUATION

The inspection and testing functions are key elements of the production process. Without accurate and specific criteria for determining that the manufacturing or service product meets the customer's requirements, we expose the organization to uncontrolled, inefficient, and expensive processing as well as negative perceptions from customers. These resulting performances can be minimized—if not eliminated—by preproduction and service evaluations.

The approach to inspection and test planning follows closely the principles of quality planning as set out in Section 3, The Quality Planning Process. Application of these principles to the inspection job has been studied extensively, and good tools are available to facilitate inspection planning.

The Nature of "Lots" of Product. It is useful here to define what is meant by *lot* and expand briefly on the term as applied to inspection and test. A *lot* is usually associated with physical product, especially in connection with sample inspection and test. Usually the product submitted for decision on conformance to standard consists of a lot. The true lot is an aggregation of product made under a common system of causes. When this ideal is met, the lot possesses an *inherent uniformity derived from the common system of causes.* The extent to which the lot conforms to this ideal greatly influences the approach to the product conformance decision and especially the kind and extent of sampling.

In its simplest form, the true lot emerges from one machine run by one operator processing one material batch, all under a state of statistical control, e.g., a single formulation of a drug product or a run of screw-machine parts turned from one piece of rod on one machine. A great deal of industrial production consists of true lots.

However, a great deal of other production consists of product mixtures that, in varying degrees, fall short of the ideal lot definition. Product made from several material batches, on several machines, or by several operators may be dumped into a common container. In shop language, this mixture is a "lot," but in more precise language it is only a "mixture." In continuous processes or in conveyor production, the process may well be common and constant, but the input materials may not be.

For precise and economic product conformance decisions, it is most helpful to *preserve the order.* This means that product is kept segregated in true lots or at least identified as to common cause. In addition, for those processes which exhibit a time-to-time variation or "drift" (e.g., the solution gradually becomes dilute. The tool gradually wears), preserving the order includes preserving the time sequence during which various portions of the lot were made. Any loss of order of manufacture also becomes a loss of some prior knowledge as to inherent uniformity. (See Section 22, Operations, under Process Capability: The Concept, Process Mixture, for a discussion of the effect of product mixture on process improvement, including application in services.)

Some products are naturally fluid and develop a homogeneity through this fluidity. Homogeneity from this new cause also can qualify the product as a true lot, with important implications for the sampling process.

When several true lots are combined for the purposes of acceptance, the combination is known as a *grand lot.* Such mixtures are very common, e.g., product from multiple cavities of molding operations or from multiple spindles of screw-machine operations. The two categories of single elements of product (e.g., discrete units and specimens) have their counterparts in two categories of lots.

The Lot as a Collection of Discrete Units. Here the lot consists of numerous bolts, teacups, or refrigerators, each one of which is governed by the product specification. In batch production, the lot is usually determined by the obvious boundaries of the batch. In continuous production, the lot is usually defined as an arbitrary amount of production or as the amount produced during an arbitrary time span, e.g., a shift, a week.

The Lot as Coalesced Mass. Here the lot also may consist of a batch, e.g., the melt of steel. In continuous production, the lot is again based on some arbitrary selection, e.g., 1 ton or a day's production.

The Inspection and Test Requirement Review. Some organizations produce a standard or proprietary product that lends itself to very little change in configuration, materials, or processing, whereas some organizations are driven by ongoing changes or modifications to their products or services, and even some organizations are a "job shop" type, producing a product specifically to a customer's requirements and specifications. Regardless of the frequency or type of change to existing inspection and testing requirements, however, a review should be made prior to first production release or any subsequent revisions.

This review should examine any measurement parameter that would require special gaging and testing equipment that is different from that presently used as the method of measurement. It also should include any other input that would provide assurance that the customer's requirements will be met or that preproduction appraisal exceptions have been resolved.

The Inspection and Test Planner. The planning can be done by anyone who understands the fitness for use of the product being inspected. Usually, however, the planning is done by a quality assurance staff planner, an inspection or test supervisor, and in some situations even the inspector or tester.

Where planning is done by a staff planner, it is recommended that his or her proposal be accepted by the inspection or testing supervisor before the plan becomes effective. The staff planner also is assigned a scope of responsibility within which to work. This scope determines which aspects of inspection or test planning are to be covered: inspection instructions, instrumentation, cost estimates, space and workplace design, documentation, etc. In large organizations, the planning is sometimes divided among specialists rather than being assigned by project. In smaller organizations, the planning may be done by the head of quality or assigned to an inspector or tester.

If the inspection or testing planning is service-oriented, it is usually conducted by the functioning individual conducting the inspection/testing or an immediate supervisor. If the inspection or testing planning is manufacturing-oriented, it is usually broken down into five categories:

1. Components completed within a single department, small series production—conducted by inspector/tester.
2. Components completed within a single department, large series production—conducted by inspection/testing supervision.
3. Simple components and services, purchased or in-house heat-treating, plating, casting—conducted by inspection/testing supervision.
4. Complex units, small series production (machine tools)—conducted by inspection/testing supervision.
5. Components produced by multiple department progression, subsystem test, or interdepartment units—conducted by quality planner.

Developing the Inspection and Test Plan. For each inspection station, the planner lists the quality characteristics to be checked. To determine these, the planner considers the various sources of pertinent product information:

- The needs of fitness for use
- The product and process specifications as published by the engineers
- The customer's order, which references the product specification but may call for modifications
- The applicable industry standards and other general-use sources

For test stations, the planner must consider the functional and reliability parameters, such as

- The industrial standards
- Third-party requirements
- Application environments
- Customer expectations

For the service application, the planner must consider the characteristic criteria for the checklist method of evaluation, such as

- The needs of the customer
- Service industry standards
- The objectives and goals of the organization.

The specification information is seldom sufficient for the inspector/tester to meet the realities to be faced. The inspector/tester planner can help to bridge this gap in several ways:

1. *Clear up the meaning of the words used.* Terminology for describing sensory qualities is often confusing. In one company, the term *beauty defects* was used generally to describe blemishes on the products. Some of these blemishes (scratches in the focal plan of an optical instrument) made the product unfit for service. Other blemishes, though nonfunctional, could be seen by the customers and were objectionable for aesthetic reasons. Still other nonfunctional blemishes could be seen by the company inspectors but not by consumers. However, because the multiple meanings of the term *beauty defect* had not been clarified, the inspectors rejected all blemishes. Data analysis showed that most of the blemishes were both nonfunctional and nonoffensive to customers. Hence new terminology was created to make the distinctions needed to describe the effect of blemishes. The clarification of terminology improved yields and opened the way to improvement in manufacturing processes as well. [Based on the consulting experience of J. M. Juran. For some added examples, see Juran (1952).]

2. *Provide supplemental information.* Make it available on matters for which the specification is vague or silent, e.g., workmanship. Usually this can be done for entire commodity or component classes, with minimum individual analysis. The greatest needs for supplemental standards arise in new and rapidly changing technology; in such cases it is common to find that vague standards are provided to the inspectors. Vague standards create confusion among departments as well as among companies. Refer to Miller (1975) for a discussion on specifying test methods and specifications.

3. *Classify the characteristics for seriousness.* This will help place the emphasis on the most important features of the product. (See Seriousness Classification, later in this section.) In the case of process characteristics, make use of the concept of dominance, as discussed in Section 22, Operations, under Planning Process Control.

4. *Provide samples, photographs, or other reference aids.* This will help explain the meaning of the specification. The greatest single need is for visual standards (see below).

Inspection and Test Equipment. Each product type requires a review of the gages and test equipment required prior to production. Many specifications can be satisfied in the inspection or testing

appraisal using standard inspection and testing equipment. However, some characteristics require gaging and testing equipment that must be designed specially. The planner must make this decision and schedule the equipment prior to production. The preproduction assessment should be verified prior to production release, and any discrepancy should be corrected immediately, before first-piece acceptance.

Inspection and Test Locations. Inspection and test stations usually are placed

At movement of goods between companies, usually called *supplier inspection or test*

Before starting a costly or irreversible operation, usually called *setup inspection*

At movement of goods between departments, usually called *process inspection* or *process testing*

As an integration of automatic inspection or testing within the process

On completion of the product, usually called *finished-goods inspection* or *final-product testing.*

For complex products, acceptance may require tests of mechanical compatibility, electrical mating, product performance under specified environmental conditions, and final configuration. These are usually called *systems tests.*

These general rules do not decide all questions of inspection and test stations. Complex supplier relations may require an inspection location at the supplier's plant. Some process operations may require a "station" from which the inspector patrols a large area. Other process operations may be sufficiently well in hand that no inspection stations are used between departments; instead, there is a station after completion of all operations. In assembly lines, inspection stations may be located on the line as well as at the end of the line. In still other situations there may be an added station after packing or at the customer's premises.

For each inspection location, instruct the inspector (or tester) what to inspect for and how to do it:

Just what the mission of that inspection or test station is, i.e., which qualities to check

How to determine whether a unit of product conforms to standard or not

How to determine whether a lot of product is acceptable or not (*lot criteria*)

What to do with conforming and nonconforming products

What records to make

While these categories of instruction are quite similar from one job to another, the degree of detail varies enormously.

In allocating the inspection work among the various inspection stations, the planner should be alert to the presence of "self-policing" operations. Some oversize parts will not enter tools or fixtures for further processing or cannot be assembled. Some parts are subjected to greater stresses during manufacture than during use. Some electrical circuit tests identify deficient components. Oil-pressure tests identify some undersize parts. [Refer to Trippi (1975) for a discussion on the optimal allocation of inspection resources; Ballou and Pazer (1982) for the optimal placement of inspection stations; and Eppen and Hurst (1974) for the optimal location of inspection stations in multistage production processes.]

INSPECTION AND TEST DOCUMENTED INFORMATION

As products have developed in complexity and technological advances, consistent and repetitive information properly documented is essential. The ability to appraise the product or service item the same exact way each and every time is imperative if acceptance or rejection criteria are to be cogently and enforceably judged.

Inputs into the Control Plan. The final results of inspection and test planning are reduced to writing in one of several ways.

Inspection and Test Procedure. This is a tailor-made plan for a specific component or product type. It always lists the characteristics to be checked, the method of check (e.g., visual, gage, etc.), and the instruments to be used. It may, in addition, include the seriousness classification of characteristics, tolerances and other piece criteria, list of applicable standards, sequence of inspection operations, frequency of inspection, sample size, allowable number of defects, and other lot criteria, as well as inspection stamps to be applied.

Inspection and test procedures are widely used in industry. In companies making complex systems or undergoing frequent design changes, these procedures become very numerous and consume extensive staff resources to prepare them.

The planner also should be alerted to the need for locating inspection stations at such operations as materials handling, storage, packing, and shipping. The fact that the departments doing these operations are not a part of production is of no consequence if product quality is affected.

Aspects that may require inspection planning include

- *Internal handling:* Use of correct containers and other handling facilities; product protection against corrosion, contamination, etc.
- *Internal storage:* Adequate identity and traceability
- *Packing:* Product identification, lot numbers, traceability; protection against adverse environments; protection against damage due to handling, shipping, and unpacking; presence of incidental small parts and information circulars
- *Shipping:* Care in loading; special markings required by customers

Once the planner has prepared the procedure, the interested departmental supervisors can be convened to reach agreement on who is to carry out which part of the inspection plan.

Inspection Data Planning. The planner also determines the data-recording needs for each inspection station. In many cases the standard inspection report forms will meet the recording needs. For finished products, a special test document is usually provided. In addition, the planner makes provision for any special recording needed for frequency distribution, control charts, certification, traceability, etc.

The concept of separating inspection planning from execution has great value if properly applied. If planning is underapplied, there is increased risk of catastrophic product failure. If overapplied, the result is excess cost and much internal friction. Striking a sound balance requires periodic reappraisal of the major forces in contention as well as analysis of the conventional alarm signals, e.g., rising staff costs or abrasion between departments. In addition, the changing job situations influence the extent of formal planning needed, notably (1) the education, experience, and training of the work force, (2) the stability of the processes, and (3) the severity of the product requirements.

Error-Proofing. The planner faces two responsibilities related to inspection error: (1) avoiding built-in sources of error and (2) providing positive means of foolproofing the inspection against error. See, for a detailed discussion, Inadvertent Inspector Errors, under Inspection Errors, below.

Overplanning. In some companies, the writing of inspection plans is done extensively. New customer orders, new product designs, new process changes, new regulations, and so on, are all occasions for scrutiny by the quality engineers, who issue inspection plans accordingly. As this goes to extremes, the cost of planning rises, and the excess formality increases the training time for inspectors, the attention to trivia, the documentation, and the control effort generally. Error rates tend to increase, with adverse effects on inspection costs and inspector morale.

Dealing with excess planning costs takes several forms. One technique is to do the planning by computer or by other means of mechanizing much of what the engineers otherwise do manually. A second approach is to minimize the amount of tailor-made planning by extending the use of inspection and test manuals that have broad application. See Instruction Manual, below.

A third approach is to delegate some of the planning itself to the inspection supervisors and the inspectors. To do this usually requires preparation of a manual on inspection planning plus training the inspection force to do the planning for all except the vital few characteristics, which are reserved for the staff planners. Still another device is to agree, case by case, on the amount of detailed planning needed.

Human, Machine, and System. A major decision in all planning is the extent to which tasks will be assigned to people or to machines and the related decision regarding delegation of tasks to people or to systems. Machines are superior for doing deeds that can be clearly defined and which require exacting attention to repetitive detail.

Table 23.3 is a list contrasting intellectual activities and proposes a division between person and machine as applied to inspection and test. (See also Thompson and Reynolds 1964.)

The study of the interrelationship of people, machines, and system masquerades under a variety of names: human factors, biomechanics, human engineering, ergonomics, and industrial psychology. Industrial managers, including quality managers, are commonly amateurs in the understanding of human capacities and especially human behavior. The behavioral scientists are the "professionals," but the subject is as yet hardly a science. In addition, communication between the practicing managers and the behavioral scientists is severely limited by differences in dialect and, especially, cultural background.

Procedure Manual (Includes Flow Diagram)

The Flow Diagram. The more complex the product, the greater is the need to prepare a flow diagram showing the various materials, components, and processes that collectively or sequentially turn out the final product. To prepare the flow diagram, the planner visits the various locations, interviews

TABLE 23.3 Assignment to Machines Versus People

Lower intellectual activities	Higher intellectual activities
Things that can be expressed exactly	Things that cannot be expressed exactly
Decisions that can be made in advance	Decisions that cannot be made in advance
Arithmetic, algebraic, and and chesslike symbolic logic	Pattern recognition, judgment, creativity, foresight, leadership, and such thinking
Highly repetitive and, therefore menial	Random, having many degrees of freedom, never exactly the same
Can be reduced to logic and therefore programmed exactly into a machine	Cannot be programmed exactly but can use heuristic approximations as an aid
Those a small machine can handle completely, faster and more positively	A machine cannot handle completely and it becomes excessively large and uneconomical in attempting to do so
Design and programming require a high level of intelligence but, once done, the mental activity need not be repeated	This problem is never exactly the same and it must be reconsidered, that is, rethought out for each new decision
Involves decisions as to what is right or wrong; the person guesses and the machine monitors to prevent him or her from making a mistake; it does this positively enough for use in safety systems	People use the display which is driven by the machine and possibly a separate computer to assist them in making the choice type of decisions as to what is best, using the most advanced mathematical techniques
Requires a high degree of orderliness	Takes care of matters which cannot be arranged into any sort of orderly procedures
Includes the decisions that must be made rapidly by the machine in periods of congestions and in emergencies	Involves situations that develop more slowly, that will, sooner or later, require a considered decision.

the key people, observes the activities, and records findings. The planner simplifies the picture by good use of symbols. One common set of symbols is

○	Operation	D	Delay
➪	Transportation	∇	Storage
□	Inspection	◎	Combined activity

(See also Section 3, The Quality Planning Process, for another view of flow diagraming.) In addition, the planner prepares proposals for improvement, sends copies of the diagram to all concerned, and then is ready to convene them for discussion of the diagram and the proposals.

Procedures for the inspection and testing activities to verify that specified requirements for the products are met are collected and organized in the inspection and testing procedures manual. A flow diagram should be incorporated into the procedure contents and should reflect the path the product takes and the types of inspection required along this path. Any changes to the procedures should be so noted by an ongoing document change system. All aspects of the throughput should be described in subsequent procedures with a clearly defined documentation trail back to the main or general procedure. Each organization should examine its flow diagram to determine if receiving, in-process (sometimes referred to as *patrol inspection*), final inspection (sometimes referred to as *finished-goods acceptance, dock auditing*), or special inspection (such as magnetic particle inspection, Zyglo, etc.) is applicable.

It is not uncommon to incorporate the in-process inspection function within the responsibility of the operators or even to integrate it into the machine or process. If so, this consolidation also should be spelled out as a procedure.

Instruction Manual. The instruction manual elaborates the work of each discipline in the inspection and testing functions, including detailed instructions on how to do specific work. The document should spell out the proper method of inspecting and/or testing and should be detailed as to how to fill out an inspection or test log, report, and any other record of data retention. All detailed information should be clearly defined, with the documentation trail extending bilaterally between the quality systems manual and the inspection and test instructions.

CRITERIA FOR INSPECTION AND TEST DEVELOPMENT

The factors for determining the methods and evaluation functions of appraising products or services are based on knowledge available from multiple sources.

Prior Knowledge of Product or Service Performance. In some cases, the concept of "audit of decisions" has been put to work so that suppliers, independent laboratories, workers, and so on have been qualified as able to give reliable product conformance decisions and in addition have accepted this very lot. In such cases, no further product inspection is necessary (beyond that inherent in "audit of decisions"). See, in this connection, Section 22, Operations, under Audit of Decisions.

Prior Knowledge of the Process. To illustrate, a press operation stamps out 10,000 pieces. If the first and last pieces contain certain specified holes of correct size and location, it follows that the intervening 9998 pieces also carry holes of correct size and in the correct locations. Such is the inherent nature of press dies. In statistical language, the sample size is two pieces, and the number of allowable defects is zero. Yet despite the tiny sample size, this is a sound way to do the inspection for these characteristics in the example given.

The press example is rather simple. In more complex cases, there is need to measure process capability and to arrange specially to take the samples with knowledge of the order of production. One organized form of this is the conventional control chart method used for process control. For product conformance, the approach is less well organized.

Prior knowledge of the "process" as used here includes knowledge of the qualifications of the suppliers and workers who run the process. Workers who have qualified for licenses require less rigorous inspection of their work than operators who have not qualified. Suppliers who have established a record of good deliveries need not be checked as severely as suppliers who lack such a record.

Product Homogeneity. When the product is a fluid, this fluidity contributes to homogeneity. The extent of this homogeneity can be established by taking multiple specimens and computing the dispersion (another form of study of process capability). The presence of uniformity through fluidity greatly reduces the need for random sampling and thereby greatly reduces the sample sizes.

Even when the product is a solid, the inspection planner should be alert to the possibility that it possesses homogeneity through former fluidity. For example, a centrifugal casting process was used to cast metal cylinders that were then destroyed during testing for strength. However, it was then found that the dispersion of several strength tests all made on one ring was not different from the ring-to-ring dispersion. This discovery made it possible to reduce the amount of product destroyed during test.

Economic Impact. When it is important to allocate limited testing resources to those parts which cost the most to replace, unique testing strategies can be designed that use past test, line, and field history. [See Wambach and Raymond (1977) for a discussion of this application for reliability-critical parts.]

Input from Outside Inspection and Test Functions. The "prior knowledge" does not automatically come to the inspection planner or the inspector. Some of this knowledge is already in existence as a byproduct of other activities and hence can be had for the procedural cost of retrieval. Other knowledge is not in existence and must be created by additional effort. However, this added effort is usually a one-time study, whereas the benefits then go on and on.

THE DEGREE OF INSPECTION AND TESTING NEEDED

It is evident that a determination of "How much inspection?" should be made only after there has been an evaluation of the other inputs to product knowledge. This evaluation can then dictate any of several levels of inspection.

