NOVATEUR PUBLICATIONS
INTERNATIONAL JOURNAL OF INNOVATIONS IN ENGINEERING RESEARCH AND TECHNOLOGY [IJIERT]
ISSN: 2394-3696
VOLUME 2, ISSUE 6 JUNE-2015

A study of the Quality Inspection Task

Prof. Vyankatesh S. Kulkarni
Department of Mechanical Engineering, Solapur University /BIT/Barshi/India

Abstract

The paper represents the evolution of quality inspection tasks in industries, and briefly discusses the types of tasks and mode of performance. The author's intention is to review the literature available in quality inspection tasks. And also discuss the advantages and limitations of these tasks. The paper is an extended version of "Advanced Manufacturing Systems: A Review" authored by same author, where discussions of AMS technologies has been carried out.

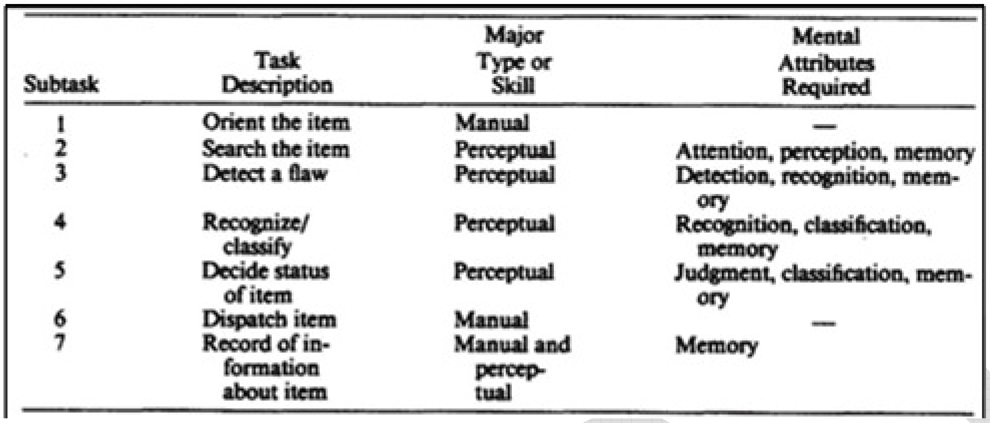
Introduction

Quality has been defined as fitness for use, or the extent to which a product meets the consumer requirements (Juran and Gryna, 1980). Inspection is the act of measuring or examining carefully the quality of a product. Sensory inspections (performed by means of the human senses to assess a product's qualitative characteristics) and physical inspections (performed by means of measuring devices to assess a product's quantitative characteristics) are the main types of quality inspections. It has been widely accepted that the quality inspection task performed by humans is prone to error (Bennet, 1975; Konz, Peterson, and Joshi, 1981; Schilling, 1982). In fact, Juran (1974) indicated that human inspectors typically find about 80% of the defects. This recognized level of inspection error has been mostly attributed to the sensory aspects of the inspection. Shingo (1985) indicated that it tends to be difficult to set criterions for sensory inspection because different people will make different judgments and even the same person might make different judgments on different days. Efforts to automate the quality inspection task have not been as successful with sensory inspections as with various physical inspections (Drury, 1992b). During the last four decades a substantial amount of human factors research has been conducted on the visual quality inspection task. The human operator's performance on the visual quality inspection task has been extensively studied from the signal detection theory (SDT) standpoint. In addition, the machine and human-plus-machine system's performance on the visual quality inspection task has been studied and compared to the human operator's performance.

Visual Quality Inspection Taxonomy

The visual quality inspection task has been described as consisting of the following subtasks (Wang and Drury, 1989): 1) orient the item, 2) search the item, 3) detect any defect, 4) recognize/classify the defect, 5) decide the status of the item, 6) dispatch the item, and 7) record the information about the item. Table 2 shows the human skills required should all the subtasks be assigned to human operators (Wang and Drury, 1989). Drury (1992) indicated that these subtasks can be combined into two main components: search and decision making. Thus, the simplest description of the visual quality inspection task is to search, recognize a defect, and make a decision on the part's acceptability within the quality limits.

Table 1. Human skills required for the quality inspection task (Wang and Drury, 1989).



