TEAM EFFICIENCIES ON A PACED VISUAL INSPECTION TASK

by

KHALID OSMAN

B. Tech., Indian Institute of Technology, 1972
Bombay, India

A MASTER'S THESIS

submitted in partial fulfillment of the
requirement for the degree

MASTER OF SCIENCE

Department of Industrial Engineering

KANSAS STATE UNIVERSITY
Manhattan, Kansas
1976

Approved by:

Major Professor

TABLE OF CONTENTS

Page

INTRODUCTION 1

LITERATURE REVIEW 1

Single vs Team Inspection 1

Experimental Evidence 2

Conclusions from Experimental Evidence . . . 11

Inspection Job Design 13

Signal Detection Theory 15

PROBLEM 23

METHOD 24

TASK 24

SUBJECTS 34

PROCEDURE 34

RESULTS 43

SPEED OF WORKING VS WORK FACTOR TIME STANDARDS 43

ANALYSIS OF TYPE I ERROR 44

ANALYSIS OF TYPE II ERROR 49

ANALYSIS FOR MULTIPLE-MONITOR TREATMENT 58

SIGNAL DETECTION THEORY ANALYSIS 61

CONCLUSIONS AND DISCUSSION 72

REFERENCES 74

APPENDICES 77

APPENDIX 1: Signal Detection Theory 77

Page

APPENDICES (cont.)

APPENDIX 2: Derivation of Multiple-monitor Equations 88

APPENDIX 3: Written Instructions for the Subjects 92

APPENDIX 4: List of Defective Slides 95

APPENDIX 5: Score Sheet 96

APPENDIX 6: List of Numbers Used 97

TABLES

Table	Title	Page
1	Variables influencing inspector behavior .	3
2	Proportions of defects detected	15
3	Determination of information content of an action Read Span for each type of message	28
4	Work Factor analysis of the Read operation	29
5	Total Work-Factor time analysis for the task "Inspect slide with no defect" . .	30
6	Total Work-Factor time analysis for the task "Inspect slide with defect no. 1" .	31
7	Total Work-Factor time analysis for the task "Inspect slide with defect no. 2" .	32
8	Total Work-Factor time analysis for the task "Inspect one slide for defect" . .	33
9	Task breakdown	41
10	Experimental design	42
11	Inspection effeciencies for detection of defect 1	45
12	Inspection effeciencies for detection of defect 2	46
13	Team efficiency for inspection of defect 1	47
14	Team efficiency for inspection of defect 2	48
15	Comparison for statistical significance between the means of the four conditions	50

Table	Title	Page
16	Analysis of variance of inspection efficiency of defect 1	51
17	Analysis of variance of inspection efficiency of defect 2	52
18	Inspection data for Type II error on defect 1	53
19	Inspection data for Type II error on defect 2	54
20	Comparison for statistical significance between the means of the four conditions on Type II errors	55
21	Analysis of variance of Type II errors on defect 1	56
22	Analysis of variance of Type II errors on defect 2	57
23	Efficiencies in various modes of team formation for inspection of defect 1	59
24	Efficiencies in various modes of team formation for inspection of defect 2	60
25	Pearson product moment correlation coefficients between the observed and expected inspection efficiencies on Condition 4	62
26	SDT measures of the index of detectability d' for inspection of defect 1	65
27	SDT measures of the index of detectability d' for inspection of defect 2	66
28	SDT measures of the decision criterion point for inspection of defect 1	67
29	SDT measures of the decision criterion point for inspection of defect 2	68
30	Correlation coefficients between the detectability index d' and inspection efficiencies	70

TABLE OF FIGURES

Figure	Title	Page
1	Example of defective and non-defective slides	25
2	Model of visual inspection task	26
3	Experimental layout for condition 1	35
4	Experimental layout for condition 2	37
5	Experimental layout for condition 3	38
6	Experimental layout for condition 4	40
7	Calculation of d' and x_c for the best subjects	64
8	Calculation of d' and x_c	83
9	Receiver operating characteristic curves	85
10	Tree diagram illustrating probabilities of the various decisions of the group of inspectors	89

INTRODUCTION

LITERATURE REVIEW

Single vs. Team Inspection

Visual inspection as a processing task is defined by Jackson (1957) as the aided or unaided observation of details, without measurement, to determine the conformance and completeness of the part or finished product. Applied psychologists classify the inspection task as a monotonous perceptual task; that is, a situation in which an inspector is monitoring a display in search of critical but infrequent signals (defectives). Such tasks are numerous and of considerable practical importance.