No Inspection. There is already adequate evidence that the product or service conforms, and hence no further inspection is needed.

Skip Lot. There is already adequate evidence that the product or service conforms, but because of the nature of the characteristics being checked or the customer's requirement for some type of verification, a need to spot check the lots in batches every so often is conducted. This skip lot will remain effective until a nonconformance occurs; then the inspection or test reverts back to lot sampling.

Sampling Plans. Where there is little or no prior knowledge and no product fluidity, the main source of product knowledge becomes product inspection through *random sampling.* The amount of this inspection can be determined "scientifically" once the tolerable level of defects in accepted product has been defined clearly. However, choice of these levels—using the sampling parameters AQL (acceptable quality level), AOQL (average outgoing quality level), etc.—is largely arbitrary

and usually is determined by negotiation. In theory, the sampling parameters can be determined from economic considerations, i.e., the cost of detecting unsatisfactory lots versus the cost of failing to detect them. In practice, the "cost of detecting" is fairly easy to determine, but the cost of "failing to detect" is difficult to determine. For intangibles such as customer goodwill, there is no way known to make the determination with any useful precision.

One Hundred Percent Inspection and Test. This alternative is usually used for final test of critical or complex products. In very critical cases, it is used to provide redundancy against the unreliability of 100 percent inspection. In these cases the amount may be 200 percent or over. In cases where "zero defects" constitute the objective or requirement, 100 percent automated inspection is required. [See Nygaard (1981) and the discussion on machine-vision systems later in this section.] In some cases, 100 percent inspection is required to satisfy legal or political requirements (Walsh 1974). In other cases, 100 percent inspection is the most cost-effective approach. [See Walsh et al. (1976) and (1978) for examples where 100 percent inspection is the cost-effective alternative in high-volume production and in testing hardness of finished steel wheels.]

One hundred percent inspection also may be used when process capability is inherently too poor to meet product specifications. Sampling is of no avail in such cases, since the accepted lots are usually no better than the rejected lots, i.e., the difference is merely the result of statistical variations in the respective samples. This does not apply in cases where the process is highly erratic so that some lots are truly conforming and others are not merely the result of statistical accidents. For such processes, sampling can be a useful way to separate the conforming lots from the nonconforming ones.

Whatever the motivation for selecting the 100 percent alternative, effective implementation goes beyond choosing the inspection methodology or test equipment. Physical arrangements and personnel procedures must be changed. In many cases, attitudes of in-house and vendor personnel must be changed through reeducation. [See Walsh et al. (1979*b*).]

With the advent of computer-based testing, 100 percent inspection is becoming more practical (Schweber 1982). However, it is not clear that it is cost-effective in all cases, even if the technology is readily available (Walsh et al. 1979*a*). This is especially true if the total costs of quality are considered (Gunter 1983).

Check List Inspection. Where there is not a hardware sort of product that requires an assessment of the dimension, appearance, or testing parameters, the main source of evaluation is the check list type. Such service industries as hotels, restaurants, banks, and health care require only evidence of compliance or no compliance to a specific standard.

OTHER TYPES OF CONFORMANCE INSPECTIONS

Faced with the objective to minimize inspection costs and achieve maximum quality control, management must be sensitive to nontraditional types of conformance decision-making inspections.

Simulation. With advancement of the digital computer, system-simulation techniques are used as an alternative to the experimental or analytical approach (Wang et al. 1981). Examples of application:

Receiving inspection simulation

Printed circuit board assembly inspection simulation

Camera subassembly and final inspection simulation

Automated Inspection and Test. The first large-scale applications of automated testing were very likely done by the Western Electric Company during the 1920s. Current developments in

microcomputers, artificial intelligence (AI), integrated computer-aided manufacturing (ICAM), robotics, and software are making automated inspection practical and cost-effective. [See "An Outlook for 3D Vision" (1984) and "ASD's Quality Assurance Program Rates in Top 10" (1984); also see the discussion on 100 percent inspection above and that on Machine Vision, in Human Factors in Inspection, below.]

Automated inspection and testing are used to reduce costs, improve precision, shorten time intervals, alleviate labor shortages, and avoid inspection monotony, among other advantages. In some industries, the labor problems now seem insoluble in the absence of automated inspection. Already in widespread use, automated inspection is still expanding, with no end in sight.

The economics of automation involve a substantial investment in special equipment to secure a reduction in operating costs. The crux of justifying the investment lies in the amount of repetitive work the equipment will be called on to perform. This estimate of the anticipated volume of testing therefore should be checked out with great care.

A common starting point in discovering opportunities for automated inspection is to make a Pareto analysis of the kinds of inspections and tests being conducted. The vital few types are identified. Estimates are then made of the personnel, costs, and other current problems associated with these tests. The economics of automation are then estimated, and the comparable figures are an aid in deciding on the feasibility of successful conversion. For complex equipment involving depot storage and field maintenance, the question of use of automated testing is itself highly complex and requires a tailor-made study of some magnitude.

Technologically, the "machine" poses many problems. It is less adaptable than the human being it replaces, so some changes may need to be made in the product to offset this rigidity. For example, the machine may hold the units to be tested by grasping certain surfaces whose dimensions were previously unimportant. Now these surfaces may need to be held to close dimensions because the machine is not as adaptable as the human inspector. Alternatively, the product design may need to be changed to provide for adequate location.

Beyond the work of original design, construction, and prove-in, the machine must be set up specially for each job. However, modular construction master test pieces and taped programs have considerably reduced setup time while improving reproducibility. Reliability generally has been high, and use of printed circuit cards and other modular components has so reduced the "mean time to repair" that downtime is generally below 5 percent for well-designed machines.

Automated gaging and testing are used extensively in the mechanical industries. They are also widely used in the electronics industries, especially for electronic components, where the problem of making connections to the automated test equipment is so severe that the original product design must provide especially for this.

In the chemical industries, the corresponding development has been the *autoanalyzer.* This already has made possible some extensive cost reductions and solutions of otherwise forbidding problems of recruitment of laboratory technicians. The autoanalyzer makes use of some equipment common to all tests—sensors, transducers, recorders, and computers. However, each type of analysis has its unique procedure for converting the material under test into a form suitable for sensing.

The rapid evolution from numerically controlled (NC) machines to computer numerical control (CNC) has brought with it related new technologies. Machines now available are capable of continuously informing the operator of the machine's state of adjustment. Specification data stored in the computer's database are compared with the continuous stream of workpiece measurements provided by the measurement sensors to warn the operator when an adjustment is needed and even, in some cases, make the adjustment automatically. This new control technology, with its programmable logic, simplifies the operator's task. In reducing the degree of nonconformance at the source and providing a richer stream of information for the inspector, it also holds the promise of reducing the volume of inspections necessary and making the inspector function much more reliable. Applying programmable logic to the control of tool cutter wear compensation makes possible longer tool life and products of higher grade and greater consistency.

These new CNC machines also can function in a "lights out" manufacturing setting, in which a single machine or machining cell can operate for long periods and many machine cycles without human intervention, like an airplane flying on automatic pilot. Technology now under development

will incorporate setup approval as an integral part of the setup process and then retain in memory a record of the setup as a mode of objective evidence documentation. Only when the setup is approved does the software release the machine for production; if the setup is rejected, the software does not respond to a "run" command. This technology will allow manufacturing organizations to create a totally paperless system and still meet the requirements of ISO 9000 standards.

As we move into the next century, we will see machines that will be programmed to react to voice recognition with output able to make tool adjustments automatically, without operator intervention. All these technologies will contribute to increased consistency and productivity; equally important, they will help to improve product quality as well.

There are many recent and successful stories of how automated gaging and testing have improved defect detection and product reliability and even reduced the overall cost per unit of inspection and reject rate. Usually custom-built and highly engineered, these automated systems are best suited for application where certain conditions exist:

- Steady production output
- History of process control
- History of low reject rate
- Where products need individual appraisal

Automated inspection and test may be categorized into five main types: postprocess gaging, in-process gaging, testing, inspection, and assembly and test systems (Quinlan 1996). In all five types of inspection and testing, the automation or semiautomated concept provides high-confidence repeatability and a faster analytical mode that can be interfaced directly to a machine or process to display visually or audibly the statistical process variance above or below the control limits and signal the immediate need for process corrective action.

Strong advancements also have been made in high-speed visual inspection processes, where automated electronic inspection stations are either integrated or slightly off line from manufacturing operations. High-speed cameras capture the image displayed on a monitor, and in microseconds each part passing under the "eye" or segment of a continuous process automatically can undergo multicharacteristic checks. At the end of the inspection cycle, a computer monitor tells the operator if the part or strip passed. The monitor also can indicate why parts are rejected, while updating quality process statistics.

Prior inspection methods could spot only 85 to 90 percent of critical defects. With the new high speed visual automation, virtually all defects are caught. This advanced technology has resulted in increased productivity, reduced waste, and fewer defective parts passing through to the customer (Lincoln 1996).

Integrated Process Inspection and Test. More and more companies are developing inspection and/or testing equipment that can be integrated into the process directly. Instead of having to physically transport the part to the inspection station, certain characteristics or testing parameters can be evaluated directly on machine or process equipment. For example, one company integrated a measurement system comprising a standard gage linked to a statistical analyzer and coupled with some specially made fixturing. This method allowed the company to make an appraisal decision without having to spend abundant amounts of time on transporting the product to the inspection station and inspection time for setup. This cost savings can be realized in productivity gains by closing the gap between manufacturing and inspection. The shorter the time required for appraisal, the more time is available for production and the greater is the impact of generating larger amounts of quality product.

Specialty gaging and control monitoring equipment has been developed to integrate directly into certain types of machinery and sold as an accessory or an attachment to an existing machine. See, for example, Figure 23.1, which shows a process-control monitor integrated into a coldforming machine for continuous monitoring of process variation. If a force load value falls outside specified parameters, the instrument signals a shutdown of the machine, and only after a quality evaluation is

conducted and conformance is determined can the monitor be reactivated by the insertion of a key—usually under quality assurance ownership.

Computer-Aided Inspection. The marriage of computer and inspector provides benefits to inspection in the form of (1) information or (2) assistance to enhance the conformance decision for almost any inspection situation, or both. The marriage of computers and inspectors takes the form of (1) providing information and/or (2) providing assistance to enhance the conformance decision for almost any inspection situation. This is especially applicable to inspecting precision machine parts and assembly inspection [see Holmes (1974) and Linn (1981) for further discussion]. The selective use of computer-aided inspection (CAI) techniques can minimize the more menial, repetitive inspection tasks and direct the human resource (the inspector) to preventive quality control.

Voice Entry. Another unique enhancement of computer-aided technology is the voice data-entry system (VDES). VDES can be applied to incoming, in-process, and final inspection functions, as

FIGURE 23.1 Process control monitor.

well as some testing operations. Yet, somewhat embryonic in its applications, VDES will continue to develop in providing the inspection and testing desciplines with many practical benefits, such as

- Saving time by eliminating manual entry
- Reducing labor requirements
- Providing instantaneous response
- Accepting multiple languages and all speakers
- Allowing for remote control
- Permitting on-line, real-time control

Video Entry. Video inspection techniques used to augment typical video applications called *Moiré contouring* are emerging as a viable method for inspection. Using an optical head with a projection system and camera, a monitor, and a computer equipped with an image-processing board, Moiré allows the operator to obtain a very dense collection of *X, Y, Z* measurements. This depth-information technology is used to find surface defects at various dimension and contours of the part.

The Moiré sensor has been used in several companies for visual enhancement of warping, to inspect hard-to-measure locations, and to examine other machining features with no sharp edges or lighting contrasts to define such shapes as bevels, tapers, etc.

Because Moiré fringe data are already separated into two dimensions (2D) and the depth information is encoded into a 2D line drawing, computer analysis is easier. Also, inspection applications let the manufacturer use the high-density *X, Y, Z* data to detect random part flaws and warpage, and the process can move from high-speed off-line inspection to high-speed on-line inspection and in-process control (see Kennedy 1996).

Optical Sensing. Optical sensing also can be designed for some applications that may require greater sensing capabilities than would otherwise be possible. This optical scanning technology can measure dimensions by counting the number of beams blocked by an object passing through the scanned region.

Optical scanning systems typically consist of three components: a transmitter, a receiver, and a controller. The transmitter includes a series of light-emitting diodes (LEDs) in a linear array. The controller sequentially switches each LED to the "on" state in succession to generate a sequence of parallel beams of infrared light.

The receiver, with a corresponding linear arrangement of phototransistors, is modulated by the same controller (multiplexed) so that each phototransistor only detects light from its corresponding LED in the transmitter. When something changes the state of one of the receivers, the controller can generate an analog, parallel digital, or serial output (Strack 1996).

INSPECTION AND TESTING FUNCTIONS

Inspection functions are usually staffed by full-time inspectors or testers responsible to the Inspection Department. This is by no means universal. Some final inspection or test functions are staffed by full-time inspectors responsible to production. Many process inspection stations are staffed by production workers whose principal job is production.

Receiving (Incoming) Inspection and Testing. The extent of inspection of products received from suppliers depends largely on the extent of prior planning for supplier quality control. In the extreme case of using surveillance and audit of decisions, there is virtually no incoming inspection except for identity. At the other extreme, many "conventional" products are bought under an arrangement that relies primarily on incoming inspection for control of supplier quality.

The inspectors and their facilities are housed in the receiving area to provide ready collaboration with other supplier-related activities, i.e., materials receiving, weighing, counting, and storage. Depending on the physical bulk and tonnage of product, entire shipments or just samples are brought to the inspection floor. The documentation routines provide the inspectors with copies of the purchase orders and specifications, which are filed by supplier name.

Inspection planning is conventional, as discussed under Inspection Data Planning above. However, there is usually a lack of prior knowledge of process capability, order of manufacture, etc. Consequently, the sampling plans involve random and (often) large samples, employing standard random inspection tables. Randomness becomes a severe problem in the case of large shipments, whether bulk or not. However, special arrangements can be made with the supplier. Setting acceptable quality levels (AQLs) has been a troublesome problem to such an extent that some industry standards have been worked up. In the absence of such standards, the AQLs are established based on precedent, past performance, or just arbitrarily. Then, as instances of rejection arise, the negotiations with vendors result in adjustment to the AQLs or other acceptable sampling criteria. Data feedback to vendors follows conventional feedback practice.

Process Inspection and Testing. This commonly serves two purposes simultaneously:

1. *Provides data for making decisions on the product*; e.g., does the product conform to specification?
2. *Provides data for making decisions on the process;* e.g., should the process run or stop?

Because of the interrelation between process and product variables, process inspection involves observation of process variables as well as inspection of the product. These observations and inspections are made by both production and inspection personnel.

Product acceptance of work in process may be done in any of several stages or by a combination of them. These stages include the following.

Setup Inspection. Some processes are inherently so stable that if the setup is correct, the entire lot will be correct, within certain limits of lot size. For such processes, the setup approval also can be used as the lot approval. Where a good deal is at stake, it is usual to formalize the setup inspection and to require that the process may not run until the inspector has formally approved the setup, e.g., by signing off, by stamping the first pieces, etc. (Garfinkel and Clodfelter 1984).

Patrol Inspection. For processes that will not remain stable for the duration of the lot, it is usual to provide for periodic sampling to be conducted during the progression of the lot, making use of various techniques described in Section 24. The numerous plans in use consist mainly of variations of the following four types:

1. Preserve the order of manufacture under an arrangement such as is depicted in Figure 23.2. In this example, the machine discharges its production into a small container called a *pan.* The production operator periodically empties the pan into one of three larger containers:
   a. Into the junk box if the parts are junk.
   b. Into the reject box if the parts are questionable or are mixed good and bad.
   c. Into the tray if the parts are presumably good.
   The patrol inspector comes to the machine and checks the last few pieces being made. (He or she also may sample the tray.) Based on this check, the tray is disposed of in one of three ways:
   a. Into the junk box if the parts are junk.
   b. Into the reject box if the parts are questionable or are mixed good or bad.
   c. Into the good box if the parts are acceptable. The good box goes on to the next operation.
   Only the inspector may dispose of the tray, and only the inspector may place any product in the good box. The reject box is gone over by a sorter, who makes three dispositions:
   a. Junk to the Junk Department.
   b. Reoperate back to the Production Department.
   c. Good parts on to the next operation.

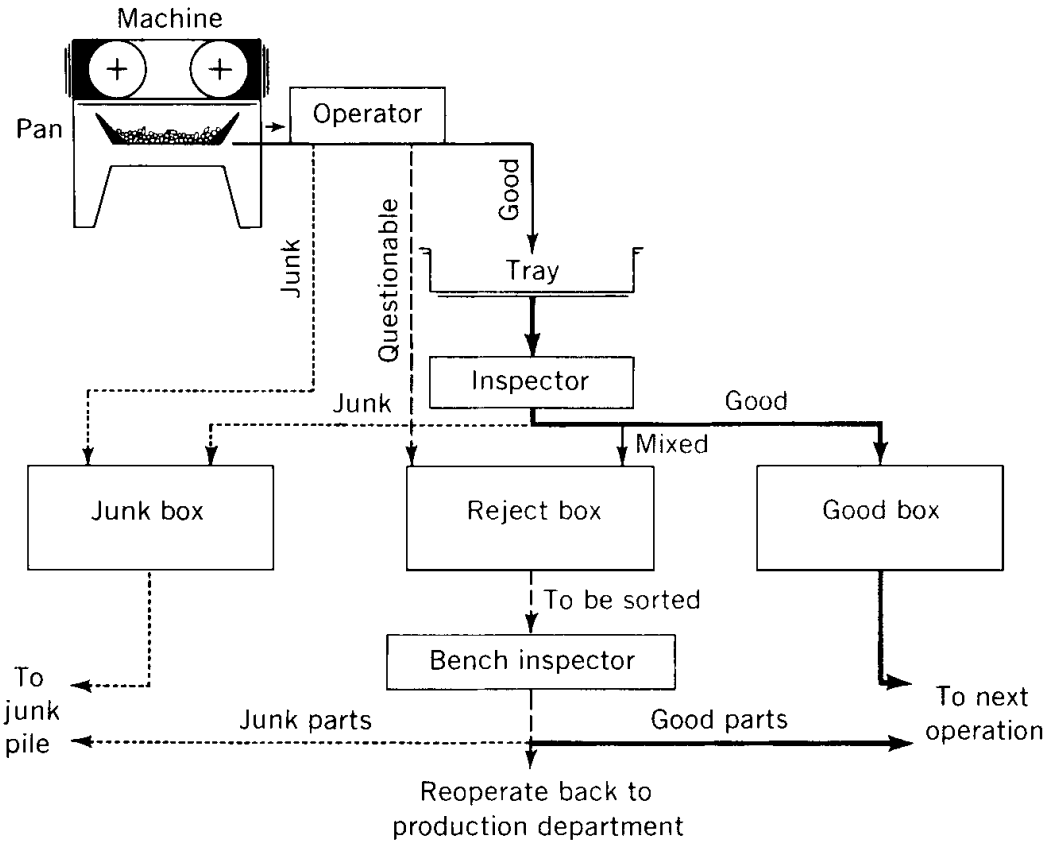

FIGURE 23.2 Patrol inspection plan based on preserving the order.

2. This method is similar to the first, but the inspection data from the last few pieces are posted to a control chart. If the process remains in control, all product made since the last check is accepted.
3. The accumulated product (e.g., in the tray of Figure 23.2) is sampled at random using some standard sampling plan, and acceptance is based on the sampling criteria.
4. The process variables are checked against a process specification, and the product is accepted provided the process conforms to specification. This method is usually restricted to cases in which there is to be a direct check on the product at later stages.

Tollgate Inspection. This is a lot-by-lot product-acceptance procedure. Commonly it is done after the Production Department has concluded its operations. Sometimes the product is moved physically to an inspection area, where it waits its turn to be inspected. Sampling is at random, using standard sampling tables.