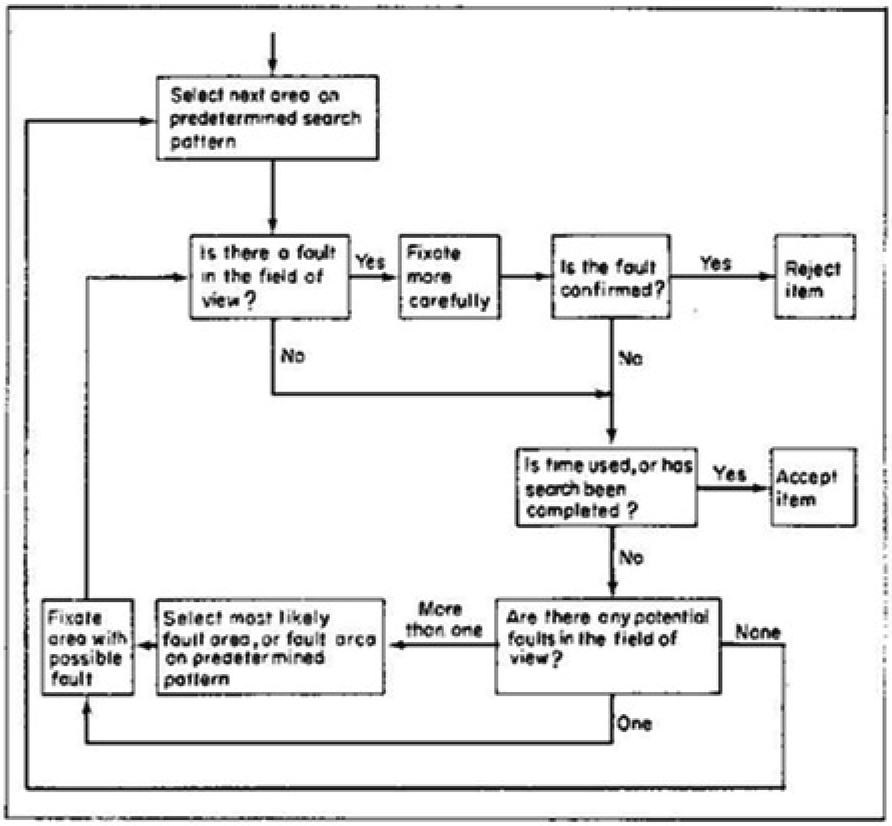
Subtask	Task Description	Major Type or Skill	Mental Attributes Required
1	Orient the item	Manual	—
2	Search the item	Perceptual	Attention, perception, memory
3	Detect a flaw	Perceptual	Detection, recognition, memory
4	Recognize/ classify	Perceptual	Recognition, classification, memory
5	Decide status of item	Perceptual	Judgment, classification, memory
6	Dispatch item	Manual	—
7	Record of information about item	Manual and perceptual	Memory

Visual search components of the quality inspection task

Visual search is a sequential process that proceeds as a series of fixations linked by eye movements and which terminates upon successful detection of a defect or the complete inspection of the unit (Drury and Prabhu, 1994). It has been indicated that almost all of the information in a visual search is obtained during the fixations which account for more than 90% of the search time (Morawski, Drury, and Karwan, 1980). In the inspector's field of view, a defect is only visible within a limited area referred to as the visual lobe. During a fixation the visual lobe is located around the central fixation point. According to Drury (1992), the visual lobe size will be affected by the luminance of the object inspected, the contrast between the object and the defect on the object, the defect size, and the distance of the defect from the inspector's eyes. Figure 2 shows a flow diagram for visual search (Howarth and Bloomfield's, 1971).

Megaw and Richardson (1979) conducted eye movement studies of inspectors and concluded that inspectors do not follow a simple pattern in searching an object. They observed that while a very random appearing search pattern was used for the inspection of complex units (e.g., circuit boards), a more systematic search pattern was used for the inspection of simpler ones (e.g., knitwear). Bloomfield (1975) suggested that the most efficient strategy will occur when: 1) the distance that the eyes move from one fixation to the next is short enough that no point on the unit is missed, and 2) the distance between fixation points is large enough to minimize the overlap between the areas in which the defect could be seen. However, Bloomfield acknowledged that when the search is for more than one type of defect, with different levels of discriminability, the most efficient strategy for one type of defect might not necessarily be efficient for the other type of defect. Drury (1992) indicated that in addition to the lobe size and the search strategy, the time available for the inspection will affect human performance in the visual search component of the inspection. The more time the inspector has to search, the better the chances are of finding the defect. Nevertheless, Schoonard, Gould, and Miller (1973) found that the best inspectors for the visual search subtask were those who detected the defect in the fewest fixations, not those with more rapid fixations. Three factors that led to a thorough investigation of search automation were: 1) the significant time consumption of the visual search, 2) the need for the selection and implementation of an optimal search strategy, and 3) the quality at the source and 100% inspection strategies strongly recommended for AMS (Drury and Prabhu, 1994; Shingo, 1985).

Figure 1. Flow diagram for visual search (Howarth and Bloomfield, 1971).