The inspection may be a hundred percent, part-by-part inspection, or a sampling scheme may be used. Researchers have designed working conditions with optimum values of the physical variables of illumination, exposure time, visual acuity, target complexity, contrast, rate of change of visual angle, etc. The studies were aimed at minimizing the errors the inspector might make and designing the conditions so as to complete the task in the minimum time or least errors. The basic idea is to improve inspection performance through tools and methods that overcome the inspector's limitations and enhance his capabilities. Colquhoun (1957) and McKenzie

(1958) have enumerated the factors that can influence inspector behavior. They are summarized in Table 1.

In many industrial and military situations where human monitors are required, it is important that no signal be overlooked. It is well known that the average human observer is a notoriously poor monitor, especially if the watch period is prolonged and the signals are infrequent and of low intensity.

Experimental Evidence

Schafer (1949) was the pioneer in the field of multi-monitor signal detection systems. He suggested that on many sorts of visual and auditory displays the chances of detecting a particular signal are significantly better for several independent observers than for any single observer alone. Schafer (1949), in fact, did collect some empirical data in support of this view. Using five observers and artificial sonar signals in seven modes of presentation, he concluded that a significant increase in detection probability can be realized by adding more observers, although little is to be gained by adding more than two. He noted further that the maximum increase in detection ordinarily cannot be expected because fluctuations in the background level caused the observers to detect the same signals. He also concluded

Table 1. Variables influencing inspector behavior.

1. <u>Task</u>
   a. Defect probability
   b. Complexity of the nature of defect
   c. Number of faults occurring simultaneously
   d. Time allowed for inspection
   e. Frequency of rest periods
   f. Illumination
   g. Time of day
   h. Objectivity of conformance standards
   i. Inspection station layout
   j. Method of inspection
2. <u>Organizational and social</u>
   a. Training
   b. Peer standards
   c. Management standards
   d. Behavior of operator or group producing the item
3. <u>Individual abilities</u>
   a. Visual acuity
   b. General intelligence and comprehension

that the gain in probability of detection of his weak auditory signals, when combining the responses of two or three isolated monitors, was less than the gain predicted by a probability model for combining independent probabilities.

Baker, Rogerware and Sipowicz (1962) extended Schafer's work to a visual rather than an auditory display as situations of visual monitoring have extensive applications. Twenty-eight trainees monitored brief interruptions of a continuous light source for a three hour period in individually isolated rooms. Individual detection probabilities for each subject were analyzed to determine the minimum number of subjects necessary to insure the detection of every one of the 72 signals presented. The minimum number was six. The findings, however, depended entirely upon the monitoring ability of the specific subjects used. For subjects of high individual detection probabilities, only three were required to ensure detection of every signal, but with poor monitors fourteen were needed.

This study was criticized by Wiener (1963) for using the independent events model in the light of Schafer's findings. Wiener claimed that Baker et. al's. study did not observe the actual performance of multi-man teams, but simply estimated it using individual detection probabilities and a mathematical model. Wiener (1964) examined the performance of multi-man teams in a visual monitoring task.

There were four groups: one, two, and three-man teams, and another three-man team in which the members monitored in isolation but hadtheir responses combined as if they were switches in a parallel circuit. The inspection time was 48 minutes (4 sessions of 12 minutes each), in which 32 signals appeared, 8 signals in each session. (As team size increased from one to two men, probability of detection significantly increased from 78% to 91%. The increase from two to three men (from 91% to 94%) was not statistically significant.) The combined performance of the three monitors in isolation was superior to the three monitors working together (99% vs 94%).

However, Bergum and Lehr (1962) reported relatively low correlations of 0.11 and -0.15 between detection rates of the isolated members of two-man simulated teams. They conducted two experiments to determine the effect of the presence of a second monitor in the same monitoring situation upon individual monitoring performance. Two groups of twenty subjects each were used in both the studies. All subjects monitored a circular light display for a period of 90 minutes without rest. Each of the three isolation booths were equipped with a 13 inch diameter circular panel, with 20 lamps which illuminated in sequence. A signal was the failure of a lamp to illuminate in its normal sequence. Experiment 1 had 24 signals/hour; Experiment 2 had 6 signals/

hour. The 40 subjects were randomly assigned to two groups of 20 subjects each. The control group was individuals working in isolation; the experimental group was pairs of individuals working independently in the same booth with freedom to converse about anything but the occurrence of signals. Separate measures of the frequency of correct signal detections were taken on all subjects in both groups. Neither experiment indicated an overall facilitation or hindering of performance resulting from pairing.