Tollgate inspection reduces congestion at the machines and clarifies departmental responsibility. The price paid is in added material handling, added floor space, loss of order of production, and greater difficulty in fixing individual responsibility.

Finished-Goods Inspection. Most finished products are tested 100 percent for minimal simulation of use. Tests are often automated, as are the data recordings. Testing may be done either at inspection stations on the production line or in separate inspection areas.

Shipping Inspection. For processes that are continuous and stable and/or do not require multiple operations before enclosure into a container, a shipping inspection and/or test will provide assurance of conformance before the containers are sealed.

Dock Audit. This is an inspection of a random selection of product that is ready to be shipped. Such audits may be made daily, weekly, semiweekly, biweekly, etc. Whatever the frequency, it is important to check not only the product characteristics but also the labeling, shipping and handling methods, and many other requirement criteria.

Destructive Test. A product that has been deemed stable and by its functional nature requires a lot or batch destruction test on its components or assemblies is subject to a random destructive evaluation. If any failure occurs, the lot or batch is automatically quarantined to future analysis.

QUALITY STANDARDS

The cornerstone of quality control is the specification. Specifications embody the minimum and maximum values (see also Bader 1980). However, they are usually *incomplete.* They tend to ignore *visual quality characteristics*, or they treat characteristics labeled as *workmanship* superficially. [Refer to the discussion of visual quality characteristics below; and see also Dodds (1967) and Alaimo (1969) for further discussion on workmanship standards.]

Seriousness Classification. Some quality characteristics and defects are very important to fitness for use; others are not. The village artisan and the small-shop proprietor, with their first-hand knowledge of fitness for use, are able to concentrate their efforts on the most important qualities. In modern, large, complex organizations, the workers, inspectors, and many of the supervisors lack complete knowledge of fitness for use and thus are not fully clear on where to place their emphasis and how to make their decisions.

For example, one company studying the fabrication and inspection of machine parts divided the quality characteristics into four classes. Table 23.4 shows the effect of this classification on product tolerances and on the number of dimensions checked. The inspection time was reduced from 215 to 120 minutes. In addition, there were greater savings through lower rework costs, lower tooling costs, and lower engineering costs for disposition of nonconforming product (see also Allen 1959).

The seriousness classification takes into consideration information from a variety of sources:

The specification: This is the primary source. Because the specifications rarely reflect the most current input from customers and other sources, it is important that the team proactively seek current information from the sources listed here.

TABLE 23.4 Results of Seriousness Classification of Characteristics

			Number of dimensions checked	
Characteristic classification	Effect of classification on design tolerance	Effect of classification on amount of inspection	Before classification	After classification
Critical	None	None	154	154
Major	None	None	110	110
Minor A	Tolerance was increased by a specified amount (doubled, etc.) provided the part assembled satisfactorily	Inspection was made normally, but to wider tolerances	66	15
Minor B	Tolerance was ignored provided the part assembled satisfactorily	Inspection was eliminated	352	0
Total			682	279

Customer input: The classification is an even more useful guide to decision making when the team broadens the information base to include direct preproduction input from the whole range of customers—the purchaser, the end user, the various internal departments involved in manufacture, inspection, packaging, shipping, and so on. Then the classification can reflect more accurately the customer's needs and expectations of the product.

Manufacturing experience with the product: Problems and inadequacies, including evidence of nonconformance during manufacturing, help identify troublesome features.

Life testing and functional testing: When these functions are part of the production cycle, they should be included in the classification process and in structuring a formal classification model. The test results are useful input to classification decisions. Where there exists the possibility of product misuse—which, for a variety of unanticipated reasons, too frequently replaces the customer's originally intended use—the tests can help anticipate and prevent such misuse or at least help mitigate its effects. The proper classification ensures that the affected features will receive proper attention.

Failure during use: Field failure is the ultimate undesirable quality outcome. Any feature for which a causal link can be established to field failure is, by definition, a critical feature.

Seriousness classification is useful input both to the control plan and the overall quality plan. Some quality characteristics are multileveled in their seriousness criteria. For example, a shaft diameter specified as "1.000 ± 0.001" gives rise to two defects: oversize and undersize. These defects may be assigned different degrees of seriousness depending on the extent and effect of nonconformance. Some extensive defect lists, e.g., the list for glass bottles, have little resemblance to the list of characteristics set out in the specifications.

Some companies use the same system of classification for both characteristics and defects. However, there is enough uniqueness about each of the two lists to suggest that adoption of a single system should be preceded by a positive examination of the nature of the two lists. For example, the effect of seriousness classification on design decisions can be quite different from the effect on inspection decisions, as is evident from Table 23.4.

Formal systems of seriousness classification were evolved originally to serve specialized purposes. (The Bell System pioneered by developing a system to permit rating of quality of finished product. The U.S. Army developed systems to simplify the administration of acceptance of goods purchased from contractors.) However, as the systems came into being, they were found to have application to the entire progression of product from design through use: in quality specification, manufacturing planning, supplier relations, tooling, production, salvage, product auditing, and executing reporting. Vital qualities could now be identified with greater confidence, and it also became feasible to delegate class decisions and actions on a broad scale. For example, all class C defects could be assigned a common sampling plan, thereby avoiding the need for publishing numerous individual plans.

The multiple uses of seriousness classification systems make it desirable that the job of developing such a system be guided by an interdepartmental committee that has the responsibility for drafting a plan, modifying it, and recommending it for adoption. Such a committee has a series of tasks:

1. Determining the number of strata or classes of seriousness to use.
2. Defining each class.
3. Classifying each defect into one of the classes.

Number of Levels or Strata. In theory, this number may be large; e.g., a defect may have any weight from 1000 down to 1. In practice, such a large number of weights is too complex to administer. The actual plans in use consist of only several classes. While choice of the actual number of classes is arbitrary, extensive experience has shown that three or four classes suffice for a wide variety of situations.

Definitions for the Classes. These will differ with the nature of the product, process, etc. However, plans in existence tend to show striking similarity in definition, the result in part of the

influence of the Bell System classification plan [see also Dodge (1928) and Dodge and Toreey (1956)]. Not only was this pioneering plan uncommonly well reasoned out; the men who devised it were later consultants to some of the U.S. armed services during World War II, and their thinking influenced the classification plans adopted by these services. These plans, in turn, influenced the plans adopted by the contractors to the armed services.

The standard definitions adopted by the Bell System [see also Dodge and Torrey (1956)] are shown in Table 23.5. Study of these definitions discloses that there is an inner pattern common to the basic definitions (Table 23.6).

A composite of definitions used in food industry companies is shown in Table 23.7. It is evident that there are industry-to-industry differences in products, markets, and so on that require a tailor-made wording for each industry. In addition, the lists are not static. The growth of government regulation has further influenced the definition, as has the problem of repairs and guarantees for long-life products.

It is also evident that the classifications must simultaneously take into account multiple considerations such as functional performance, user awareness, and financial loss. For example, the effects and awareness of a radio receiver's defects may be as follows:

Defect	Effect	User awareness
Open circuit in power supply	Set is inoperative	Fully aware
Short circuit in resistor	Excess power consumption	Seldom aware
Poor exterior finish	No effect	Usually aware
Poor dress or internal wiring	No effect	Seldom aware

Classifying the Defects. This essential task is time-consuming, since there are always many defects to be classified. If the class definitions have been well drawn, the task becomes much easier.

During classifying, much confusion is cleared up. It is found that the seriousness of important visual defects depends not so much on whether the inspector can see them as on whether the consumer can see them. It is found that some words describing defects must be subdivided; i.e., a *stain* may be placed in two or three classes depending on severity and location. In many ways, the work of classifying defects is rewarding through clearing away misconceptions and giving a fresh view to all who participate.

Classification of Characteristics. In some companies the formal "seriousness" classification is not of defects but of characteristics in the specifications. The classification may be in any of several alternatives:

1. *Functional or nonfunctional.* Where a single set of drawings carries both functional ("end use") requirements and nonfunctional ("means to an end") requirements, it is important to make clear which is which. (This is not to be confused with mechanical, chemical, or electrical functioning. In products such as jewelry or textiles, the most important functional requirement is *appearance.*) The purposes served by these two classes are generally alike throughout industry:

Functional requirements are intended to

- Ensure performance for intended use
- Ensure long, useful life
- Minimize accident hazards
- Protect life or property
- Provide interchangeability in the field
- Provide competitive sales advantage

Nonfunctional requirements are intended to

- Inform the shop as to method of manufacture
- Reduce cost of manufacture
- Facilitate manufacture
- Provide interchangeability in the shop
- Provide information to toolmakers

TABLE 23.5 Serious Classification of Defects (Bell System)

Class A: Very serious (demerit value, 100)
1. Will surely cause an operating failure of the unit in service that cannot be readily corrected in the field, e.g., open relay winding, or 2. Will surely cause intermittent trouble, difficult to locate in the field, e.g., loose connection, or 3. Will render unit totally unfit for service, e.g., dial finger wheel does not return to normal after operation, or 4. Liable to cause personal injury or property damage under normal conditions of use, e.g., exposed part has sharp edges.
Class B: Serious (demerit value, 50)
1. Will probably cause an operating failure of the unit in service that cannot be readily corrected in the field, e.g., protective finish missing from coaxial plug, or 2. Will surely cause an operating failure of the unit in service that can be readily corrected in the field, e.g., relay contact does not make, or 3. Will surely cause trouble of a nature less serious than an operating failure, such as substandard performance, e.g., protector block does not operate at specified voltage, or 4. Will surely involve increased maintenance or decreased life, e.g., single contact disk missing, or 5. Will cause a major increase in installation effort by the customer, e.g., mounting holes in wrong location, or 6. Defects of appearance or finish that are extreme in intensity, e.g., finish does not match finish on other parts, requires refinishing.
Class C: Moderately serious (demerit value, 10)
1. May possibly cause an operating failure of the unit in service, e.g., contact follow less than minimum, or 2. Likely to cause trouble of a nature less than an operating failure, such as substandard performance, e.g., ringer does operate within specified limits, or 3. Likely to involve increased maintenance or decreased life, e.g., dirty contact, or 4. Will cause a minor increase in installation effort by the customer, e.g., mounting bracket distorted, or 5. Major defects of appearance, finish, or workmanship, e.g., finish conspicuously scratched, designation omitted or illegible.
Class D: Not serious (demerit value, 1)
1. Will not affect operation, maintenance, or life of the unit in service (including minor deviations from engineering requirements), e.g., sleeving too short, or 2. Minor defects of appearance, finish, or workmanship, e.g., slightly scratched finish.

TABLE 23.6 Inner Pattern: Seriousness Classification System

Defect class	Demerit weight	Cause personal injury	Cause operating failure	Cause intermittent operating trouble difficult to locate in field	Cause substandard performance	Involve increased maintenance or decreased life	Cause increase in installation effort by customer	Appearance, finish, or workmanship defects
A	100	Liable to	Will surely*	Will surely				
B	50		Will surely		Will surely	Will surely	Major increase	
C	10		May possibly		Likely to	Likely to	Minor increase	Major
D	1		Will not		Will not	Will not	Minor increase	Minor

*Not readily corrected in the field.

TABLE 23.7 Composite Definitions for Seriousness Classification in Food Industry

Defect	Effect on consumer safety	Effect on use	Consumer relations	Loss to company	Effect on conformance to government relations
Critical	Will surely cause personal injury or illness	Will render the product totally unfit for use	Will offend consumer's sensibilities because of odor, appearance, etc.	Will lose customers and will result in losses greater than value of product	Fails to conform to regulations for purity, toxicity, identification
Major A	Very unlikely to cause personal injury or illness	May render the product unfit for use and may cause rejection by the user	Will likely be noticed by consumer, and will likely reduce product salability	May lose customers and may result in losses greater than the value of the product; will substantially reduce production yields	Fails to conform to regulations on weight, volume, or batch control
Major B	Will not cause injury or illness	Will make the product more difficult to use, e.g., removal from package, or will require improvision by the user; affects appearance, neatness	May be noticed by some consumers, and may be an annoyance if noticed	Unlikely to lose customers; may require product replacement; may result in loss equal to product value	Minor nonconformance to regulations on weight, volume, or batch control, e.g., completeness of documentation
Minor	Will not cause injury or illness	Will not affect usability of the product, may affect appearance, neatness	Unlikely to be noticed by consumers, and of little concern if noticed	Unlikely to result in loss	Conforms full to regulations

"Which is which" becomes important because it directs the priorities of process design and many aspects of economics of manufacture, as well as the jurisdiction over waivers. When the engineers make this classification, they commonly add a designation such as *E* (for engineering) to the functional characteristics. All others are then assumed to be nonfunctional.

A comparable situation prevails in process specifications, where the need is to distinguish mandatory from advisory requirements, which correspond roughly to functional and nonfunctional requirements as applied to the product. However, the process specifications seldom make this distinction. (For further discussion, see Section 22, Operations, under Knowledge of "Supposed to Do.")

2. *Seriousness classification.* When this method is used, it parallels closely the classification into critical, major, and minor as used for classification of defects. (The contention is often raised that the tolerances on the specifications are an automatic form of seriousness classification, i.e., anything with assigned tolerances must be met and is "therefore" critical. An alternative contention is that the closeness of the tolerances is a key to seriousness classification; i.e., the narrowest tolerances are assigned to the most critical characteristics. When these contentions are examined more closely, they are found to contain too many exceptions to serve as firm rules for classifications.)

The resulting classifications then become a supplement to the specification or are shown on the drawings themselves by some code designation, for example:

Critical ⊕
Major A ⦶
Major B ○
Minor Not marked

One large automotive company differentiates between regulatory and nonregulatory critical characteristics by the use of two different symbols.

3. *Segregation of functional requirements.* Place in a separate document, such as an "engineering specification" or a "test specification."

4. *Shop practice tolerances versus special tolerances.* This method is based on preparation of a shop practice manual that sets out general-use tolerances derived from the process capability of general-use machines and tools. Once published, these shop practice tolerances govern all characteristics not specially toleranced.

Who Classifies? For defect classification, an interdepartmental committee is the ideal choice. This provides each department with the benefits derived from the process of active review, and it also produces a better final result (sometimes the committee goes further and establishes a plan for product rating, including demerit weights for each class; see Section 8, under Appraisal Costs). However, some companies assign a staff specialist to prepare a proposed classification, which is then reviewed by all interested departments. The specialist is usually a quality control engineer.

When the classification is limited to specified characteristics, e.g., functional versus nonfunctional, the designer usually prepares the draft.

SENSORY QUALITIES

Sensory qualities are those for which we lack technological measuring instruments and for which the senses of human beings must be used as measuring instruments. (For some special purposes, e.g., tests of toxicity, the test panel may consist of animals.) Sensory qualities may involve

- Technological performance of the product, e.g., adhesion of a protective coating, friction of a sliding fit
- Esthetic characteristics of consumer products, e.g., taste of food, odor of perfume, appearance of carpets, noise of room air conditioners

In common with other qualities, sensory qualities require:

1. Discovery of which characteristics are required and in what degree to meet the needs of fitness for use
2. Design of products that will possess these characteristics
3. Establishment of product and process standards and of tests that will simulate fitness for use
4. Judgment of conformance to the product and process standards

This multiplicity of tasks requires a corresponding multiplicity in type of sensory test panel used, choice of test design, and so on.

Customer Sensitivity Testing. In this form of test, the purpose is to discover the "threshold" level at which customers can detect the presence of sensory qualities. The qualities under test may be "desirable." For example, if an expensive ingredient is used in a product blend, it is very useful to know the threshold concentration level that ensures customer recognition of the ingredient. The qualities under test may be "undesirable." For example, a product exhibits varying degrees of visual blemish. It is very useful to know the threshold degree of defectiveness that makes the customer respond negatively to the product.

In customer-sensitivity testing, a graduated set of samples is prepared, each exhibiting a progressively greater extent of the quality or deficiency under investigation. These samples are submitted to a customer panel as part of an organized study.

For example, in two companies—one making sterling silverware and the other making costume jewelry—studies were conducted to discover customer sensitivity to visual defects. In both companies, a committee of key people (from Marketing, Design, Manufacture, and Quality) structured a plan of study as follows:

1. An assortment of product was chosen to reflect the principal visual defects, the principal products in the product line, and the principal price levels.
2. These samples were inspected in the factory by the regular inspectors to determine the severity of the defects as judged by the frequency with which the inspectors rejected the various units of product.
3. The assortment of products was then shown to a number of customer panels chosen from those segments of the buying public which constituted important customer classes, e.g., suburban women, college students, etc. These panels reviewed the products under conditions that simulated use of the product, e.g., silverware in place settings on a dining room table. The customers were instructed (by printed card) somewhat as follows: "Assume you have previously bought these products and they have been delivered to you. Naturally, you will want to look them over to see that the merchandise is satisfactory. Will you be good enough to look it over, and if you see anything that is objectionable to you, will you please point it out to us?"

The resulting data showed that the customer panels were highly sensitive to some defects. For such defects, the strict visual standards were retained. For certain other defects, the customer sensitivity was far less than factory-inspector sensitivity. For such defects, the standards were relaxed. In still other instances, some operations had deliberately been omitted, but the customers proved to be insensitive to the effect of the omission. As a result, the operations were abolished.

Customer sensitivity testing is an extension of the principle that "the customer is right." This principle may be subdivided as follows:

1. The customer is right as to qualities he or she can sense. As to such qualities, the manufacturer is justified in taking action to make such qualities acceptable to the customer.
2. The customer is also right as to qualities he or she *cannot* sense. The manufacturer is not justified in adding costs to create an esthetic effect not sensed by the customer.
3. Where, for a given quality, the customer is sensitive to a limited level but not beyond, the manufacturer should take action to make the quality to that level but not beyond.

The intermediate marketing chain sometimes interferes with these principles. Sales clerks are proficient in emphasizing product differences, whether important or not. In turn, dealers are alert to seize on such differences to wring concessions out of competing manufacturers. A frequent result is that all manufacturers are driven to adopt wasteful standards, resulting in a needlessly high cost, e.g., finishes on nonworking or nonvisible surfaces. Elimination of such perfectionism commonly requires that the manufacturer secure data directly from consumers and then use the data to convince the distribution chain. These same data may be needed to convince other nonconsumers who exhibit perfectionist tendencies: upper management, designers, salespeople, inspectors, etc.

Visual Quality Characteristics. These constitute a special category of sensory qualities. (Visual inspection remains the largest single form of inspection activity.) For these characteristics, the written specifications seldom describe completely what is wanted, and often inspectors are left to make their own interpretation. In such cases, inspectors are really making two judgments simultaneously:

1. What is the meaning of this visual characteristic of the specification, e.g., what is the standard?
2. Does this unit of product conform to the standard?

Where inspectors understand fitness for use, they are qualified to make both these judgments. If a particular inspector lacks this knowledge, he or she is qualified to make only judgment 2, no mat-

ter how long on the job. Extensive experience has shown that inspectors who lack this knowledge differ widely when setting standards and, in addition, do not remain consistent. (As an example, from the consulting experience of J. M. Juran, in an optical company, study of the methods used by 18 different inspectors, engineers, and so on disclosed the existence of six methods of counting the number of "fringes of irregularity.") Several methods are available to planners to clarify the standard for visual characteristics.

Visual Inspection and Test Standards. The most elementary form of visual standard is the *limit sample*—a unit of product showing the worst condition acceptable. In using this standard, the inspector is aided in two ways:

1. The sample conveys a more precise meaning than does a written specification.
2. The inspection is now made by comparison, which is well known to give more consistent results than judgment in the absence of comparison.

A more elaborate form of visual standards involves preparation of an exhibit of samples of varying degrees of defects ranging from clearly defective to clearly acceptable. See, for an example involving solder connections, Leek (1975 and 1976), who describes how the companies Martin Marietta and Northrup supported series or "ranges" to provide manufacturing latitude in processing material while at the same time identifying minimum and maximum limits on visual attributes. This exhibit is used to secure the collective judgments of all who have a stake in the standard—consumers, supervisors, engineers, and inspectors. Based on these judgments, standards are agreed on, and limit samples are chosen. (It is also feasible to estimate, by sampling, what would be the yield of the process, and thereby the cost of defects, for any one of the various degrees of defectiveness.)