Decision making component of the quality inspection task

Given its strict relevance to decision making, signal detection theory (SDT) has been used to explain the decision making component of the quality control inspection task (Wallack and Adams, 1969; Drury and Fox, 1975). In the fundamental signal detection problem, as defined by Swets, Tanner and Birdsall (1964), an observation is made of events occurring in a fixed interval of time, and a decision is made whether the interval contained only the always present noise (N) or the signal plus noise (SN). As opposed to classic methods of psychophysics, SDT provides both an independent quantitative measure of the criterion (β) that the observer uses in making a perceptual judgment, and a relatively pure measure of sensitivity (d'). The major assumptions of SDT are: 1) the observations (sensory data) on which the decision is based may arise from either conforming or nonconforming items, 2) the observations may be represented as varying continuously along a single dimension forming two probability density functions (N and SN), and 3) both probability density functions can be described by two normal distribution with equal variances. The criterion (ß) or policy to determine whether the observation results from the N or the SN distribution is established by a cutoff value X_c on the continuum of observations. In a quality inspection context, as described by Drury and Fox (1975), SDT proposes that the human, functioning as a defect detection device, builds up in the neural system two distributions of activity: one relating to the probability of accepting an unit, the other to the probability of rejecting it. The degree of separation of these two distributions' means is a measure of the inspector's discriminability of the defects (d′). The criterion level (β), which is the ratio of the two ordinates of the curves at a given level X_c , delineates the boundary between accepting and rejecting a unit, and in doing so takes in some good units to be rejected and some faulty units to be accepted. Inspectors make a correct decision either by accepting a good unit (correct rejection) or by rejecting an unacceptable unit (hit). They fail either by not detecting a rejectable defect (miss) or by falsely reporting the presence of a reject able defect (false alarm). A theoretical SDT distribution is shown in Figure 3 (Wickens, 1992).

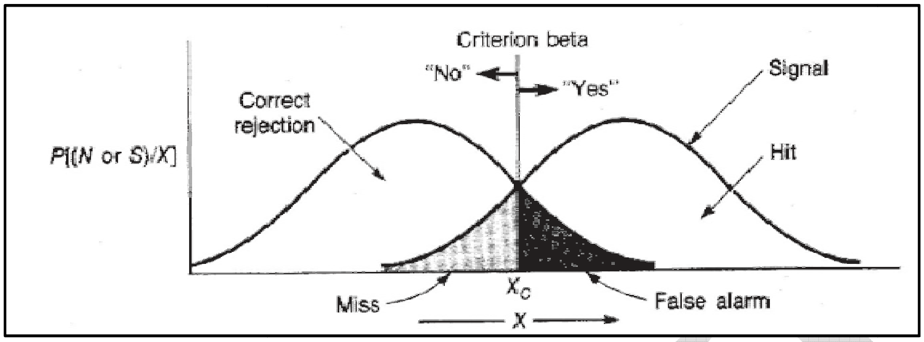


Figure 3. Theoretical SDT distribution (Wickens, 1992).

Both decision making performance measures (d′ and ß) are derived from the hit rate and the false alarm rate. Drury (1992) indicated that the pure decision making component of the quality inspection task can be measured by concentrating on tasks that require no search. The general conclusion of quality inspection studies reviewed by Drury and Fox (1975b) is that "the decision making component is among those rare tasks where a human being behaves like a rational economic decision maker, balancing the costs and payoffs involved to arrive at an optimum performance." As a normative model, SDT defines the optimal criterion (ß$_{\text{opt}}$) used by the ideal observer to optimize economic gains (Green and Swets, 1966). Based upon the values of a hit and a correct rejection, and the costs of a miss and a false alarm, the ß$_{\text{obs}}$ of a rational observer can be calculated and compared to that of the theoretical ideal observer (ß$_{\text{opt}}$). Gescheider (1985) reported that after comparing ß$_{\text{obs}}$ with ß$_{\text{opt}}$, it has been generally found that rational observers do fairly well at optimizing their winnings. However, they tend not to set extremely low or high criterion, even in situations where these strategies would lead to optimal performance.