Morgan and Alluisi (1965) had 24 subjects perform a visual watchkeeping over a 30 minute period. Eight subjects worked singly and 16 worked as pairs. The proportion of signals detected in the paired-subject condition (0.91) was significantly greater than in the single-subject condition (0.74), but the estimated likelihood of an individual watchkeeper's detecting a signal was essentially identical in the two conditions (0.69 and 0.74 for the 2 and 1-man systems, respectively). The results were interpreted as supporting the inference that the watchkeeping behaviors of paired subjects are operationally "independent".

Hornseth and Davis (1967) examined one and two-member teams on three target finding tasks. For task 1 they used $8\frac{1}{2}$ x 11 inch aerial photographs and their corresponding cue cards. The target cue card was a cut-out of a target to be found in the photograph. The stimulus material for task 2 and task 3 was generated by a digital computer. A 15 x 15 matrix of I's and T's was generated. Then a space in the matrix was selected by a random process, then the experimenter cleared the spece and placed a target in the space. For task 2 the target was a single letter I and for task 3 the target was a 2 x 2 matrix of I's. Seventy-eight undergraduate students were tested (26 individually and 52 as 26 teams of 2 each). All the photographs with incorrect identifications were eliminated from the analysis. Team performance was significantly faster than individual for task 2 and task 3 but not for task 1. For task 1, mean performance scores (in seconds) for teams and individuals respectively were 17.27 and 20.66; for task 2, the mean performance scores were 37.06 and 50.52; and for task 3, the mean performance scores were 14.64 and 23.59. The target cue cards provided general target location information as well as target identification information. Target location information provided by the target cue cards may have restricted the search area to such an extent that the task could be accomplished more effectively by a single individual.

Waag (1972) investigated two variables; 1) team size, and 2) the decision rule employed in defining the requirements of a team response.

Team size varied from two to four members. The decision rule was varied from "parallel" (a response by any one or more members producing a team response) to "series" (a team response occurred only if all members responded). In the parallel mode, the detection rates were .86, .94 and .97 for teams of two, three and four members respectively. In the series mode, the detection rates were .43, .28 and .21 for teams of two, three and four members respectively. "Parallel" teams maximized correct detections while "series" teams eliminated all false alarms. For each decision rule, detection rate increased as a function of team size. For each team size, detection rate deteriorated as the decision rule required more members to respond correctly.

Schlegel, Boardman and Purswell (1973) compared a single inspector system with a two-inspector system on a 45-minute task. The two inspector system consisted of two inspectors in series, each of whom examined all the targets. Only three subjects were used. Inspection effectiveness was measured in terms of the number of defective targets accepted and the number of good targets rejected. The targets were black Landolt rings (defective) and closed circles (good) on a grey background. The probabilities of committing a Type II error were 0.09, 0.16 and 0.10 for individual inspectors compared to 0.0, 0.0 and 0.03 for the two-inspector combinations. Both systems showed a decrement in performance with time, but the two-inspector system showed a much smaller change compared with the single-inspector system.

Morrissette et. al. (1975) studied the team organization and monitoring performance of one and two-man teams using division of labor and redundancy to detect randomly presented signals on one of four milk-glass display windows. Twenty subjects were tested as individuals and 40 were tested in pairs. The four groups were 1) individuals, manning four displays each, 2) individuals, manning two displays each, 3) 2-man team in redundancy, each manning four displays, 4) 2-man team with division of labor, each manning two displays.

Eight signals occured during each of four 12-minute work periods (2 signals on each display). Critical signals were repeated until the subject responded, or until 5 seconds had elapsed.

The time scores for 2-man team-redundancy were obtained by selecting the fastest of the two responses made by the two team members to each critical signal. Using this time (response times), team performance was superior to individual performance. The mean response times by individuals was 1.22 seconds; for teams, 0.88 second. The mean detection time scores for the two team conditions (0.88 second for both redundancy and division of labor) were not significantly different.