In products sold for esthetic appeal, appearance becomes a major element of fitness for use and commonly a major element of cost as well. In such cases, an exhibit of samples with varying degrees of defects intermingled with perfect units of product becomes a means of measuring consumer sensitivity (or insensitivity) to various defects. Use of consumer panels to judge such mixtures of product invariably confirms some previous concepts but also denies some long-standing beliefs held by managers as well as by the inspectors.

In the sterling silverware case mentioned earlier, consumers were quite sensitive to several types of defects—they held out 22 percent of the defects present. However, for the bulk of defects, the consumers were quite insensitive and found only 3 percent of such defects. The salespeople generally found twice as many defects as consumers but still considerably fewer than factory inspectors.

A further use of samples of various defects is to establish *grades* of defects. The concept of different grades is vital when a plant makes products that, while outwardly similar, are used in widely different applications, e.g., ball bearings used for precision instruments and those used for roller skates, lenses for precision apparatus and lenses for simple magnifiers, or sterling silverware and plated silverware. Unless the grades are well defined and spelled out in authoritative form, the risk is that the inspectors will apply one standard to all grades.

Once limit samples have been agreed to, there remains a problem of providing working standards to the inspection force. Sometimes it is feasible to select duplicates for inspection use while retaining the official standard sample in the laboratory. An alternative is to prepare photographs (sometimes stereoscopic) of the approved standards and to distribute these photographs instead.

Standardizing the Conditions of Inspection and Test. Visual inspection results are greatly influenced by the type, color, and intensity of illumination, by the angle of viewing, by the viewing distance, and so on. Standardizing these conditions is a long step in the direction of securing uniform inspection results. In the case of esthetic visual qualities, the guiding rule for conditions of inspection is to simulate the conditions of use, but with a factor of safety.

Establish a *fading distance.* In some products the variety of visual defects is so great and the range of severity so wide that the creation of visual standards becomes prohibitively complex. An

alternative approach is to standardize the conditions of inspection and then to establish a fading distance for each broad defect class. The definition for a defect becomes "anything that can be seen at the fading distance." (This technique appears to have been evolved in 1951 by N. O. Langenborg of St. Regis Paper Company. See also Riley 1979.)

Sensory Tests: Design and Analysis. There are numerous designs of sensory tests, some of them quite complex. Some of the basic forms are described below.

Tests for Differences or Similarities. These include

1. *The paired-comparison test.* Product is submitted to members of a panel in pairs of samples. One sample is identified to each panelist as the standard or "control"; the other is the test sample. The panelist is asked to judge and record the difference on a scale of differences (such as no difference, slight difference, or pronounced difference). Some of the pairs have no difference; i.e., both are "controls."
2. *The triangle test.* The panelist is asked to identify the odd sample in a group of three, two of which are alike. He or she also may be asked to estimate the degree of difference and to describe the difference between the odd sample and the two like samples.
3. *The duo-trio test.* The panelist is asked to identify which of two samples is like the "control" to which she or he has been subjected previously. For example, in liquor manufacture, the aim is to make each batch indistinguishable in taste from past batches. The duo-trio test is used as a product-acceptance test. Each panelist tastes the "control," which he or she is told comes from previous product. Then each panelist tastes the two remaining samples, one of which is "control" and the other of which is the batch under test. However, the panelists are not told which is which. If the data make clear that the panel cannot distinguish the new batch from the control, the batch is accepted. Otherwise, it is reblended.
4. *Ranking test.* Coded samples are submitted to each panelist, who is asked to rank them in the order of concentration.

Creating New Instruments to Measure Sensory Qualities. Many sensory qualities formerly judged by human perception are now measured by instruments. This development of new instrumentation goes on apace using essentially the following approach (based on a procedure set out by Dr. Amihud Kramer 1952; see also Hains 1978):

1. Define precisely what is meant by the quality characteristic under discussion. This must be done with participation of all interested parties.
2. Discover, through analytical study, the subcharacteristics, and define them in a way that permits, in theory, measurement by some inanimate instrument.
3. Search the literature to become informed about methods already in existence or under the development for measuring these subcharacteristics. This search will disclose a number of such possible measurement methods.
4. Choose or create product samples that vary widely for the subcharacteristics. Test a limited number (10 to 50) of samples with each of the various measurement methods, and correlate these tests with evaluation by panels of human testers. The human evaluation here aims not to measure personal preferences but to rate the degree to which the samples possess the variable under study. Hence the main requirement of the panel is that it be able to discriminate the subcharacteristics under study. Discard those measurement methods which lack precision or which fail to reflect human evaluation.
5. For the remaining, more promising measurement methods, conduct tests on a larger number of samples (100 to 1000) also chosen to reflect the entire range of quality variation. In addition, conduct tests of duplicate samples using evaluation by human test panels.

6. Correlate the results of measurement against the human test panel evaluation; select the method that gives a high correlation. (Multiple correlation methods may be necessary.)
7. Improve and simplify the selected measurement method through further tests and correlations.
8. Establish a scale of grades through use of a human sensory test panel. At this stage, the prime purpose of the human test panel is to state preferences along the scale of measure. Hence the main requirement of this panel is that it be representative of the producers and users of the product.
9. Weigh the various subcharacteristics in accordance with their rated performance [see also Montville (1983) and Papadopoulos (1983)].
10. Develop the sampling procedures needed to apply the resulting method of measurement.

MEASUREMENT QUALITY: AN INTRODUCTION

Conduct of the quality function depends heavily on qualification of product and process characteristics. This quantification is done through a systematic approach involving

1. Definition of standardized units called *units of measure* that permit conversion of abstractions (e.g., length, mass) into a form capable of being quantified (e.g., meter, kilogram).
2. *Instruments* that are calibrated in terms of these standardized units of measure.
3. Use of these instruments to quantify or *measure* the extent to which the product or process possesses the characteristic under study. This process of quantification is called *measurement.*

The word *measurement* has multiple meanings, these being principally

1. The *process* of quantification; e.g., "The measurement was done in the laboratory."
2. The resulting number; e.g., "The measurement fell within the tolerances."

Measurement rests on a highly organized, scientific base called *metrology,* i.e., the science of measurement. This science underlies the entire systematic approach through which we quantify quality characteristics.

MEASUREMENT STANDARDS

The seven fundamental units of the International System (SI) of measurement are defined as shown in Table 23.8. It is seen that except for the kilogram, all units are defined in terms of natural phenomena. (The kilogram is defined as the mass of a specific object.)

Primary Reference Standards. In all industrialized countries there exists a national bureau of standards whose functions include construction and maintenance of *primary reference standards.* These standards consist of copies of the international kilogram plus measuring systems that are responsive to the definitions of the fundamental units.

In addition, professional societies (e.g., the American Society for Testing and Materials) have evolved standardized test methods for measuring many hundreds of quality characteristics. These standard test methods describe the test conditions, equipment, procedure, and so on, to be followed. The various national bureaus of standards, as well as other laboratories, then develop primary reference standards that embody the units of measure corresponding to these standard test methods.

Primary reference standards have a distinct legal status, since commercial contracts usually require that "measuring and test equipment shall be calibrated . . . utilizing reference standards . . . whose calibration is certified as being traceable to the National Institute of Standards and

TABLE 23.8 Definitions of Fundamental Units of the SI System

Unit	Definition
Meter, m	1/650 763.73 wavelengths in vacuo of the unperturbed transition $2_{p10} \rightarrow 5d_5$ in ^{36}Kr
Kilogram, kg	Mass of the international kilogram at Sevres, France
Second, s	1/31 556 925 974 7 of the tropical year at 12^h ET, 0 January 1900, supplementarily defined in 1964 in terms of the cesium F, 4; M, 0 to F, 3; M, 0, transition, the frequency assigned being 9 192 631 770 Hz
Kelvin, K	Defined in the thermodynamic scale by assigning 273.16 K to the triple point of water (freezing point, 173.15 K = 0°C)
Ampere, A	The constant current which, if maintained in two straight parallel conductors of infinite length, of negligible circular sections, and placed 1 m apart in a vacuum, will produce between these conductors a force equal to 2×10^{-7} mks unit of force per meter of length.
Candela, cd	1/60 of the intensity of 1 cm^2 of a perfect radiator at the temperature of freezing platinum.
Mole, mol	An amount of substance whose weight in grams numerically equals the molecular formula weight.

Technologies." In practice, it is not feasible for the U.S. National Institute of Standards and Technologies to calibrate and certify the accuracy of the enormous volume of test equipment in use in shops and test laboratories. Instead, resort is made to a hierarchy of secondary standards and laboratories, along with a system of documented certification of accuracy.

Hierarchy of Standards. The primary reference standards are the apex of an entire hierarchy of reference standards (Figure 23.3). At the base of the hierarchy there stands the huge array of *test equipment,* i.e., instruments used by laboratory technicians, workers, and inspectors to control processes and products. These instruments are calibrated against *working standards* that are used solely to calibrate these laboratory and shop instruments. In turn, the working standards are related to the primary reference standards through one or more intermediate secondary reference standards or *transfer standards.* Each of these levels in the hierarchy serves to "transfer" accuracy of measurement to the next lower level in the hierarchy.

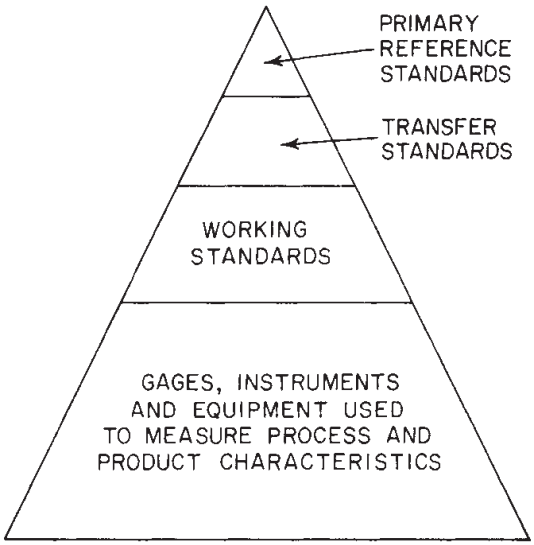

FIGURE 23.3 Hierarchy of standards.

Within the hierarchy of standards there are differences both in the physical construction of the standards and in their precision. The primary reference standards are used by a relatively few highly skilled metrologists, and their skills are a vital commitment to the high precision attained by these standards. As we progress down the hierarchy, the number of technicians increases with each level, until at the base there are millions of workers, inspectors, and technicians using test equipment to control product and process. Because of the wide variation in training, skills, and dedication among these millions, the design and construction of test equipment must feature ruggedness, stability, and foolproofing so as to minimize errors contributed by the human being using the equipment.

Precision of measurement differs widely among the various levels of the hierarchy of standards. At the level of primary reference standards, the precision is determined by the state of the art. For example, Figure 23.4 shows the precision attained by the U.S. National Bureau of Standards (now

the National Institute of Standards and Technology) when weighing loads across the spectrum of 10^{-10} to 10^{6} (National Bureau of Standards 1965).

At the base of the hierarchy, the precision of measurement is determined by the needs of fitness for use as reflected by the product and process tolerances. While some specialists urge that the test equipment be able to "divide the tolerance into tenths," this ideal is by no means always attained in practice. However, the tolerances themselves have been tightened drastically over the centuries, and this tightening generally has paralleled the advances made in the state of the art of measurement. For example, accuracy of measurement of a meter of length has progressed from an error of 1000 per million (at the end of the fifteenth century) to an error of 0.0001 per million as we move into the twenty-first century.

Allocation of measurement errors among the working and transfer standards has been discussed widely but has not been well standardized. The precision gap between primary reference standards and product test equipment may be anywhere from one to several orders of magnitude. This gap must then be allocated among the number of levels of standards and laboratories (transfer plus working) prevailing in any given situation. (Some models have been worked out to show the interrelation among the cost of developing greater precision in the primary reference standard, cost of attaining precision at each level of transfer laboratory, and number of laboratories at each level. See, for example, Crow 1966.) When this problem of allocation was first faced, there was a tendency to conclude that each level should have a precision 10 times greater than the level it was checking. More recently, there has been growing awareness of how multiple levels of precision combine; i.e., their composite is better represented by the square root of the sum of the squares rather than by the arithmetic sum. This new awareness has caused many practitioners to accept a ratio of 5:1 rather than 10:1 for precision of working standards to product tolerances. This same ratio also has been tolerated among transfer standards as well.

ERROR OF MEASUREMENT

Product and process conformance is determined by measurements made by the test equipment at the bottom of the hierarchy of standards. Obviously, any error in these measurements has a direct bearing on the ability to judge conformance. On examination, the nature of measurement error is quite complex; even the terminology is confused. A clear understanding of the meaning of the

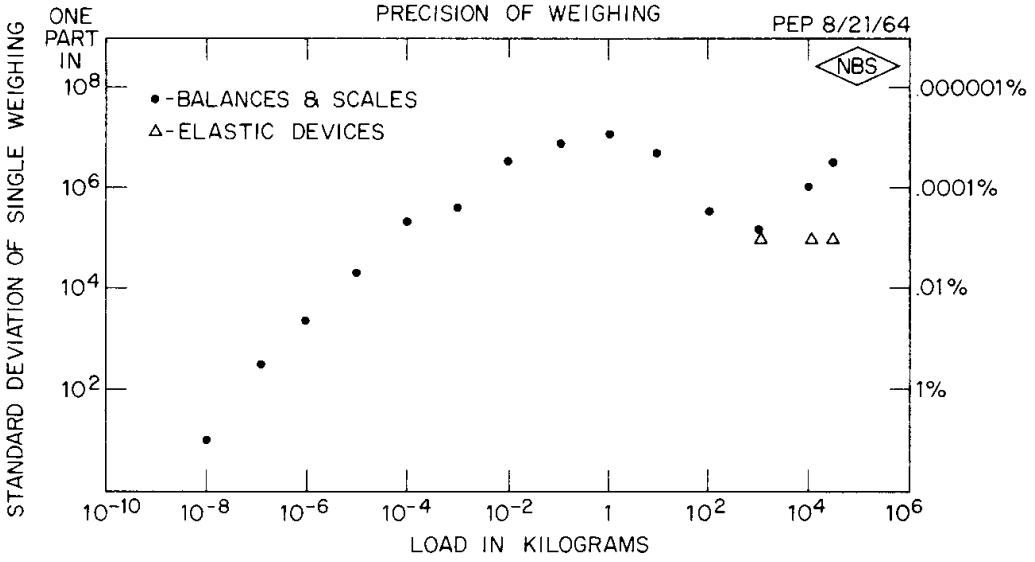

FIGURE 23.4 Precision of weighing.

measurements requires a minimal degree of understanding of the nature of measurement error. The starting point is to understand the nature of accuracy and precision. Figure 23.5 shows the meaning of these terms by example and by analogy. See, in this connection, ASTM 177-71 (listed in Appendix III) on use of the terms *precision* and *accuracy* as applied to measurement of the properties of materials. See also Mathur (1974) for a discussion on the influence of measurement variation in mass production of parts.

Accuracy. Suppose that we make numerous measurements on a single unit of product and that we then compute the average of these measurements. The extent to which this average agrees with the "true" value of that unit of product is called the *accuracy* of the instrument or measurement system that was employed. The difference between the average and the true value is called the *error* (also *systematic error, bias,* or *inaccuracy*) and is the extent to which the instrument is out of calibration. The error can be positive or negative. The *correction* needed to put the instrument in calibration is of the same magnitude as the error but opposite in sign. The instrument is still considered *accurate* if the error is less than the *tolerance* or maximum error allowable for that grade of instrument.

Accuracy and *error* are quantified as a difference between (1) the average of multiple measurements and (2) the true value. As will be seen, each of these is surrounded by a fringe of doubt. Consequently, the expression of accuracy must show the extent of these doubts if the full meaning of the numbers is to be conveyed.

Precision. Irrespective of accuracy of calibration, an instrument will not give identical readings even when making a series of measurements on one single unit of product. Instead, the measure-

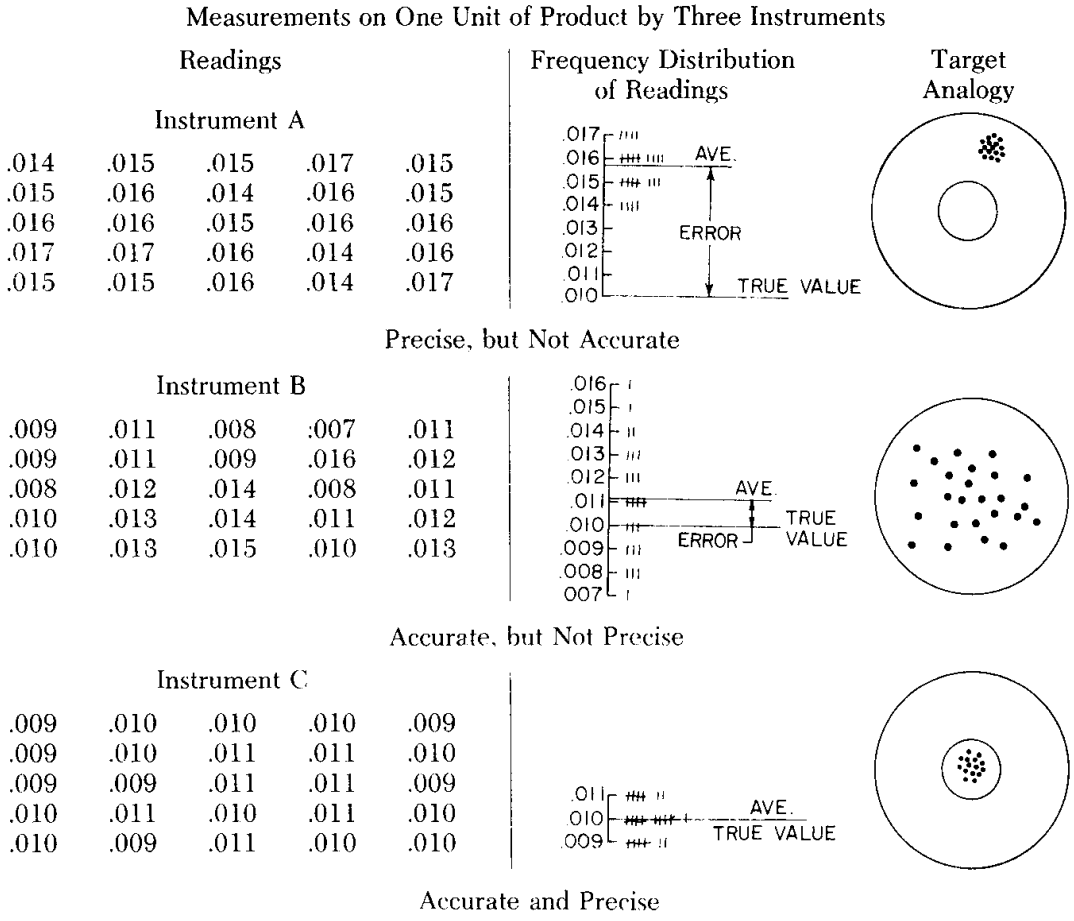

FIGURE 23.5 Accuracy and precision.

ments scatter about the average, as exemplified in Figure 23.5. The ability of the instrument to reproduce its own measurements is called its *precision,* and this varies inversely with the dispersion of the multiple (*replicated*) measurements.

Experience has shown that any measurement system has an inherent dispersion that is itself reproducible, measurable, and therefore (once known) predictable. This inherent precision of measurement parallels the inherent process capability of a machine tool. (The parallel extends to the requirement that the system be in a state of statistical control.) Quantification of precision is in terms of the standard deviation of replicated measurements and is expressed by σ (sigma), the statistical symbol for standard deviation of a population.

Normally, recalibration can improve the accuracy of an instrument by reducing its error. However, recalibration normally does not improve the precision of the instrument, since this precision remains relatively constant over the working range.