SDT Studies on Visual Quality Inspection

Signal Detection Theory (SDT) was first used to model the decision making performance of the quality inspection task by Wallack and Adams in 1969 (Thapa, Gramopadhye, and Melloy, 1996). After using SDT to study the performance of industrial electronics inspectors in a visual, subject-paced task, they concluded that SDT performance measures (d′ and ß) were more useful than the other available measures. Wallack and Adams (1969) concluded that SDT is useful because, in addition to relating performance to payoff, it also indicates the magnitude and the direction of improvement required. Although not all of the research using SDT to study the decision making performance has been conducted in an industrial inspection context, its findings have been beneficial in understanding human quality inspector performance. More recently another measure of criterion level or decision making response bias called the index *C* has been developed (Macmillan and Creel man, 1991). The main difference between ß and the index *C* is in the way these bias indices locate the criterion (X_c). The likelihood ratio measure, ß, locates X_c by the ratio of the heights of the SDT distributions (N and SN), while the index *C* locates X_c by its distance from the intersection of the two distributions. The range of c is therefore the same as that of d′, although zero is at the center rather than an endpoint. This parametric index is considered to be more effective than ß over a full range of sensitivity in recognition memory experiments and in vigilance (Snodgrass and Corwin, 1988; See, Warm, Dember, and Howe, 1997). Unlike the index *C*, ß has a tendency to produce extremely high values for conservative observers resulting in a scale that does not produce equal intervals. See et al. (1997) conducted three experiments involving factors that affect response bias (signal probability, payoff matrix, and signal salience) in order to calculate and compare ß and *C*. All three experiments employed a 40 minute task without performance feedback in which the subjects were required to detect increments of 3 mm in the height of a single 32 x 4 mm white line. The white line flashed on and off at a rate of 20 events/minute in the center of a computer screen. The researchers found that the index *C* was generally more sensitive than ß to the no perceptual manipulations of signal probability, payoff, and probability shifts. Also, they reported that ß was less effective than *C* at differentiating variations in response bias tendencies when sensitivity approached chance. The computational formulas for d′, ß, and the index *C* are presented in particular the normal distributions with equal variance assumption, are not always true in every inspection situation. Swets (1977) indicated that this assumption is only necessary for the calculation of d′. From a theoretical perspective Green and Swets (1966) identified the Central Limit Theorem as the reason for the normal distribution assumption. From a practical and empirical perspective the reason for SDT assumptions is that when checks are made on the results of SDT studies on quality inspection, they seem to be in basic agreement with predictions from the theory (Chapman and Sinclair, 1975; Drury and Fox, 1975). Decision making performance changes over time. Based on the review of twelve studies conducted between 1969 and 1975, Swets (1977) indicated that all twelve experiments showed an increasingly strict criterion (ß) over time when the signal-to-noise ratio was low; eight of the experiments showed a constant sensitivity (d′) over time. In the four studies in which d′ did not remain constant it was found to decrease by 20%. An increase in ß (conservative criterion) represents a decrease in signals detected as well as a decrease in false alarm errors. As Williges (1969) indicated, a performance change over time is characterized by a shift in the subject's response criterion (ß). Williges (1969) conducted a study in which the subjects were required to detect a signal (the longer of two fixed-duration changes from a standard brightness) during a 60 minute monitoring session. A distraction element was introduced by asking the subject to time share a secondary task (four numerical processing operations) with the primary monitoring task. The events in both tasks were occurring at intervals of 10 seconds. Two signal-to-noise ratios (1/5 and 5/1), a correct and an incorrect set of instructions (expected number of signals in the sample), and the distraction vs. no distraction element were combined factorially to measure the hit and false alarm rates. The ANOVA conducted for the performance measures (d′ and ß) showed the following results: 1) d′ remained constant over time and 2) ß increased over time when the signal-to-noise ratio was low. The same results were obtained by Colquhoun and Edwards (1970). They conducted a study in which the subjects were required to indicate which one of six disks, if any, was larger than the others. They found d′ to be constant and ß to increase over time. On the other hand, Guralnick's (1972) results differed from those obtained by Williges (1969). He conducted an experiment in which the signal was the longest of a pair of vertical lines, and the events were presented at a rate of eight per minute. Although his results showed a ß increase over time, d′ was found to decrease by 20%.