Lion et. al. (1975) studied the visual arrangement of material and of working singly or in pairs, upon performance at simulated industrial inspection. They used three different types of conveyor belts: 1) one line, one inspector, 2) three-line, at 1/3 speed, one inspector, and 3) six-line at 1/3 speed, two inspectors. Three 12 minute test sessions (one test session at each conveyor) were administered on 96 subjects to inspect plastic discs for surface flaws.

The subjects working together had a significantly higher probability of detection (.975) than when working alone (.973). Performance on the three machines differed highly significantly ($p < .001$), with probability of detection of .969 on the single belt, .973 on the 3-line belt and .975 on the 6-line belt.

For Type II errors (perfect discs rejected), performance on the three machines also differed highly significantly ($p < .001$), with 2% errors on the single line belt, 1.8% on the 3-line belt and 1.1% on the 6-line belt.

McFarling and Heinstra (1975) examined motivational differences between self-paced and machine-paced inspection tasks, and measured subject perceptions of inspection tasks. Twenty women subjects in both self-paced and machine-paced conditions inspected 225 slides of simulated printed circuits. In the machine-paced task each circuit was displayed for 8 seconds, during which the subject visually inspected the circuit and determined whether it was acceptable or defective. At the end of 8 seconds the circuit disappeared from the subject's view and in the following six seconds the next circuit was automatically advanced and readied for inspection. This cycle was repeated for all 225 circuits.

In the self-paced condition, circuit presentation rate was inspector controlled. They were asked (but not required) to attempt to inspect all 225 circuits within the same time frame as in the machine-paced condition (52.5 minutes). The self-paced subjects performed better, on both criteria, defects detected and the good circuits accepted (95.3% vs. 99.2% and 98.0% vs. 99.7%). Both groups reported the task as basically dull and uninteresting.

Conclusions from Experimental Evidence

Individual performance in a visual signal detection task has been shown to reflect reliable scanning strategies or processing rules which are similar for different subjects. The existance of such stable data on a particular task for individual performance provides a basis for prediction of group performance under different models of behavior. (Group performance is better than individual performance under most conditions on most tasks.) However, models of group performance often have predicted better performance than that actually attained by real groups. A characteristic of the cited research is the apparent lack of concern for commissive errors or false alarms. It is not difficult to imagine systems in which a false alarm may be just as important as the correct detection of a signal, e.g. fire alarms and seismic equipments.

Inspection Job Design

An industrial inspector's typical job is to divide a batch of product into acceptable items and rejectable items, where 'acceptable' and 'rejectable' are defined in terms of standards agreed between the manufacturer and the customer.

Researchers in the field of signal detection have shown that performance improves when a number of inspectors were used.

Team construction can be approached in one of two ways. First, a team can be designed to consist of 'n' individuals operating in arrangements in which they are free to communicate with one another. This is perhaps the most widely used approach to team configuration. An alternative approach makes use of 'n' individuals, each performing the same task independently of one another. In most cases, the individual members are isolated in order to prevent communication. Teams formed in this manner are sometimes referred to as simulated teams. Regardless of which approach is used, team performance is based on some combination of the responses of the individual members. The procedures for arriving at a team response from the individual member responses is referred to as the "decision rule".

Statistically, the probability of detection of a signal by at least one inspector of the 'n' independent inspectors, is given by

$$P_n = 1 - \prod_{i=1}^{n} (1 - P_i) \qquad (1)$$

where P_i = the probability of detection of the ith inspector. Thus if P_i of each of 10 independent inspectors is equal to 0.5, P_n = 0.999. The specification of the inspector is general; it may be a man, a machine, or a man-machine combination.

A combination of inspectors may operate under a variety of decision rules. The rule underlying Equation (1) may be termed a (n,n) decision rule in that every inspector must respond "no signal". (A "parallel" system in which the "current" can flow through any path.) Equation (1) applies to the increase in both the correct acceptance rate and the incorrect acceptance rate.