Sources of Error. Systematic error and dispersion of measurements have their origin in several well-known components of measurement error. (In some industries, e.g., chemical processes, the measurement problems are so severe that development of valid test procedures is a major step in the launching of a new product or process.)

Within-Operator Variation. The same operator, inspector, or technician, even when using the same measuring system on the same unit of product, nevertheless will come up with a dispersion of readings. This variation is usually referred to as *within-operator variation.*

Between-Operators Variation. When two operators use the same measuring system on the same products, they usually will exhibit differences traceable to differences in operator technique. These differences are called *between-operators variation* and can be exhibited both as systematic error and as differences in dispersion.

Materials Variation. In many cases it is not feasible to conduct replicated tests on the same "unit of product"; i.e., the product is changed or destroyed by testing. In other cases, the standard itself is consumable (e.g., hardness test blocks), so material variation affects the standard as well. In these cases, where replicate testing is not feasible, the variations due to operator, equipment, and test method are composited with the materials variation. Sometimes it is feasible to resolve these composites into their components, and sometimes it is not. A further complication is the case of perishable materials, which may require use of calibrations that relate time elapsed to degradation suffered.

Test Equipment Variation. Instruments are subject to numerous sources of error, both within a single instrument and between instruments: nonlinearity, hysteresis (e.g., gear backlash), drift due to temperature rise, and sensitivity to extraneous "noise" (e.g., magnetic, thermal, electrical fields). Each technology is subject to its own unique array of troubles. These instrument troubles are multiplied by the *fixturing* troubles of connecting the instruments into the larger test equipment units and of connecting the test specimens for tests. These fixturing troubles include such problems as making good electrical connections, fastening mechanical linkage, locating probes precisely, and so on.

Test Procedure Variation. In those cases where more than one test procedure is available to conduct measurement, it is essential to determine the relative variations, since these comprise one of the criteria for judging the adequacy of the procedure used.

Between-Laboratories Variation. This is a major problem both within companies and between companies. Some major programs must await resolution of this problem before they can be concluded, e.g., industry standardization of materials, test equipment, and test procedures. In like manner, variation between vendor and buyer laboratories may be at the root of a major quality problem. So extensive is the need to reduce between-laboratories variation that standard procedures have been evolved for the purpose (see *Quality Control Handbook,* 4th ed., Section 18, under Interlaboratory Test Programs).

Composite Errors. The observed measurements are, of course, a resultant of the contributing variations. Generally this resultant or composite is related to the component variables in accordance with the formula

$$\sigma_{obs}^2 = \sigma_w^2 + \sigma_b^2 + \sigma_m^2 + \sigma_e^2 + \sigma_f^2 + \cdots$$

where σ_{obs} is the standard deviation of the observed measurements and σ_w, σ_b, σ_m, σ_e, σ_f, etc. are the standard deviations reflecting the size of the variables that affect precision, i.e., within operator, between operators, material used, in test equipment, in test procedure, etc., respectively.

This relationship is valid provided the variables are independent of each other, which they often are. Where two or more of the variables are interrelated, then the equation must be modified. If, for example, variables *A* and *B* are interrelated, then

$$\sigma_T^2 = \sigma_A^2 + \sigma_B^2 + \rho_{AB}\sigma_A\sigma_B$$

where σ_T^2 = total variance
σ_A^2 = variance of *A*
σ_B^2 = variance of *B*
ρ_{AB} = the correlation coefficient (ρ) of *A* and *B*

In many cases it is feasible to quantify the effect of some component sources of variation by simple designs of experiment. When an instrument measures a series of different units of product, the resulting observations will have a scatter that is a composite of (1) the variation in the system of measurement and (2) the variation in the product itself. This relationship can be expressed as

$$\sigma_{obs} = \sqrt{\sigma_{prod}^2 + \sigma_{meas}^2}$$

where σ_{obs} = σ of the observed data
σ_{prod} = σ of the product
σ_{meas} = σ of the measuring method

Now, solving for σ_{prod},

$$\sigma_{prod} = \sqrt{\sigma_{obs}^2 - \sigma_{meas}^2}$$

It is readily seen that if σ_{meas} is less than one-tenth σ_{obs}, then the effect on σ_{prod} will be less than 1 percent. This is the basis of the rule of thumb that the instrument should be able to divide the tolerance into about 10 parts.

To illustrate, in one shop the validity of a new type of instrument was questioned on the ground that it lacked adequate precision. The observed variation σ_{obs} was 11 (coded). An experiment was conducted by having the instrument make replicate checks on the same units of product. The σ_{meas} was figured out to be 2. Then, since

$$\sigma_{prod}^2 = \sigma_{obs}^2 - \sigma_{meas}^2$$

$$\sigma_{prod} = \sqrt{121 - 4} = \sqrt{117} = 10.8$$

This was convincing proof that the instrument variation did not significantly inflate the product variation.

In another instance involving the efficiency of an air-cooling mechanism, the observed variation σ_{obs} was 23, and the variation on retests σ_{meas} was 16. Thereupon,

$$\sigma_{prod} = \sqrt{23^2 - 16^2} = \sqrt{529 = 256} = \sqrt{273} = 16$$

This showed that the measurement variation was as great as the product variation. Further study disclosed that the measurement variation could be resolved into

Variable	σ of that variable	σ^2
A	14	196
B	5	25
All other	7+	52
		273

It became clear that real progress could be made only by improving variable *A*, and the engineers took steps accordingly.

To quantify the individual components of variation requires still more elaborate analysis, usually through a special design of experiment (see Section 47, Design and Analysis of Experiments). [See also McCaslin and Grusko (1976) for a discussion on an attribute gage study procedure and Ezer (1979) for statistical models for testing vial-to-vial variation in medical laboratories.]

Statement of Error. In publishing results, it is necessary to make clear the extent of error in those results. Lacking clear conventions or statements, those who review the results simply do not know how to evaluate the validity of the data presented. To make clear the extent of error present in the data, metrologists have adopted some guidelines that are ever more widely used.

Effect of Reference Standards. Accuracy of an instrument is expressed as the difference between *T*, the "true" value, and $\bar{X}_m$, the average of the replicated measurements. The reference standard used is *assumed* to be the true value, but of course, this is not fully valid; i.e., the standards laboratory is able to make only a close approximation. In theory, the "true" value cannot be attained. However, the extent of error can be ascertained through the use of replication and other statistical devices. As long as the systematic error of the standard is small in relation to the error of the instrument under calibration (the usual situation), the error of the standard is ignored. If there is some need to refer to the error of the standard, the published measurements may include a reference to the standard in a form similar to "as maintained at the National Institute of Science and Technology."

Effect of Significant Systematic Error. When the systematic error is large enough to require explanation, the approved forms of explanation consist of sentences appended to the data, stating (for example), "This value is accurate within $\pm x$ units," or "This value is accurate within $\pm y$ percent." [See National Bureau of Standards (1965).] If necessary, these statements may be further qualified by stating the conditions under which they are valid, e.g., temperature range.

It is a mistake to show a result in the form $a \pm b$ with no further explanation. Such a form fails to make clear whether *b* is a measure of systematic error, an expression of standard deviation of replicate measurements, or an expression of probable error, etc.

Effect of Imprecision. The quantification of precision is through the standard error (standard deviation), which is the major method in use for measuring dispersion. In publishing the standard error of a set of data, care must be taken to clear up what are otherwise confusions in the interpretation.

1. Does the standard error apply to individual observations or to the average of the observations? Unless otherwise stated, it should be the practice to relate the published standard error to the published average, citing the number of observations in the average.
2. If uncertainty is expressed as a multiple of the standard error, how many multiples are used? An approved form of expression is "…with an overall uncertainty of ±4.5 km/s derived from a standard error of 1.5 km/s."
3. Is the standard error based solely on the data presented or on a broader history of data? To clarify this requires still more intricate wording, since a dispersion based solely on the current data is itself uncertain [see Eisenhart (1968)].

Effect of Combined Systematic Error and Imprecision. In these cases the expression of the published result must make clear that both types of error are present and significant. Eisenhart (1968)

recommends a phraseology such as "…with an overall uncertainty of ±3 percent based on a standard error of 0.5 percent and an allowance of ±1.5 percent for systematic error."

Errors Negligible. Results also may be published in such a way that the significant figures themselves reflect the extent of the uncertainty. For example, in the statement, "The resistance is 3942.1 Ω correct to five significant figures," the conventional meaning is that the "true" value lies between the stated value ±0.05 Ω.

CALIBRATION CONTROL

Measurement standards deteriorate in accuracy (and in precision) during use and, to a lesser degree, during storage. To maintain accuracy requires a continuing system of calibration control. The elements of this system are well known and are set out below.

(The terminology associated with *calibration control* is not yet standardized. To put an instrument into a state of accuracy requires first that it be tested to see if it is within its calibration limits. This test is often referred to as *checking* the instrument. If, on checking, the instrument is found to be out of calibration, then a *rectification* or *adjustment* must be made. This adjustment is called variously *calibration*, *recalibration*, or *reconditioning.* In some dialects, the word *calibration* is used to designate the combination of checking the instrument and adjusting it to bring it within its tolerances for accuracy.)

While the same system can be applied to all levels of standards, as well as to test equipment, there are some significant differences in detail of application. Transfer standards are exclusively under the control of standards laboratories staffed by technicians whose major interest is maintaining the accuracy of calibration. In contrast, test equipment and, to some extent, working standards are in the hands of those production, inspection, and test personnel whose major interest is product and process control. This difference in outlook affects the response of these people to the demands of the control system and requires appropriate safeguards in the design and administration of the system.

New-Equipment Control. The control system regularly receives new elements in the form of additional standards, new units of test equipment, and expendable materials. These elements should be of proven accuracy before they are allowed to enter the system. The approach varies depending on the nature of the new item:

1. *Purchased precision standards.* These include high-accuracy gage blocks, standard cells, etc. Control is based on the supplier's calibration data and on his or her certification that the calibration is traceable to the National Institute of Standards and Technology. Where such purchased standards represent the highest level of accuracy in the buyer's company, any subsequent recalibration must be performed by an outside laboratory, i.e., the supplier, an independent laboratory, or the National Institute of Standards and Technology.
2. *Purchased working standards.* These are subjected to "incoming inspection" by the buying company unless the demonstrated performance of the supplier merits use of an audit of decisions (see Section 21).
3. *New test equipment.* This equipment is intended for use in checking products and processes (Figure 23.6). However, it usually embodies measuring instruments of various sorts and may well include working standards as well, i.e., test pieces ("masters" for in-place check of calibration).
4. *Test materials.* These include consumable standards as well as expandable supplies such as reagents or photographic film. Variability in such materials can affect the associated measurements and calibrations directly.

For example, a manufacturer of sandpaper needed a uniform material on which to test the abrasive qualities of the sandpaper. The engineer investigated the possibility of using plastic blocks and

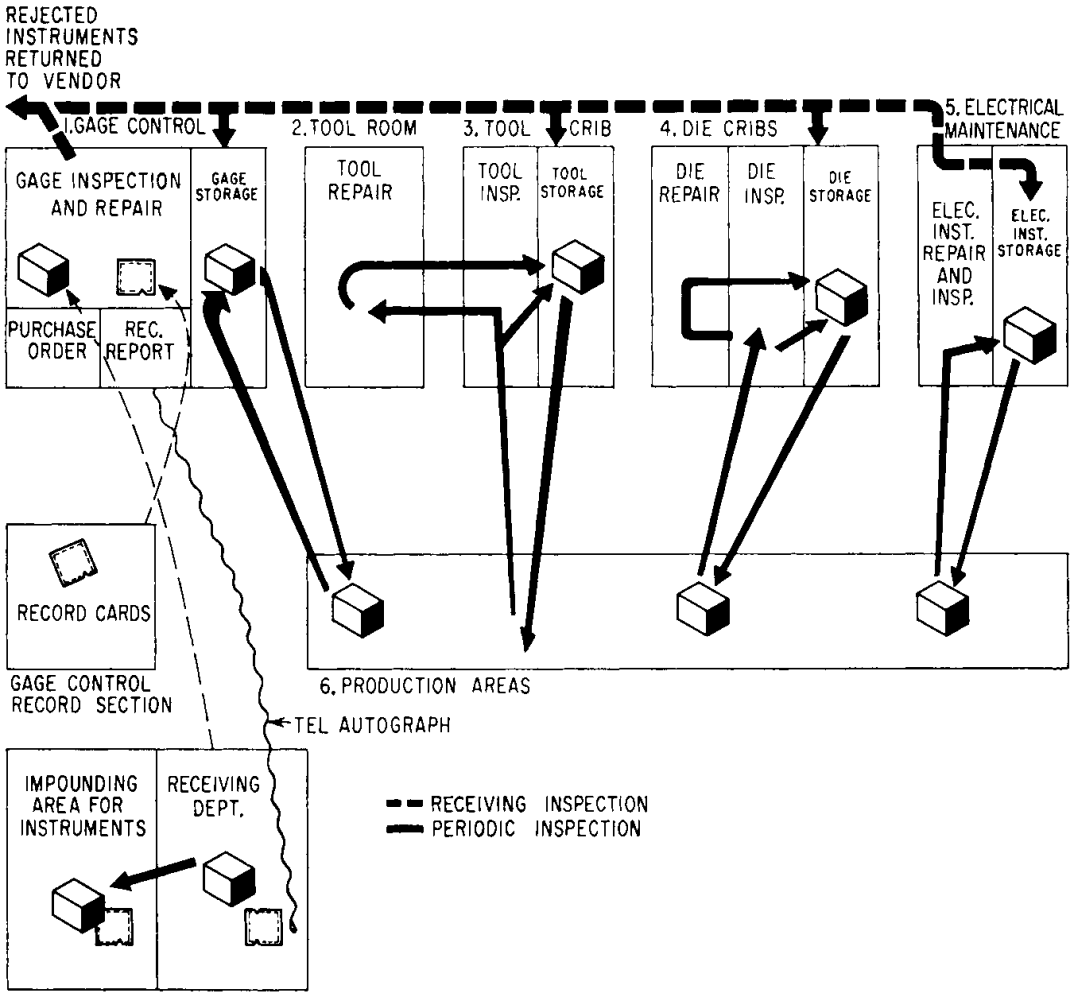

FIGURE 23.6 Flow diagram for gage control.

found that the plastics manufacturer was using the same sandpaper as a means of testing the toughness of the plastic. For some of these materials, the suppliers could provide data on variability. For the rest, it was necessary to discover the variability through analysis, as discussed earlier under Error of Measurement.

Inventory and Classification. A systematic approach to calibration control starts with a physical inventory of all standards, instruments, and test equipment. (Where tooling is used as one of the means of product inspection, such tooling is commonly included in the list of items to be systematically controlled for accuracy.)

For each item that is to enter the system, a database record is created. This record contains the historical origin of the item, its assigned serial number, the checking schedule, and related information. The record is also designed to accommodate information on the results of checking and the repairs needed [see also Woods (1978)]. The physical test equipment is also marked with an assigned serial number for identification and traceability in the system.

Calibration Schedules. These are established by class of equipment and are varied to reflect precision, nature, and extent of use and still other factors. At the outset, these schedules are established by judgment and bargaining. Later, as data become available on the results of checking, it becomes feasible to change the schedules in the interest of greater effectiveness and economy.

The broad intent of calibration schedules is to detect deterioration beyond tolerable levels of accuracy. This deterioration takes place primarily though use and secondarily through the passage of time. As a result, the calibration schedules describe the extent of use or of elapsed time in several ways:

1. *Elapsed calendar time.* This method is in widest use. It establishes a fixed calendar time, e.g., 3 months, as a checking interval. At the end of the 3 months, steps are taken to check the equipment in accordance with schedule.
2. *Actual amount of use.* This is based on counting the actual use, e.g., number of units of product checked by the equipment. The count may be made (*a*) manually, by the inspectors, (*b*) through automatic counters installed in the equipment, or (*c*) by programming the computers to show the amount of testing performed based on production schedules.
3. *Test accuracy ratio* (*TAR*) *control.* This is a systems approach that analyzes the degree to which interrelated parameters are identified and controlled within a *traceability cone.* Minimum TARs for each traceable parameter are measured and controlled [see also Tobey (1979)].

Adherence to Schedule. This vital detail makes or breaks the entire system of calibration control. Generally, the transfer standards and most working standards pose no problem of adherence to schedule, since they are in the custody of a few standards laboratories and a relatively few associated technicians. In contrast, the test equipment (and some working standards) are widely scattered over numerous locations and are in the custody of thousands of workers, inspectors, and testers. Some of these individuals can be relied on to see that the checking schedule is followed, but many cannot.

In part, the problem is one of lack of knowledge of when the recalibration is due. The shop personnel require the aid of a memory system if they are to know which piece of equipment is due to be checked that day. They may recall what the checking intervals for each class of equipment are, but they cannot recall what the date of the last calibration was.

Some systems for adherence to schedule make use of ingenious color codes or labels that mark on each unit of equipment the date it was put back into service. (These codes are often extended to identify the grades of the standards themselves, whether primary, secondary, etc.) For large units, the expiration date also may be entered on a maintenance card that is attached to the unit. Such dates are an aid to personnel for adhering to the checking schedule.

However, an added problem is that of motivation. The numerous users of test equipment are quite concerned with recalibration when trouble is encountered but less concerned when things seem to be going smoothly. In these latter cases, interruptions for calibration can even be a nuisance.

The solution is to give responsibility (for adherence to schedule) to the standards laboratory rather than to the production, inspection, and test personnel. When this proposal is made to practicing managers, they seldom accept it purely on grounds of theory of organization. However, when it is proposed that a sample of instruments be taken at random and checked for calibration (as a test of the existing "system" of calibration control), these same managers are quite willing to conduct such a test. The resulting disclosure of the actual state of calibration of the sample (of 25 to 100 instruments) is then decisive in convincing the managers of the need for a revision in the system of adherence to schedule. Under this assignment, thelaboratory organizes a plan of checking that will keep up with the scheduled load [see also Gebhardt (1982)].

In administering the checking plan in the time-interval system, the database is queried, using the calibration dates, to provide a list showing which standards and equipment are due to be checked in the forthcoming week.

Calibration Practice. To ensure accuracy and to establish traceability, control laboratories have evolved some widely used procedures. Individual responsibility is established by requiring that all concerned sign for their actions. The equipment record, retained either in hardware or software, carries these signatures, as do the labels on the equipment. Dates are recorded for all actions in view of the role of elapsed time in the calibration procedures.

Manuals of practice are established, including tolerances for accuracy and methods to be used in calibration. In some types of test, these methods must be spelled out in detail, e.g., temperature or

humidity controls, time cycle, human technique, etc. (Witness the detail of some of the ASTM standards on test method.)

Training programs are established for personnel, including (in some cases) formal qualification certificates to attest to proficiency. Equipment is tamperproofed through sealing the adjusting screws. (The seals are then imprinted with the stamp of the laboratory.) In like manner, panels and drawers of test equipment are lock-wired, and the wires are lead-sealed together. (The laboratory takes no responsibility when seals are broken, and the company takes stern measures against tampering with the seals.) As a means of assisting enforcement, quality assurance audits are conducted to review the calibration control procedures.

Record and Analysis of Results. It is most useful to keep a record of the results of checking calibration and of the extent of work done to restore accuracy. Typically, such a record lists

Observed deficiencies in the equipment

Causes of out-of-calibration conditions

Repair time and recalibration time

Periodic analysis of these data then becomes the basis for

1. Reducing checking for equipment shown to be stable
2. Redesigning equipment to eliminate causes of repetitive failure

Organization for Calibration Control. Ensuring that measuring devices are calibrated correctly is critical to ensure that product is conforming to the customer's requirements. The best inspector is only as good as his or her gages. If the gage is in error, then we could reject good parts or accept bad parts. Both results are costly to the company; both send a message to the customer that the company does not have a basic control system.