Determinants of decision making performance

A study of on-line quality inspection for mass-produced jam tarts was conducted in the baking industry by Chapman and Sinclair (1975). They tested the hypothesis that speed of movement of the tarts on the conveyor and the presentation time of the tarts had no effect upon the performance of the inspectors. The two levels of the tarts movement factor (moving and stationary) and the three levels of the presentation time (22.5 sec., 15 sec., and 7.5 sec.) were combined factorially to measure the hit and false alarm rates. The ANOVA conducted for the performance measures (d′ and ß) showed the following results: 1) d′ decreased as the presentation time was reduced, and 2) ß increased as the presentation time was reduced. The authors acknowledged that since slowing down the production line was not an option the only possible solution was to double the number of inspectors. The outcome of this research suggested the need for conducting the assessment of quality inspection agents' capabilities early during the design stage of the particular manufacturing system in order to avoid solutions incompatible with world class AMS. Smith and Barany (1970) conducted an experiment in which the varied inspection task variables were task pace, percentage of defective, and payoff matrices (consequences of the decision made). The experimental task was the on-line inspection of aluminum disks (1 inch in diameter) containing four equally spaced holes (1/16 inch in diameter). The subjects inspected the product units and indicated a defective (three holes instead of four) by pressing a micro switch. It was found that d′ decreased as the task pace increased and remained invariant with respect to the other two independent variables. The decision criterion (ß) utilized by the inspectors increased (became more stringent) as the percentage of defective increased and remained invariant with respect to the other two independent variables. The researchers attributed the unexpected result (conservative ß when signal probability was high) to the failure of the inspectors to adopt the payoff system as instructed. Williges (1971) conducted a study in which the subjects were required to detect a signal (the longer of two fixed-duration changes from a standard brightness) during a 60 minute monitoring session. The events were presented on a 6 second time base. Two signal-to-noise ratios (1/9 and 1/1) and three payoffs (lax, neutral, and strict) were the varied inspection task variables. Based on the results of the study he concluded that the signal probability rather than payoffs seem to play the major role in determining inspection performance. When the signal probability is high the observers maintain a “lax” response criterion, but, when the probability is low the observers adopt a “conservative” one. The effect of knowledge of results (KR) on decision making performance is beneficial but complex. Craig and Colquhoun (1975), referencing the work of Mack worth (1964, 1965), indicated that the performance increase caused by false KR suggests that at least some part of the effect of providing true KR may be due to an initial increase in arousal or alertness. Williges and North (1972) conducted a 2^3 factorial experiment in which the presence or absence of correct detection KR, false alarm KR, and cumulative total event KR were randomly assigned to subjects. The experimental task was the same one used in Williges’ previous SDT studies (1969, 1971). They concluded that: 1) the specificity of KR influences d’ more than ß, 2) positive KR on the correctness of a detection response is used by the observer to increase their d’, 3) observers use signal-to-noise ratio KR to manipulate their ß, 4) false alarm KR results in fewer detections overall when cumulative total events KR are provided, and 5) subjects appear to use the cumulative total events KR to guide their frequency of responding. The results of a study on the effect of information on industrial inspection performance conducted by Zunzanyika and Drury (1975) are consistent with some of Williges and North (1972) findings. In their experiment, the independent variables were the percentage of defective (10%, 20%, and 30%) and the method of transmitting information about the batch quality (feedback, feedforward, and both). Based on the analysis of the results, the authors stated that the percentage of defective KR had a significant effect on d’ and a more obvious effect on ß. Various researchers argue that the traditional laboratory findings about vigilance task performance are not particularly relevant to real operational problems, partly because real world tasks are of a more complex nature than the simple tasks which have been used in the laboratory studies (Alluisi, Coates, and Morgan, 1977; Craig and Colquhoun, 1977; Craig, 1979). As an example of a complex vigilance task they mentioned the industrial quality inspection that frequently requires the operator to look for the occurrence of several kinds of signals. Craig and Colquhoun (1977) conducted an experiment in which two different type of defects (with equal probability of 10% each) were presented to the subjects. Each defect consisted of a circle with a single radial spoke. One of the defects needed to be rejected whenever a change in the angle of the radial spoke was observed. The second defect needed to be rejected whenever a change in diameter was observed. The defects were presented on a television monitor. A control group only inspected one type of defects, while a second group inspected both types simultaneously. Based on the results of this experiment the researchers concluded that the effect of what they called complex inspection (inspection of both defects simultaneously) on the overall level of detections was not significant. Craig (1979) conducted the same type of experiment, but used unequal probabilities (5% and 15%) for the two types of defects, and reached the same conclusion. He indicated that these research findings (Craig and Colquhoun, 1977; Craig, 1979) do not invalidate the traditional laboratory findings about vigilance task performance, but rather suggest the value of further research on this issue.