A technician trained in metrology—the science of gaging—is the determining factor in the accuracy of the gage by calibrating to a recognized master, traceable in the United States to the National Institute of Standards and Technology or similar international organization. The calibration process is performed by following a set of test procedures developed for scientific instruments.

The quality system can elect to calibrate all measuring devices in house, send them to a commercial calibration service, return the device to the original manufacturer, or a combination of all of the above. The determining factor is usually based on master calibration up-front cost, frequency of needed calibration control, and turnaround time of outside source.

Regardless of how the actual calibration function is allocated, the frequency of calibration of each measurement device must be specified by a sytematic calibration frequency program. This program defines the categories of inspection, measuring, and test equipment to be covered and assigns responsibility for operating the program and maintaining records.

There are no standard rules and no body of knowledge that dictate how often measuring instruments should be inspected. The determining factors of calibration are use and mishandling (Palumbo 1997).

Physical design of the laboratory workplace has been greatly complicated by the proliferation of many varieties of specialized testing: ultrasonic, x-ray, vibration, shock, acceleration, heat, humidity, etc. The details of these designs are beyond the scope of this handbook. The practitioner must consult with the available experts: equipment manufacturers, researchers, metrologists, and still others. This must be a continuing process, since there is continuing progress in development of new tests and standards.

HUMAN FACTORS IN INSPECTION

A myriad of factors can influence inspector behavior; they are summarized by Baker (1975), as shown in Table 23.9. One factor, visual acuity or sight, is the dominant sense in human beings, and

great reliance is placed on it in inspection tasks. The effectiveness of the use of sight depends largely on eye movements that bring the images of significant features of the material being inspected to the most sensitive part of the retina. However, experience and studies have shown that the other factors in Table 23.9 have an interrelated effect on the effectiveness, productivity, reliability, and accuracy of the inspector. [See Megaw (1978) for related studies carried out in a textile factory.]

A detailed discussion of inspector errors below follows the discussion on human factors; however, it appears appropriate to first discuss the techniques and measurements developed to improve the reliability of the inspection function.

Machine Vision. The term *machine vision,* or *noncontact inspection,* is applied to a wide range of electrooptical sensing techniques from relatively simple triangulation and profiling to three-dimensional object recognition and bin picking, techniques based on sophisticated computerized image-analysis routines. The applications are broad, ranging from relatively simple detection and measuring tasks to full-blown robot control. (See Table 23.10.)

The incentive to introduce machine-based systems, e.g., robots, is obvious—to eliminate human error. [See Spow (1984) for a discussion of robots in an automatic assembly application.] The key influences behind the growth of the machine-vision industry also include inspector capability, inspector productivity, and inspection costs. [See also Nelson (1984) for a review of machine-vision equipment, some designed to eliminate process contamination, in addition to human error.] Table 23.11 discusses these factors as they relate to inspecting printed circuit boards. [See Ken (1984) for a detailed discussion of automated optical inspection (AOI) of printed circuit boards (PCB); see also Denker (1984) for a detailed discussion on justifying investments in automatic visual PCB testing.]

Vision-based systems—human or machine—involve an inspection procedure: the examination of a scene. The examination, in turn, can lead to recognition of an object or feature, to a quality decision, or to the control of a complex mechanism (Schaffer 1984). When a human inspector is involved, human judgment and perception have an influence on the quality assessment process.

The most common application for vision systems involves measuring critical dimensions, detecting flaws, counting/sorting, assembly verification, position analysis, character or bar-code reading/verification, and determination of presence/absence of features on small parts.

TABLE 23.9 Variables Influencing Inspector Behavior

1. Individual abilities
   - *a.* Visual activity
   - *b.* General intelligence and comprehension
   - *c.* Method of inspection
2. Task
   - *a.* Defect probability
   - *b.* Fault type
   - *c.* Number of faults occurring simultaneously
   - *d.* Time allowed for inspection
   - *e.* Frequency of rest periods
   - *f.* Illumination
   - *g.* Time of day
   - *h.* Objective of conformance standards
   - *i.* Inspection station layout
3. Organizational and social
   - *a.* Training
   - *b.* Peer standards
   - *c.* Management standards
   - *d.* Knowledge of operator or group producing the item
   - *e.* Proximity of inspectors
   - *f.* Reinspection versus immediate shipping procedures

TABLE 23.10 Machine-Vision Applications

1. Inspection
   a. Dimensional accuracy
   b. Hole location and accuracy
   c. Component verification
   d. Component defects
   e. Surface flaws
   f. Surface-contour accuracy
2. Part identification
   a. Part sorting
   b. Palletizing
   c. Character recognition
   d. Inventory monitoring
   e. Conveyor picking (overlap, no overlap)
   f. Bin picking
3. Guidance and control
   a. Seam-weld tracking
   b. Part positioning
   c. Processing/machining
   d. Fastening/assembly
   e. Collision avoidance

TABLE 23.11 Selected Example of Factors and Effect of Automated Optical Inspection (AOI) on Printed Circuit Board

Manufacturing inspection costs	Inspection accounts for up to 30 percent of the manufacturing costs of complex double-sided and multilayer boards.
Inspection capability	Human inspection capability decreases well before 5 mil because of fatigue
Inspection productivity	AOI reduces the number of inspections from 30 to 7 per shift while obtaining a 1 percent yield improvement and fivefold increase in inspection speeds.

If parts are moving and not indexed, an electronic shutter camera may be needed. Many systems can handle randomly oriented parts, but throughput rates will be lower than for fixtured parts, and if the need is to inspect more than one side or area of the parts, separate cameras may be required.

Repetitive Function. In highly repetitive subjective inspection (e.g., inspecting parts on an automated paint line), the process of perceiving can become numbed or hypnotized by the sheer monotony of repetition. In some way the scanning process and the model become disconnected, and the observer sees only what he or she expects to see but does not see anything not actively expected. The situation often can result in bad parts passing through the process and should be evaluated closely for effectiveness before being initiated. The known remedies are

- Break up the benumbing rhythm with pauses.
- Introduce greater variety into the job.
- Increase the noticeability of the faults through enhancing.
- Provide background and contrast cues.
- Arrange for frequent job rotation.

Vigilance becomes a big problem when faults are obvious and therefore serious but infrequent and unpredictable. What traveler has not wondered about the protection provided by 100 percent x-ray and metal-detection inspection of airline passenger carryon baggage and personal effects? In 1993, an audit by the U.S. General Accounting Office of airport security measures in the United States revealed the ease with which the "secure" areas of America's airports could be penetrated. Of the hundreds of attempts to breach security, 75 percent were successful; some of these successes were attributable to a failure of 100 percent inspection. The auditing agent who succeeded in passing a live hand grenade through the screening system provided an especially troublesome example (Gleick 1996).

INSPECTION ERRORS

The inspector, as the human element in the inspection process, contributes importantly to inspection errors. Inspection errors due to the inspector, called *inspector errors*, are discussed here. The reliability of the human inspector was discussed previously under human factors (as opposed to noncontract machine inspection or the machine-vision system). Other sources of inspection error, e.g., vague specifications, lack of standards, inaccurate instruments, etc., are discussed elsewhere in this handbook. (*Note:* The problem of human error is common to operators, inspectors, and anyone else. For an extensive discussion of problems of worker error, see Section 22, under Concept of Controllability: Self-Control.)

Inspector errors are of several categories:

Technique errors

Inadvertent errors

Conscious errors

Each of these categories has its own unique causes and remedies. Collectively, these inspector errors result in a performance of about 80 percent accuracy in finding defects; i.e., inspectors find about 80 percent of the defects actually present in the product and miss the remaining 20 percent. [*Note by the editor* (*Juran*): Numerous studies in various countries have yielded the 80 to 20 ratio as a broad measure of quantified inspector accuracy. For example, Konz and coworkers (1981) found that inspectors were only catching about 80 percent of the defects in glass subassemblies.] However, little is known about the relative importance of each category of inspector error (e.g., lack of technique, inadvertence, conscious). For further discussion, see Tawara (1980).

Technique Errors. Into this category are grouped several subcategories: lack of capacity for the job, e.g., color blindness; lack of knowledge due to insufficient education or job training; and lack of "skill," whether due to lack of natural aptitude or to ignorance of the knack for doing the job. Technique errors can be identified in any of several ways:

Check Inspection. A check inspector reexamines work performed by the inspector, both the accepted and rejected product. Figure 23.7 shows an example of the results of such check inspection of the work of several inspectors. It is evident that inspectors C and F operate to loose standards, whereas inspector B operates to tight standards. Inspector E shows poor discrimination in both directions.

Round-Robin Inspections. In this analysis, the same product is inspected independently by multiple inspectors. The resulting data, when arrayed in a matrix (usually with defect type along one axis and the inspectors along the other axis), shows the defects found by each inspector in relation to the inspectors as a group.

Repeat Inspections. In this method, the inspector repeats his or her own inspection of the product without knowledge of his or her own prior results. The analysis of the resulting data discloses the extent of the consistency or lack of consistency of the inspector's judgments.

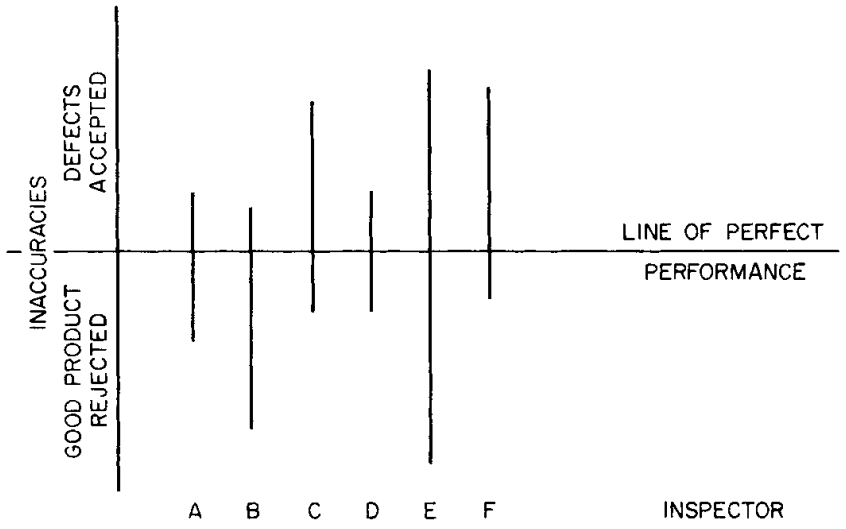

FIGURE 23.7 Analysis of inspection errors.

Standard Sample Array. In this method, the inspector makes an "examination" by inspecting a prefabricated mixture of product consisting of good units plus various kinds of defects. The standard sample array is known by various names, including *job sample.* For added discussion, refer to Harris and Chaney (1969). All units were previously carefully graded by a team of experts and numbered for ready analysis of results. The inspector's score and his or her pattern of errors all point to the need, if any, for further training or other remedial steps. (In effect, the check inspection is conducted before the inspection.)

For example, in a company making glass bottles, an attempt was made to correlate process variables with the frequency and type of defects found by inspectors stationed at the cold end of the annealing lehr. The experiment failed because inspector variability from shift to shift exceeded product variability. This also threw suspicion on the accuracy of the inspection performed by the final product sorters at the end of the line. A standard sample array of 500 bottles was created and was used to examine the inspectors. (The examinations were conducted in the Training Department on a miniature lehr.) The suspicions turned out to be well founded (consulting experience of J. M. Juran).

Remedies for Technique Errors. The need is to provide the missing skill or know-how and to answer the inspector's proper question, "What should I do differently from what I am doing now?" Unless the inspector is in a position to discover the answer for himself or herself, the answer must be provided by management. If no answer is provided, there will be no change in performance.

The various methods of analysis discussed earlier all can provide some clues that suggest the type of remedial action needed. In particular, use can be made of the concept of finding the knack. Under this concept, the data on inspector performance are analyzed to discover which inspectors give consistently superior performances and which are inferior. Next, a study is made of the work methods used by both types of inspectors to identify the differences in methods. Analysis of these differences often discovers what is the secret know-how (knack) being used to get the superior performance (or what is the secret ignorance that results in poor performance). Finally, the knack can be transferred to all inspectors through retraining (Czaja and Drury 1981 and Cooper 1980) or through being embodied in the technology (Kusch 1979).

Where the technique errors are the result of lack of job capacity, the foregoing may be of no avail, and the need may be to foolproof the operation (see below) or to reassign the inspector to a job for which he or she does have adequate job capacity.

Certification of Inspectors. In critical inspections involving inspector judgment (e.g., interpreting x-rays of critical welds), it is increasingly the practice to require that the inspector be formally certified as qualified to do this job. [See Gibson (1983) for a further discussion of inspector qualifications in offshore industries.] The qualification process follows a well-standardized series of steps:

A formal training program on how to do the job

A formal examination, including a demonstration of successful performance of the job

A formal certificate attesting to the success in the examination

A license to do the job for some designated period of time

A program of audit to review performance and to serve as a basis for renewing the license

In some companies this concept of certification has been based on the "escape rate," i.e., the extent to which defects escape detection, as determined by subsequent check inspection. (See Measure of Inspector and Test Accuracy, below.) When this concept is used, a limited "license" (e.g., 2 months) is given to the inspector, subject to renewal if check inspection results continue to be favorable.

Inadvertent Inspector Errors. These errors are characterized by the fact that at the time the error is made, the inspector is not even aware he or she is making an error. Also, the best intentions are present—the inspector wants not to make any errors. The term *inadvertent* or *unavoidable* is used to connote the fact that the human being is simply unable to achieve perfection, no matter how good his or her intentions. (This topic is closely related to inadvertent worker errors. For added discussion, see Section 5, under Inadvertent Errors.)

The theory of inadvertence has wavered up and down. For years it was the sincere belief of many inspection supervisors that when product was inspected 100 percent, the inspectors would find all the defects. Numerous unpublished and published studies have since demonstrated that human inspectors do not find all defects present. By and large, the human inspector finds about 80 percent of the defects present and misses the remaining 20 percent.

While the 80 to 20 ratio is widely accepted, there are numerous aspects that are not fully researched, i.e., how this ratio changes with percentage defective in the product, with types of inspection (e.g., visual, mechanical gaging, electrical testing), with product complexity, with amount of time allotted for inspection, etc. For a discussion from the viewpoint of human factors plus some supporting data (e.g., that increased product complexity results in increased inspector error), see Harris and Chaney (1969, pp. 77–85).

Inspection fallibility can be demonstrated easily in the industrial classroom. The following sentence has been used thousands of times:

FEDERAL FUSES ARE THE RESULTS OF YEARS OF SCIENTIFIC STUDY COMBINED WITH THE EXPERIENCE OF YEARS

The sentence is flashed before the audience for 30 s or for a full minute. Each member is asked to count and record the number of times the letter *F* appears. When the record slips are collected and tallied, the result is invariable. Of the *F* 's present, only about 80 percent have actually been found.

The existence of so extensive an error rate has stimulated action on several fronts:

1. *To discover why inspectors make these errors.* To date, the research has not been adequate to provide conclusive answers, so industrial psychologists have not agreed on what the main causes are.
2. *To measure the extent of the errors.* Techniques for this are now available. See Measure of Inspector and Test Accuracy, below.
3. *To reduce the extent of these errors.* There is a wide assortment of remedies, as discussed below.

Remedies for Inadvertent Inspector Errors. In the absence of convincing knowledge of the causes of these errors, managers have resorted to a variety of remedies, all involving job changes in some form. These remedies include the following:

Error-Proofing. There are several forms of error-proofing that are widely applicable to inspection work: redundancy, countdown, and fail-safe methods. These are discussed in detail in Section 22, Operations, under Error-Proofing the Process. See also Inspection Planning, above.

Automation. This is really a replacement of the repetitive inspection by an automation that makes no inadvertent (or other) errors once the setup is correct and stable. The economics of automation and the state of technology impose severe limits on the application of this remedy. See Automated Inspection, above.

Sense Multipliers. Use can be made of optical magnifiers, sound amplifiers, and other devices to magnify the ability of the unaided human being to sense the defects. Development of a new instrument to do the sensing is the ultimate form of this multiplication. Evidently there is an optimum to the level of magnification, and this optimum can be discovered by experimentation. [For some industrial studies, see Harris and Chaney (1969, pp. 124–126, 137–142).]

Conversion to Comparison. In many types of inspection, inspectors must judge products against their memories of the standard. When such inspectors are provided with a physical standard against which to make direct comparison, their accuracy improves noticeably. For example, in the optical industry, scratches are graded by width, and tolerances for scratches vary depending on the function of the product element (lens, prism, etc.) To aid the inspectors, plates are prepared exhibiting several scratches of different measured widths so that the inspectors can compare the product against a physical standard.

Standards for comparison are in wide use: colored plastic chips, textile swatches, forging specimens, units of product exemplifying pits and other visual blemishes, etc. Sometimes photographs are used in lieu of product. There are also special optical instruments that permit dividing the field of view to permit comparison of product with standard.

In some cases it is feasible to line up units of product in a way that makes any irregularities become highly conspicuous, e.g., lining up the holes in a row. (The childhood row of tin soldiers makes it obvious which one has the broken arm.) Some practitioners advocate inspecting units of product in pairs to utilize the comparison principle. See Shainin (1972) for further discussion.

Templates. These are a combination gage, magnifier, and mask. An example is the cardboard template placed over terminal boards. Holes in the template mate with the projecting terminals and serve as a gage for size. Any extra or misplaced terminal will prevent the template from seating properly. Missing terminals become evident because the associated hole is empty.

Masks. These are used to blot out the view of characteristics for which the inspector is not responsible and concentrate attention on the real responsibility. Some psychologists contend that when the number of characteristics to be inspected rises to large numbers, the inspector error rate also rises.

Overlays. These are visual aids in the form of transparent sheets on which guidelines or tolerance lines are drawn. The inspector's task of judging the size or location of product elements is greatly simplified by such guidelines, since they present the inspector with an easy comparison for judging sizes and locations.

Checklists. These may be as simple as a grocery shopping list used to verify that you purchased all the items you originally planned to buy. At the other extreme, a checklist may consist of the countdown for the lofting of a new space shuttle. [See Walsmann (1981) for a discussion of the purpose, advantages, and drawbacks as well as the development and implementation of various checklists.]

Reorganization of Work. One of the theories of cause of inadvertent inspector errors is fatigue, due to inability to maintain concentration for long periods of time. Responses to this theory have been to break up these long periods in any of several ways: rest periods, rotation to other inspection operations several times a day, and job enlargement, e.g., a wider assortment of duties or greater responsibility. Some behavioral scientists urge reorganization of work on the broader ground of motivation theory, and they offer data to support this theory [see generally Harris and Chaney (1969, pp.

201–229)]. However, to date, there is no conclusive evidence that in Western culture reorganization of work (to provide greater participation, etc.) will give measurably superior results in work accomplishments. (See, in this connection, *Quality Control Handbook,* 4th ed., Section 10, under Processing System Design.)

Product Redesign. In some instances the product design is such that inspection access is difficult or that needless burdens are placed on inspectors. In such cases, product redesign can help to reduce inspector errors as well as operator errors. For some examples, see Section 22, Operations, under Error-Proofing the Process. [See also Smith and Duvier (1984).]

Errorless Proofreading. Beyond the techniques described in Section 22, Operations, under Error-Proofing the Process, there are special problems of error-proofing in inspection work. A major form of this is proofreading of text of a highly critical nature, i.e., critical to human safety and health. In such cases, the low tolerance for error has driven many companies to use redundant checking, despite which some errors still get through.

A closer look makes it clear that proofreading is of two very different kinds:

1. *Passive proofreading.* Here, the proofreader takes no overt action. For example, he or she silently reads the copy while someone else reads the master aloud. Alternatively, the proofreader silently reads both documents and compares them. In such cases it is quite possible for extraneous matters to intrude and dominate his or her attention temporarily.
2. *Active proofreading.* Here, the proofreader must take an overt action, e.g., he or she reads aloud, performs a calculation, etc. Such positive actions dominate the proofreader's attention and reduce the chance of error.

For a second example, in blood donor centers it was once usual to remove the whole blood from the donor, take it to a separate location, centrifuge it to remove a desired component, and then return the remaining fraction of the blood to the donor. This return demanded absolute assurance that the remaining fraction was being returned to that donor and to no one else. The system used involved piping both the donation and the return through tubing on which there were repeats of 10-digit numbers. The tubing was cut when the donation went to the centrifuge. Prior to return of the cells, two technicians checked to compare the two 10-digit numbers on the two ends of the cut tubing. One technician actively read the number on the tubing end attached to the donor. The other technician passively listened while comparing the number he or she heard with the number seen on the end of the tubing that was attached to the bag of blood cells.

It was sometimes feasible to use technology to make both technicians "active." For example, where the equipment was available, each was required to enter on a keyboard the number he or she saw. These signals went to a computer that compared the two numbers and signaled either a go-ahead or an alarm. This same principle of comparing two independent active sets of signals can be extended to any problem in proofreading.

Today another technology has eliminated altogether the need for this application of proofreading. The donor is now usually linked directly to the separating equipment. The blood is drawn from one arm, mixed with an anticoagulant, and passed through a separating machine that collects the specific component. The remaining blood components are returned to the donor in the opposite arm (Sataro 1997).

While the foregoing are listed as remedies for inadvertent inspector errors, most of them also can be used to reduce errors due to lack of skill or errors of a willful nature.

Procedural Errors. Aside from inadvertent failures to find defects, there are inadvertent errors in shipment of uninspected product or even shipment of rejected product. These errors are usually the result of loose shipping procedures. For example, a container full of uninspected product may be moved inadvertently in with the inspected product; a container full of defectives may be moved inadvertently into the shipping area. Such errors can be reduced by error-proofing the identification and shipping routines:

1. The inspector must *mark the product* at the time of inspection. Sometimes the inspector places the good product in one box and the bad product in another or places the good product lengthwise and the bad product crosswise. Lacking the markings, there is always the risk that between shifts, during rest periods, etc., the unmarked product will go to the wrong destination.
2. The product markings should be so distinctive that the product "screams its identity" to packers and shippers. Bar codes should be used where appropriate.
3. The colors used for markers that identify good product should be used for no other purpose. These markers should be attached only when the inspector finds that there remains nothing to be done but ship the product.
4. Issuance of markers used to identify good product should be restricted to specially chosen personnel. (These markers are a form of company seal.) For products of substantial value, serial numbers may be used as a further control.
5. The markers should provide inspector identity. In some companies, the system of identification includes the operators and packers.
6. Shipping personnel should be held responsible for any shipment of goods failing to bear an inspector's approval.

Conscious Inspection and Test Errors

Management-Initiated. The distinguishing features of the willful inspector error are that the inspector knows that he or she is committing the error and intends to keep it up. These willful errors may be initiated by management, by the inspector, or by a combination of both. However, with few exceptions, the major notorious quality errors and blunders have been traceable to the decisions of managers and engineers rather than to those of the inspectors at the bottom of the hierarchy. Management-initiated errors take several forms, all resulting in willful inspector errors.

Conflicting Management Priorities. Management's priorities for its multiple standards (quality, cost, delivery, etc.) vary with the state of the economic cycle. When the state of management priorities is such that conformance to quality standards is subordinated to the need for meeting other standards, the inspectors' actions are inevitably affected, since they also are given multiple standards to meet.

Management Enforcement of Specifications. When management fails to act on evidence of nonconformance and on cause of defects, the inspectors properly judge management's real interest in quality from these deeds rather than from the propaganda. For example, if the supervision or the material review board consistently accepts a chronic nonconformance condition as fit for use, the inspectors tend to quit reporting these defects, since they will be accepted anyhow.

Management Apathy. When management makes no firm response to suggestions on quality or to inspector complaints about vague information, inadequate instruments, etc., the inspectors again conclude that management's real interests are elsewhere. Consequently, the inspectors do the best they can with information and facilities that they believe to be deficient.

Management Fraud. Periodically a company manager attempts to deceive customers (or the regulators, etc.) through fictitious or deceitful records on quality. Seldom can a manager acting alone perpetrate such a fraud. The manager requires confederates who submit themselves to orders, usually in a way that makes clear to them the real character of what is going on. An inspector who is a willing accomplice (e.g., for a bribe) shares in the legal responsibility. However, the inspector also may be a most reluctant accomplice, e.g., an immediate superior gives orders to prepare nonfactual reports or to take actions clearly contrary to regulations. In such cases the inspector cannot escape taking some kind of risk, i.e., participation in a conspiracy versus the threat of reprisal if he or she

fails to participate. These management-initiated errors parallel closely those associated with conscious worker errors. See Section 22, Operations, under Concept of Controllability: Self-Control.

Inspector-Initiated. These errors likewise take multiple forms, and some take place for "good" reasons. It is important to understand the distinctions among these forms, since any misunderstandings are a breeding ground for poor industrial relations.

Inspector Fraud. The inspector is subjected to a variety of pressures. The most rudimentary forms are those by production supervisors and operators pleading for a "break." Sometimes this extends to a collusion where piecework payments are involved, both for quality and for quantity certification. At higher levels are cases in which an inspector is exposed to suppliers who have a good deal at stake in the lot of product in question. Even a situation in which inspectors dealing directly with production supervisors who outrank them involves substantial pressures to which inspectors should not be subjected.

Another form of inspector fraud consists of reporting false results solely to improve the outward evidence of one's own efficiency or to make life more convenient. For example, Figure 23.8 shows the results reported by an inspector after taking a sample of *n* pieces from each of 49 lots. There is a large predominance of three defects per lot reported in the sample (exactly the maximum allowable number). The reason was found to be the inspector's reluctance to do the paperwork involved in a lot rejection.

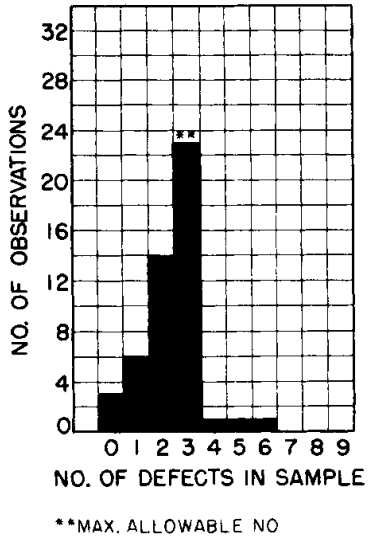

FIGURE 23.8 Inspector error made for personal convenience.

In the example in Figure 23.9, the inspector was to take a sample of 100 pieces from each lot, with no defects allowable. If one or more pieces were defective, an added sample of 165 pieces was to be taken, with a total of 3 defects allowed in the combined sample of 265. It is seen that the inspector reported defects in virtually every first sample of 100 pieces.

However, no defects were reported in most of the second (larger) samples. It was found that the inspector could improve personal efficiency by taking second samples, since the time allowance for taking the second samples was liberal. Inspector fraud can be minimized by

1. Filling inspection jobs only with persons of proved integrity
2. Restricting the down-the-line inspector to the job of fact finding, and reserving to the inspection supervision the job of negotiating and bargaining with other supervisors or executives
3. Rotating inspectors and workers to have both employee categories evaluate first hand the consequences of actions and decisions
4. Including inspectors in customer presentations or visits
5. Conducting regular check inspections and periodic independent audits to detect fraud
6. Taking prompt action where fraud is discovered

Inspector Shortcuts. These may be unauthorized omissions of operations that the inspector has reason to believe are of dubious usefulness; e.g., accidental omissions had failed to give evidence of trouble. In some cases there is a shared blame; i.e., management has imposed a highly disagreeable task. For example, in a company making "tin cans," one inspection involved cutting up a can with hand-held tin shears, submerging the pieces in chloroform to remove the enamel, and measuring the thickness of the bare pieces with a micrometer. The cutting process was tedious and the chloroform

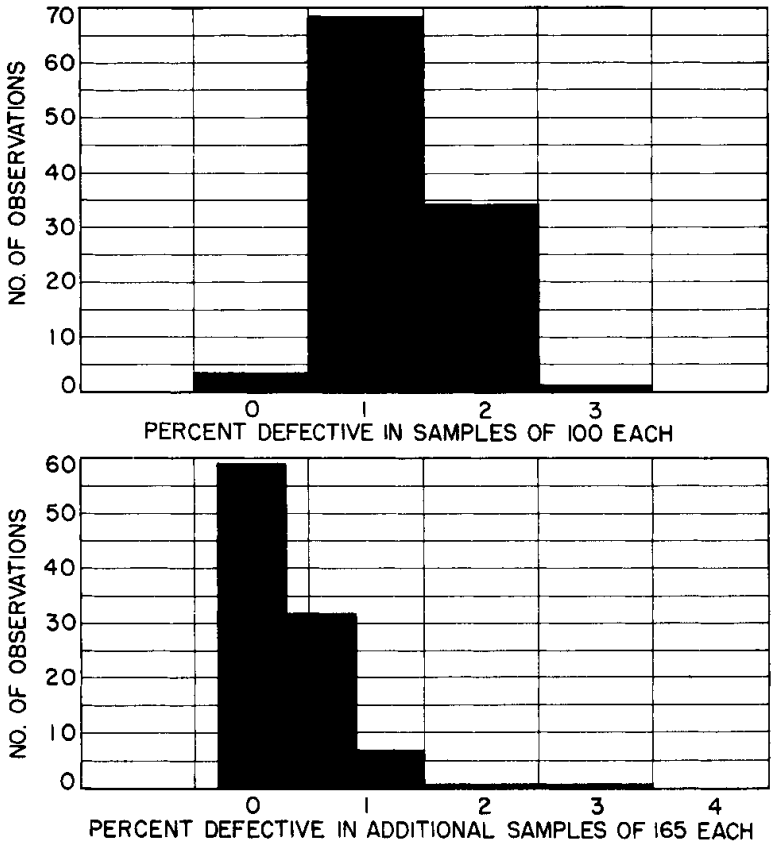

FIGURE 23.9 Inspector error made to improve personal efficiency.

was irritating to the skin, so the inspectors avoided the operation as much as possible. When better cutting tools and a different solvent were provided, the problem became minimal.

Flinching. This is the tendency of inspectors to falsify the results of inspection of borderline product. Flinching is actually widespread among all persons who report on performance versus goals and especially when it is their own performance. Figure 23.10 shows a frequency distribution of measurement on volume efficiency of electronic receivers. There is an "excess" of readings at the specification maximum of 30, and there are no readings at all at 31, 32, or 33. Retest showed that the inspector recorded these "slightly over" units of product at 30. By this flinching, the inspector, in effect, changed the specification maximum from 30 to 33. This is a serious error (Juran 1935).

Flinching during variables measurements is easy to detect by check inspection, which is also conducted on a variables basis. Analysis of the inspector's variables data likewise will detect flinching (as in the preceding example).

The remedy for flinching is an atmosphere of respect for the facts as the ethical foundation of the Inspection Department. The main means for achieving this are examples set by the inspection supervisors.

One way *not* to deal with flinching is to criticize the inspector on the basis that the pattern of readings does not follow the laws of chance. Such criticism can be interpreted as being aimed at the symptom (the unnatural pattern of reading) rather than the disease (recording fictitious instead of factual readings). The risk is that inspectors will try to meet such criticisms by trying to make the false results look more natural, hence eliminating the symptom but not the disease.

Flinching also takes place during attribute inspection. Numerous studies have shown that inspector errors in rejecting good product outnumber the errors of accepting bad product. In part, this arises because the good product outnumbers the bad and hence affords greater opportunity for error. However, it also arises in part from the fact that acceptance of defects often comes dramatically to the attention of higher management, whereas rejection of good product seldom does so. These same

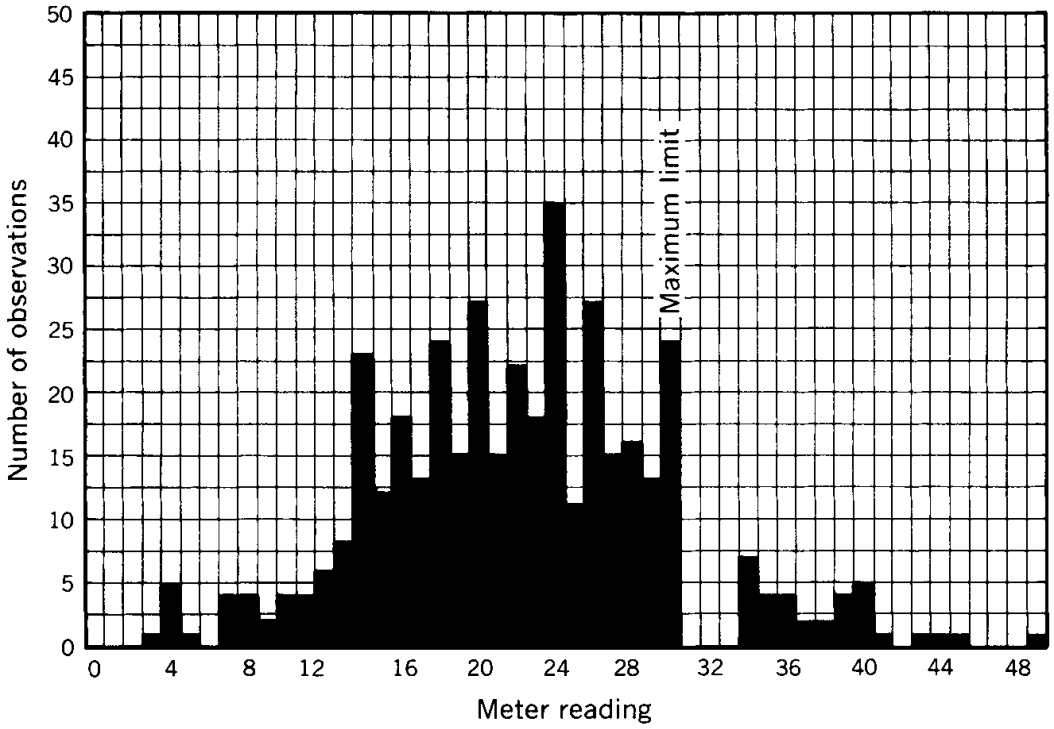

FIGURE 23.10 How inspector flinches at design limit.

studies show that when check inspection is introduced and check is made both of product rejected and of product accepted, the rejection of good product is reduced without affecting the acceptance of bad product.

Another form of flinching is to modify the inspection results to conform to the results that the inspector expects. For example, a visual inspection was being performed following a specified lapping operation. An experiment to omit the lapping operations, conducted without inspector awareness of omission of the operation, resulted in rejection of less than a third of unlapped product.

In some cases, flinching by inspectors is actually management-initiated through manager pressures that seem to the inspectors to leave no alternative. In one company, the inspectors making hardness tests were discovered to be flinching to an astonishing degree. This practice had been going on for years. It developed that the manufacturing vice president had designed this hardening process himself when he was the process engineer. At the time, he had deluded himself as to its capabilities and thereby had been the author of this long-standing practice (Juran, early consulting experience).

Rounding Off. The process of dispensing with unneeded accuracy is generally referred to as *rounding off.* Inspectors commonly round off their meter readings to the nearest scale division, as shown in Figure 23.11. The effect of rounding off is seen in the "picket fence" frequency distribution of Figure 23.10.

Rounding off is easy to detect from analysis of inspection data. A good analyst can, from the data alone, reconstruct the pattern of scale markings of an instrument without ever having seen the scale itself.

Rounding off is often a good thing, since it avoids undue attention to individual readings. Sometimes, however, the need for precision on individual readings is great enough that rounding off should not be practiced. The planner and inspection supervisor should be on the alert to identify situations in which rounding off is not tolerable, and they should provide accordingly.

Instruments and gages should be selected properly for the application. [Churchill (1956) gives a quantitative discussion on scale interval length and pointer clearance.] One practice is to require "readings to be recorded to the nearest . . . . "

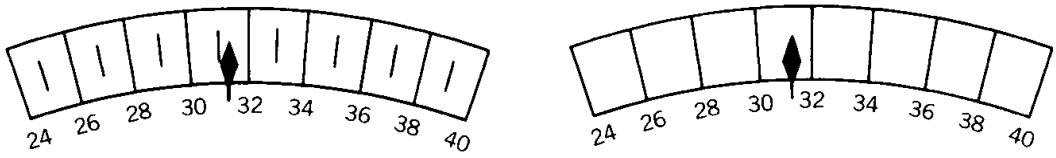

FIGURE 23.11 Rounding off.

Measure of Inspector and Test Accuracy. The collective effect of inspector errors, from all causes, is so extensive that there is a need for measuring the extent of errors and for use of the data in controlling the effectiveness of inspectors. If this measurement is made only occasionally, use can be made of standard sample arrays (see above, under Technique Errors) and cross-check among inspectors, as well as check inspection. If the measurement is to be conducted regularly so as to discover trends in performance, then check inspection is necessary.

In conventional check inspection, a second inspector, i.e., a check inspector, reviews the decisions of the inspector by reexamining the product after it has been inspected. For an early example, refer to Taylor (1911). The best practice is to reexamine the rejected product as well as the accepted product.

Inspection errors may consist of accepting defective units of product or rejecting good units of product. If, in addition, the check inspection reviews the procedure followed by the inspector, other errors may be found, e.g., use of wrong issue of the specification, wrong instrument, improper filling out of documents, etc.

The convenient use of check inspection data to quantify inspector accuracy is to count the errors, to assign weights, and to use the composite of errors as an index of accuracy (of inaccuracy, usually). [See, for example, Gilman (1963).] In some schemes, the errors discovered in later operations or in customer complaints are included in the data. The composite of errors may be expressed in terms of percentage defective (found to exist in the inspected product) or in terms of demerits per unit. [Refer to Weaver (1975), which reports on a study of inspector accuracy during the production process, based on accept/reject decisions involving the product currently produced.] Either way, the scoring system is open to the objection that the inspector's accuracy depends, to an important degree, on the quality of the product submitted to him or her by the process; i.e., the more defects submitted, the greater is the chance of missing some.

A plan for measuring inspectors' accuracy in a way that is independent of incoming quality is that evolved in 1928 by J. M. Juran and C. A. Melsheimer. [See Juran (1935) for the original published description of this plan.] Under this plan, the check inspector, as usual, reexamines the inspected product, both the accepted and the rejected units.

In addition, the check inspector secures the inspector's own data on the original makeup of the lot, i.e., total units, total good, total defective. From these data, the following formulas emerge as applied to a single lot, which has been check inspected:

$$\text{Accuracy of inspector} = \text{percent of defects correctly identified}$$

$$= \frac{d - k}{d - k + b}$$

where d = defects reported by the inspector
k = number of defects reported by the inspector but determined by the check inspector not to be defects
$d - k$ = true defects found by the inspector
b = defects missed by the inspector, as determined by check inspection
$d - k + b$ = true defects originally in the product

Figure 23.12 illustrates how the percentage of accuracy is determined. The number of defects reported by the inspector, d, is 45. Of these, 5 were found by the check inspector to be good; that is,

$k = 5$. Hence $d - k$ is 40, the true number of defects found by the inspector. However, the inspector missed 10 defects; that is, $b = 10$. Hence the original number of defects, $d - k + b$, is 50, that is, the 40 found by the inspector plus the 10 missed. Hence

$$\text{Percentage of accuracy} = \frac{d - k}{d - k + b} = \frac{45 - 5}{45 - 5 + 10} = 80\%$$

In application of the plan, periodic check inspection is made of the inspector's work. Data on *d*, *k*, and *b* are accumulated over a period of months to summarize the inspector's accuracy, as for example:

Job no.	Total pieces	*d*,	*b*	*k*
3	1000	10	0	0
19	50	3	1	0
42	150	5	1	0
48	5000	10	4	0
Total		200	30	0

The totals give, for percentage accuracy:

$$\frac{d}{d + b} = \frac{200}{230} = 87\%$$

As is evident, the plan lends itself to simple cumulation of data. However, some compromise is made with theory to avoid undue emphasis on any one lot checked. Over a 6-month period, where the cumulative checks may reach 50 or more, the need for such compromise or weighting is diminished.