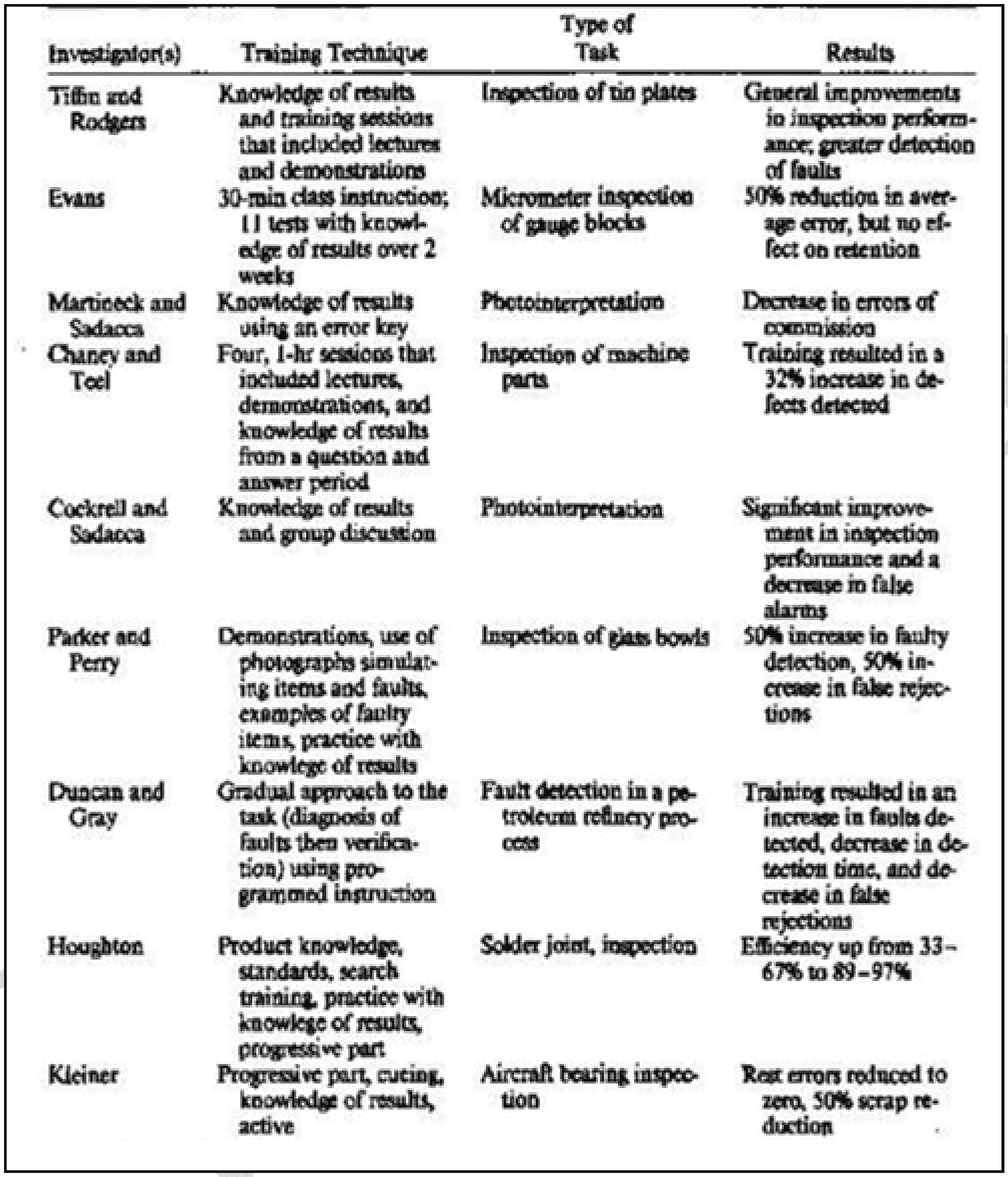
Quality Inspection Training

Training has been identified as essential to improving the decision making performance of human operators (Embrey, 1975; Wiener, 1975). Juran and Gryna (1980) indicated that chronic deficiency in knowledge (e.g., education or training) is a cause of errors consistently made by inspectors. Czaja and Drury (1981) claimed that despite the existence of a profound knowledge base on training, it is a neglected area for improvement in industrial inspection performance. They indicated that KR is fundamental to learning. Thapa, Gramopadhye, and Melloy (1996) mentioned that the results of the KR transmission strategies (feedback, feedforward, or both) used in training programs are consistent with the research findings of the effect of KR on decision making performance. Drury (1992) listed five techniques that have proven effective in training for inspection: 1) Cueing, 2) Feedback, 3) Active training, 4) Progressive parts, and 5) Develop schema. Table 3 summarizes various applications of industrial inspection training programs (Drury, 1992).

Hybrid Inspection Systems

Zero-defect products and shorter lead-time production are vital for the survival and success of AMS in a highly competitive WCM. According to Drury and Sinclair (1983), “this can often be achieved only by 100% inspection, which is known to be unreliable when performed by humans.” This dilemma prompted an industry movement towards automated inspection systems. Various automated inspection systems are listed in Table 4 (Drury and Prabhu, 1994). The advent of these microprocessor-based automated inspection devices, at prices competitive with human inspection, called for a human factors reassessment of the human-machine function allocation possibilities in quality control (Drury and Sinclair, 1983). Gramopadhye, Drury, Sharit, and Sudit (1992) proposed a framework for function allocation in inspection (see Figure 4). They recommended accuracy, speed, flexibility, and reliability as the performance criterion for the inspection system.