The check inspector also makes errors. However, these have only a secondary effect on the inspector's accuracy. In the preceding example, if the check inspector were only 90 percent accurate, only 27 of the 30 defects missed by the inspector would be found. The inspector's accuracy would become

$$\frac{200}{227} = 88.1\% \quad \text{instead of} \quad 87.0\%$$

In some situations, *k* is small and may be ignored. However, in other situations, notably for sensory qualities, the inspector may have a bias for rejecting borderline work. In such cases, it is feasible to use, as an added measure, the inspector's accuracy due to rejection of good pieces. This has been termed *waste.* Under the terminology used here:

$$\text{Waste} = \text{percentage of good pieces rejected} = \frac{k}{n - d - b + k}$$

where *n* is the total pieces inspected.

The *percent accuracy* is also equal to the percentage of material correctly inspected. This feature permits use of the plan in the pay formula of the inspector.

Some investigators have developed variations on the foregoing measures of inspector accuracy as applied to visual inspection. These include measures based on probability theory (Wang 1975) and on use of signal-detection theory in the analysis of industrial inspection (Ainsworth 1980).

The proportion of correct decisions made by the inspector is an intuitively good index of the inspector's efficiency when the costs of rejecting a good item and accepting a bad item are equal.

The two measures for evaluating inspector's efficiency, introduced by Wang (1975) are

$$N_{\alpha} = \frac{\text{number of true defects detected by the inspector}}{\text{number of true defects}}$$

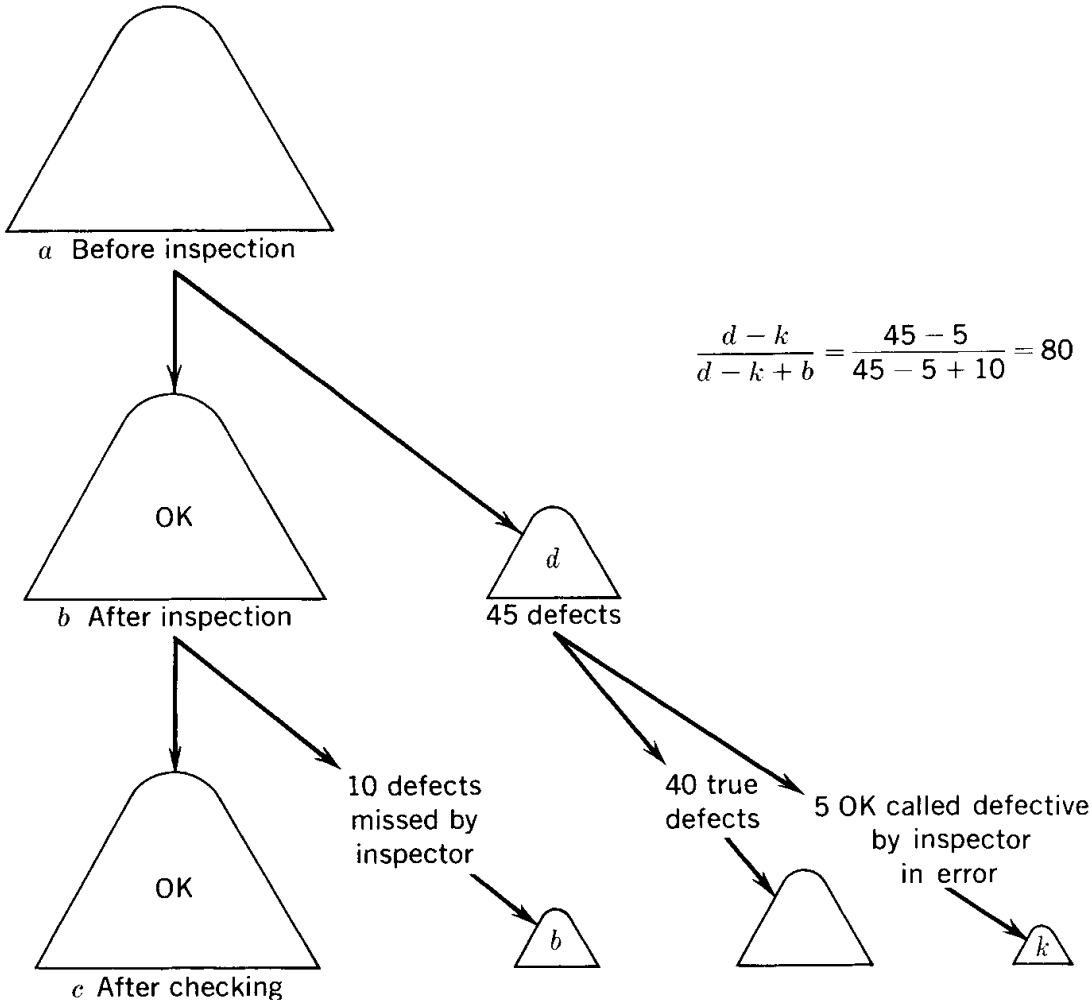

FIGURE 23.12 Process for determining accuracy of inspectors.

$$N_{\beta} = \frac{\text{number of true defects detected by the inspector}}{\text{number of all defects detected by the inspector}}$$

If the two costs are not equal, the theory of signal detectability (TSD) provides a better measure for analyzing an inspector's performance. However, the use of TSD requires assumptions about normal and equal variant population distributions, and the probability density functions for the "good" and "defective" populations must be calculated. See Johnson and Funke (1980) for a discussion of the advantages and disadvantages of a number of human performance measures, including Wang's (1975) approach.

In the application of any plan to check the accuracy of inspectors, it is essential that the checks be at random. Neither the inspector nor the check inspector should know the schedule in advance. Random dice, cards from a pack, etc. should be used. It is also essential that the responsibility be clear. The inspector who has accepted defects under orders or through inaccurate instruments, etc. cannot be held responsible for the results.

INSPECTION AND TESTING SOFTWARE

Training. As we move into the twenty-first century, we will be required to learn the concepts and application of computer-based technologies. Computer-based training has increasingly become a popular training tool for many organizations that do not require bilateral communications between

instructor and trainee and can provide alternate training schedules without having to rely on facility and instructor availability. Software for computer-based programs range from PC disks and interactive CD/ROMs to Internet links and cover topics for everything from gage calibration to 100 percent on-time delivery analysis (Kennedy 1996).

Statistical Process Control Interface. Software organizes collected data and tracks them to measure performance and standards in manufacturing processes. Many manufacturers have integrated the measurement system directly onto the machine to provide for the operator an immediate analysis of the process variation through charts and graphic data.

Data Collection. Software collects data from peripheral devices and retrieves data from the collection devices and downloads them onto a host computer. These data can then be analyzed for trends or problem solving.

Gage Calibration. Software is used for monitoring and recording of measurement device hardware accuracy, reportability, and gage performance. Features include gage type, gage due date, next calibration date, location, assignee, gage inspector (metrologist), and calibration history.

Simulation. Software is used for preproduction simulation to check for machine tool collisions, verify material removal of a machine tool, tool life analysis, inspection and test station layouts, and many other inspection and test interfaces into production processing. See "Software for Manufacturing" (1996).

REFERENCES

Ainsworth, L. (1980). "The Use of Signal-Detection Theory in the Analysis of Industrial Inspection." *Quality Assurance,* vol. 6, no. 3, September, pp. 63–68.

Alaimo, A. P. (1969). "A Total Inspection System." *ASQC Technical Conference Transactions,* Los Angeles, pp. 71–78.

Allen, P. E. (1959). "Evaluating Inspection Cost." *ASQC Technical Conference Transactions,* Los Angeles, pp. 585–596.

"An Outlook for 3D Vision." (1984). *Assembly Engineering, Newsline,* August, p. 6.

"ASD's Quality Assurance Program Rates in Top 10." (1984). *Assembly Engineering, Newsline,* August, p. 6.

Bader, M. E. (1980). "Quality Assurance and Quality Control, Part 1: Specifications." *Chemical Engineering,* vol. 83, no. 3, Feb. 11, pp. 92–96.

Baker, E. M. (1975). "Signal Detection Theory Analysis of Quality Control Inspector Performance." *Journal of Quality Technology,* vol. 7, no. 2, April, pp. 62–71.

Ballou, D. P. and Pazer, H. L. (1982). "The Impact of Inspector Fallibility on the Inspection Policy in Serial Production Systems." *Management Science,* vol. 28, no. 4, April, pp. 387–399.

Churchill, A. V. (1956). "The Effect of Scale Interval Length and Pointer Clearance on Speed and Accuracy of Interpolation." *Journal of Applied Psychology,* vol 40, December, pp. 358–361.

Cooper, J. E. (1980). "The Care and Training of Your Inspectors." *ASQC Technical Conference Transactions,* Milwaukee, pp. 674–676.

Crow, E. L., (1966). "Optimum Allocation of Calibration Errors." *Industrial Quality Control,* November, pp. 215–219.

Czaja, S. J., and Drury, C. G. (1981). "Training Programs for Inspectors." *Human Factors,* vol. 23, no. 4, pp. 473–484.

Denker, S. P. (1984). "Justifying Your Investment in Automatic Visual PCB Testing." *Circuits Manufacturing,* October, pp. 56, 58, 60, 62.

Dodds, L. B. (1967). "Workmanships." *ASQC Technical Conference Transactions,* Chicago, pp. 249–252.

Dodge, H. F. (1928). *A Method of Rating Manufactured Product.* Reprint B-315, May, Bell Telephone Laboratories.

Dodge, H. F., and Torrey, M. N. (1956). "A Check Inspection and Demerit Rating Plan." *Industrial Quality Control,* vol. 13, no. 1, July, pp. 5–12.

Eisenhart, C. (1968). "Expression of the Uncertainties of Final Results." *Science,* June 14, pp. 1201–1204.

Eppen, G. D., and Hurst, E. G., Jr. (1974). "Optimal Location of Inspection Stations in a Multistage Production Process." *Management Science,* vol. 20, no. 8, April, pp. 1194–1200.

Ezer, S. (1979). "Statistical Models for Proficiency Testing." *ASQC Technical Conference Transactions,* Houston, pp. 448–457.

Garfinkel, D., and Clodfelter, S. (1984). "Contract Inspection Comes Into Its Own." *American Machinist,* October, pp. 90–92.

Gebhardt, C. (1982). "Color Me Calibrated." *Quality,* March, pp. 62–63.Gibson, J. D. (1983). "How Do You Recognize a Qualified Inspector?" *Quality Assurance for the Offshore Industry* (*London*), April, pp. 55–56.

Gilman, J. R. (1963). "Quality Reports to Management." *Industrial Quality Control,* May, pp. 15–17. [In this demerit scheme of check inspecting the work of inspectors (who are paid by piece work), the accuracy is expressed in a form equivalent to the number of demerits found per lot checked.]

Gleick, Elizabeth (1996). "No Barrier to Mayhem," *Time,* July 29, p. 42.

Gunter, B. (1983). "The Fallacy of 100% Inspection." *ASQC Statistical Division Newsletter,* vol. 5, no. 1, September, pp. 1–2.

Hains, R. W. (1978). "Measurement of Subjective Variables." *ASQC Technical Conference Transactions,* Chicago, pp. 237–244.

Harris, D. H., and Chaney, F. B. (1969). *Human Factors in Quality Assurance.* John Wiley & Sons, New York, pp. 107–113.

Holmes, H. (1974). "Computer Assisted Inspection." *The Quality Engineer,* vol, 38, no. 9, September, pp. 211–213.

Johnson, S. L. and Funke, D. J. (1980). "An Analysis of Human Reliability Measure in Visual Inspection." *Journal of Quality Technology,* vol. 12, no. 2, April, pp. 71–74.

Juran, J. M. (1935). "Inspector's Errors in Quality Control." *Mechanical Engineering,* vol. 59, no. 10, October, pp. 643–644.

Juran, J. M. (1945). *Management of Inspection and Quality Control.* Harper & Brothers, New York.

Juran, J. M. (1952). "Is Your Product Too Fussy?" *Factory Management and Maintenance,* vol. 110, no 8, August, pp. 125–128.

Juran, J. M., and Gryna, F. M., Jr. (1980). *Quality Planning and Analysis from Product Development Through Use.* McGraw-Hill, New York, pp. 357, 360–361.

Juran, J. M., and Gryna, F. M., Jr. (1988). *Juran's Quality Control Handbook,* 4th ed. McGraw-Hill, New York.

Karabatsos, N. (1983). "Serving Quality." *Quality,* vol. 22, no. 9, September, pp. 65–68.

Ken, J. (1984). "The (Artificial) Eyes Have It." *Electronic Business,* Sept. 1, pp. 154–162.

Kennedy, M. S. (1996). "Inspection-Focused for the Future." *Quality in Manufacturing,* vol. 7, no. 4, May-June, pp. 22–23.

Konz, S., Peterson, G., and Joshi, A. (1981). "Reducing Inspector Errors." *Quality Progress,* vol. 14, no. 7, July, pp. 24–26.

Kramer, Amihud. (1952). "The Problem of Developing Grades and Standards of Quality." *Food Drug Cosmetic Law Journal,* January, pp. 23–30.

Kusch, J. (1979). "Robots and Their Advantage in Inspection." *Proceedings, Society of Photo-Optical Instrumentation Engineers,* vol. 170: *Optics in Quality Assurance,* vol. II, pp. 40–42.

Leek, J. W. (1975). "See It as It Really Is." *ASQC Technical Conference Transactions,* San Diego, pp. 41–43.

Leek, J. W. (1976). "Benefits from Visual Standards." *Quality Progress,* December, pp. 16–18.

Lincoln, M. (1996). "Automated Inspection Boots Productivity, Quality," *Quality in Manufacturing,* vol. 7, no. 5, July/Aug., p. 18.

Linn, R. D. (1981). "Computer-Aided Inspection—Its Time Has Come." *ASQC Quality Congress Transactions,* San Francisco, pp. 599–602.

Mathur, C. P. (1974). "The Influence of Measurement Variation in Mass Production of Precision Parts." *Q.R. Journal,* January, pp. 1–5.

McCaslin, J. A., and Gruska, G. F. (1976). "Analysis of Attribute Gage Systems." *ASQC Technical Conference Transactions,* Toronto, pp. 392–400.

Meckley, D. G., III (1955). "How to Set Up a Gaging Policy and Procedure." *American Machinist,* vol. 99, no. 6, March 14, p. 133.

Megaw, E. D. (1978). "Eye Movements in Visual Inspection Tasks." *Quality Assurance,* vol. 4, no. 4, December, pp. 121–125.

Miller, E. M. (1975). "Test Methods and Specification Requirements." *ASQC Technical Conference Transactions,* San Diego, pp. 229–230.

Montville, V. L. (1983). "Color Control from Start to Finish." *Quality,* March, pp. 36–38.

National Bureau of Standards (1965). "Expression of the Uncertainties of Final Results." Chapter 23 in *Experimental Statistics,* NBS Handbook 91. U.S. Goverment Printing Office, Washington.

Nelson, A. V. (1984). "Machine Vision: Tomorrow's Inspections with Today's Equipment." *Evaluation Engineer,* vol. 23, no. 9, October, pp. 21, 24, 26, 29, 30, 34, 37.

Nygaard, G. M. (1981). "Why 100 Percent Inspection?" *Quality,* October, pp. 38–39.

Ohta, H., and Kase, S. (1980). "Evaluation of Inspectors in Sensory Tests—Qualification by Geometrical Methods and Classification by Bayesian Diagnosis Rule." *Journal of Quality Technology,* vol. 12, no. 1, January, pp. 19–24.

Palumbo, R. (1997). "Gage Calibration Is an Investment in Quality Production." *Quality Digest,* vol. 17, no. 2, February, p. 60.

Papadopoulos, N. (1983). "Instrumental Color Control—Where to Start." *Quality,* December, pp. 44–46.

"Portable Gage Lab Provides Flexibility and Minimizes Downtime." (1971). *Quality Management and Engineering,* November, p. 23.

Quinlan, J. C. (1996). "Automated Gaging, Test, and Inspection." *Quality in Manufacturing,* vol. 7, no. 3, April, pp. 8–9.

Riley, F. D. (1979). "Visual Inspection—Time and Distance Method." *ASQC Technical Conference Transactions,* Houston, pp. 483–490.

Sataro, Pat (1997). American Red Cross Blood Service, Farmington, Conn., in private correspondence with the editors.

Schaffer, G. (1984). "Machine Vision: A Sense for CIM." *American Machinist,* Special Report 767, June, pp. 101–120.

Schweber, W. (1982). "Programming for Control—Computerizing Measurement and Control." *Quality,* March, pp. 38–42.

Shainin, D. (1972). "Unusual Practices for Defect Control." *Quality Management and Engineering,* February, pp. 8, 9, 30.

Smith, J. R., and Duvier, H., III (1984). "Effect of Inspector Error on Inspection Strategies." *ASQC Quality Congress Transactions,* Chicago, pp. 146–151.

"Software for Manufacturing." (1996). *Quality in Manufacturing,* vol. 7, no. 3, April, p. 36.

Spow, E. E. (1984). "Automatic Assembly." *Tooling and Production,* October, pp. 46–47.

Strack, Charles M. (1996). *Scientific Technologies,* vol. 2, no. 3, November.

"Tape Controlled Machines at Sunstrand Aviation." (1970). *Quality Assurance,* June, pp. 30–34.

Tawara, N. (1980). "A Case Study on Measuring Inspection Performance for Inspection Job Design." *International Journal of Production Research,* vol. 18, no. 3, May-June, pp. 343–353.

Taylor, F. W. (1911). *Principles of Scientific Management.* Harper & Brothers, New York, pp. 80–96.

"The Vision Thing." (1996). *Quality in Manufacturing,* vol. 7, no. 4, May-June, p. 10.

Thompson, H. A., and Reynolds, E. A. (1964). "Inspection and Testing as a Problem in Man-Machine Systems Control Engineering." *Industrial Quality Control,* July, pp. 21–23.

Tobey, D. (1979). "Metrology = Calibration." *ASQC Technical Conference Transactions,* Houston, pp. 513–520.

Trippi, R. R. (1975). "The Warehouse Location Formulation as a Special Type of Inspection Problem." *Management Science,* vol. 21, no. 9, May, pp. 986–988.Walsh, L. (1974). "Back to One Hundred Percent Inspection??" Editorial comment. *Quality Management and Engineering,* March, p. 9.

Walsh, L., et al. (1976). "100% Inspection at Production Line Rates." *Quality,* November, pp. 30–31.

Walsh, L., et al. (1978). "Steel Wheel Maker Tests Hardness 100%" *Quality,* November, p. 37.

Walsh, L., et al. (1979*a*). "Can 100% Testing Be Eliminated?" *Quality,* May, pp. 42–43.

Walsh, L., et al. (1979*b*). "100% Inspection Plus." *Quality,* September, pp. 102, 104.

Walsmann, M. R. (1981). "The Check List—A Powerful Inspection Tool." *ASQC Technical Conference Transactions,* pp. 348–351.

Wambach, G. W., and Raymond, A. S. (1977). "The Optimum Sampling Plan." *ASQC Technical Conference Transactions,* Philadelphia, pp. 574–578.

Wang, S. C. (1975). "Human Reliability in Visual Inspection." *Quality,* September, pp. 24–25.

Wang, S. H. S., and Seppanen, M. S. (1981). "A System Simulation for Inspection Planning." *ASQC Quality Congress Transactions,* San Francisco, pp. 769–776.

Weaver, L. A. (1975). "Inspection Accuracy Sampling Plans." *ASQC Technical Conference Transactions,* San Diego, pp. 34–39.

Woods, D. G., and Zeiss, C. (1981). "Coordinate Measuring and Finite Metrology." *ASQC Quality Congress Transactions,* San Francisco, pp. 232–237.

Woods, K. C. (1978). "Calibration System for Measuring and Testing Equipment." *Quality Progress,* March, pp. 20–21.