Table 2. Applications of industrial inspection training programs (Drury, 1992)



Investigator(s)	Training Technique	Type of Task	Results
Tiffin and Rodgers	Knowledge of results and training sessions that included lectures and demonstrations	Inspection of tin plates	General improvements in inspection performance; greater detection of faults
Evans	30-min class instruction; 11 tests with knowledge of results over 2 weeks	Micrometer inspection of gauge blocks	50% reduction in average error, but no effect on retention
Martineck and Sadacca	Knowledge of results using an error key	Photointerpretation	Decrease in errors of commission
Chaney and Teel	Four, 1-hr sessions that included lectures, demonstrations, and knowledge of results from a question and answer period	Inspection of machine parts	Training resulted in a 32% increase in defects detected
Cockrell and Sadacca	Knowledge of results and group discussion	Photointerpretation	Significant improvement in inspection performance and a decrease in false alarms
Parker and Perry	Demonstrations, use of photographs simulating items and faults, examples of faulty items, practice with knowlege of results	Inspection of glass bowls	50% increase in faulty detection, 50% increase in false rejections
Duncan and Gray	Gradual approach to the task (diagnosis of faults then verification) using programmed instruction	Fault detection in a petroleum refinery process	Training resulted in an increase in faults detected, decrease in detection time, and decrease in false rejections
Houghton	Product knowledge, standards, search training, practice with knowlege of results, progressive part	Solder joint, inspection	Efficiency up from 33–67% to 89–97%
Kleiner	Progressive part, cueing, knowledge of results, active	Aircraft bearing inspection	Rest errors reduced to zero, 50% scrap reduction

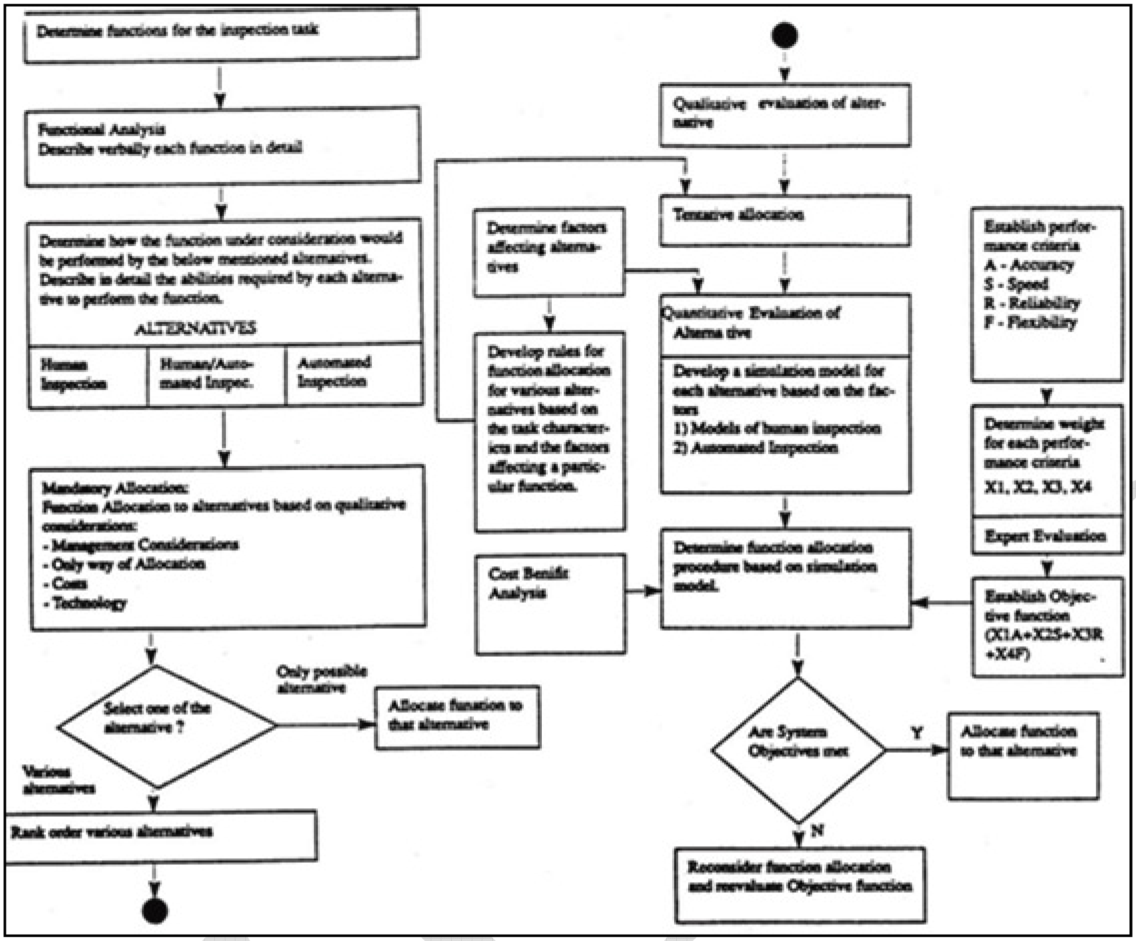


Figure 4. Framework for function allocation in inspection (Gramopadhye, Drury, Sharit, and Sudit, 1992).

One of the first efforts to reassess the human-machine function allocation possibilities in quality control was conducted by Drury and Sinclair (1983). They compared the performance of experienced inspectors and a prototype optical/microprocessor inspection device in an inspection task of small steel cylinders with four different possible defect types. The comparison was done in terms of both hit rate and false alarm rate. These SDT metrics were plotted one against the other to form a Receiver-Operating Characteristic (ROC) curve. Drury and Sinclair's (1983) main findings were:

1) Neither the human nor the automated systems achieved an outstanding performance and

2) The automated system was better at locating the defects (search) but could not classify them as acceptable or reject able (decision making) as well as the human inspectors.

Hou, Lin, and Drury (1993) conducted a 3^3 factorial experiment in which the experimental task was the quality inspection of printed circuit boards containing surface mounted devices. The independent variables were the complexity of the product (number of components in the board), the contrast level between components and background, and the visual noise in the displayed image (the noise coming from the lighting conditions and the imaging system). Using the accuracy and speed performance criterion measures combined with the false alarm rate and the hit rate, they calculated a cost-based evaluation function. This cost-based evaluation function coupled with the Drury and Sinclair (1983) findings led to the conclusion that allocating the search function tomachines and the decision making function to humans results in better performance than pure human or pure machine inspection. This computer-search/human-decision making system is known as a Hybrid Inspection System (HIS). The idea behind HIS is to capitalize on the machine speed and precision to scan the inspection unit, and on the decision making ability exhibited by humans.

Quality Inspection Task Load

Quality inspection tasks that impose a sustained load on working memory (to recall what the quality acceptability criterion looks like) will demand the continuous supply of processing resources (Parasuraman, 1979; Parasuraman, Warm, and Dember, 1987, Wickens, 1992). Parasuraman (1979) conducted an experiment using a successive-discrimination task (that imposed a memory load) in which the signal was specified as the decrease in the intensity of a flashing light. The signals were presented irregularly at a mean rate of two signals per minute, and the event rate was 30 events per minute. The duration of the task was 45 minutes. He concluded that the performance in such areas of vigilance application as radar monitoring and industrial quality inspection could be adversely affected when the operator has to discriminate a signal from a standard represented in memory and when the event rate is high. This performance decrement may result either from signal-data limits (weak signal in noise), or memory-data limits (quality of stored representation of the standard in delayed comparison memory tasks). Like many other tasks, quality inspection has been identified as having an inverted-U shaped relationship between task demand and performance level (McGrath, 1965, Wiener, Curry, and Faustina, 1984). The Inverted-U theory states that for a given task, there is an optimal level of workload or demand that yields the highest level of performance. A departure in either direction from the optimal level of work will result in a performance decrement. Wiener (1975) indicated that while most of the results of vigilance research support the right-hand side of the inverted-U theory (overload) there is a lack of support for the left-hand branch (assertion that the task performance level can be improved by increasing the load). One of the first experiments that supported the left- Hand side in a vigilance task was conducted by McGrath (1965).

After comparing easy and hard visual monitoring tasks conducted concurrently he concluded that the presence of the hard task facilitated performance on the easy one. Wiener et al. (1984) conducted an experiment in which a control group performed a vigilance task (the signal was the decrease in distance between two dots presented on a computer screen), and a second group performed a one-dimensional compensatory tracking task in addition to the vigilance task. They found that performance of the second group (vigilance and tracking tasks) in terms of signal detection exceeded the performance of the control group (vigilance task only). They concluded that these research results provided support for the facilitating effect of increasing the task load (left side of the inverted-U). Some researchers describe the quality inspection task as being intrinsically boring (Craig, 1984; Poulton, 1977). According to them, this explains why it is often the case that mild stress will increase the performance in terms of detection and response time. However, Wickens (1992) has indicated that vigilance tasks with working memory load are susceptible to interference from concurrent tasks.

Conclusion

The paper has discussed quality inspection methods to complete tasks in industries. The paper is useful for the industry as well as academic and professional practitioners as a guide.

References:

1. Karwowski, Salvendy, Bad ham, Brodner, Clegg, Hwang, Iwasawa, Kidd, Kobayashi, Koubek, LaMarsh, Nagamachi, Naniwada, Salzman, Seppala, Schallock, Sheridan, and Warschat, 1994
2. Human flexibility and adaptability (Salvendy, 1992; Pinochet, et al., 1996
3. Social communication and interaction (Corbett, 1987)
4. the attempt to eliminate the gaps or deficiencies in machine designs by means other than skilled human operators Unterweger (1988)
5. Factory of the future." Davies (1986)
6. "Implicit in WCM developments involving AMT and JIT and similar principles of manufacturing organization Wilson (1992)"