At the opposite extreme, the decision rule underlying Equation (2) may be termed a (y,y) rule in that every inspector must respond "yes, a signal has been received" for the combination of inspectors to respond "signal". Otherwise, the combination of inspectors responds "no signal". (A "series" system in which "current" flow is prevented by any circuit break.) Equation (2) also applies to the decrease in both the correct acceptance rate and the incorrect acceptance rate.

$$P_n = \prod_{i=1}^{n} P_i \qquad (2)$$

Speed of working or time to complete the inspection task has been considered to be of lesser importance in situations where group inspection is thought to be a possible solution to the problem of high inspection reliability. This is because, from an economic point of view, the gain by saving on time might be more than offset by the expense of a defect not being detected.

The inspection task may be paced, unpaced or self-paced. In this study sufficient time was given to the subjects for completion of the machine-paced task. McCormick (1964) says,

> The rate of presentation should be within acceptable bounds as far as human performance is concerned. Further, where spacing of the signals to be presented can be controlled, one should avoid short intervals, the bunching of signals and short intervals between signals following previous responses. Where feasible, it seems desirable to permit the individual to control the rate of signal input.

Table 2 summarizes the literature.

Signal Detection Theory

Signal detection theory not only provides a theoretical framework for the inspection task, but identifies performance parameters that may be more suitable for task evaluation than those previously used. Originally, Wald's theory of statistical decision (1950) was translated into a theory of signal

Table 2. Proportion of targets detected.

		Parallel mode			Series mode		
Reference	Individual	2	3	4	2	3	4
Baker et. al.(1962)	0.69	0.90	0.97	0.99			
Wiener (1964)	0.78	0.91	0.94	0.99			
Bergum & Lehr (1962)	0.81	0.86					
Morgan & Alluisi (1965)	0.74	0.91					
Waag (1972)		0.86	0.93	0.97	0.43	0.28	0.21
Schlegel et. al. (1973)	0.99				0.98		
This study defect 1	0.78	0.94			0.62		
This study defect 2	0.85	0.96			0.73		

detection, with applications in the field of radar design. Considerable success has been achieved in recent years by treating the human operator as a signal detection system, as if he were making a statistical decisions of whether a 'signal' is present or whether 'noise' alone is present. Swets (1961) shows how such a treatment is an accurate description of human performance in a variety of experimental signal detection tasks.

Signal detection theory considers detection performance as a judgment process in which observations are made of events occurring in a fixed interval of time. The events are classified by the observer either as signals (observation events composed of signal plus any extraneous noise) or as nonsignals (observation events composed entirely of extraneous noise) depending on a criterion which the observer adopts. The term "observation" refers to the sensory data on which the observer bases his decision, and it is assumed that any observation may arise with specific probabilities either from noise or from signal plus noise.

An additional assumption, often made in signal detection theory applications and one that was used in this thesis, is that the observation probability density functions are normal with equal variance. Signal detection theory also assumes the observer's observation is converted through a psychological mechanism to a value on a continum of likelihood ratio for purposes of comparison against the decision criterion point.

Signal detection theory has been formulated to explain the phenomenon of signal detection under vigilance conditions. Basically the human operator in an inspection task performs two operations: first, the detection of defects and second, the classification of the defectives according to the severity of the defects. Since the basic task demanded by the inspection process is detection of defects, it seems logical to approach the question of human reliability in this type of work by considering the efficiency of the human operator as a detector of small deviations from the product specifications, called signals in his working environment. The ability of human beings to make judgements regarding perceived variations in the nature of these signals also is of interest.

The use of detectability indexes from signal detection theory (Green and Swets, 1966) as measures of performance has the advantage of separating sensitivity from response bias and allowing prediction of group sensitivity under various assumptions, such as that of independent observations by members of the group. Two models in particular are relevant to the way in which individuals might combine observations to make a group decision in a signal-detection task. Behavior under each model could be considered the behavior of a pseudogroup (since each uses individual rather than group performance), and could provide a basis for suggesting the mechanisms used by a real group in such a task.

One model is an integration model (Green and Swets, 1966) which assumes that the information available to individuals is used in an optimal combination to provide information for a group decision. Individual measures of sensitivity (d') can be combined to yield a predicted value (d_n') for a group decision assuming that individual observations are independent events. The second model is a decision-threshold model, which also assumes independence of individual observations, but which assumes that the decisions reached by individuals are combined in a logical fashion to reach a group decision.

Signal detection theory is discussed in Appendix 1.

Smith (1972) has shown the tasks of signal detection and of inspection to be analogous. In placing the "fundamental detection problem" within the context of industrial inspection, it is important to recall that the SDT observer simply reports whether or not the sensory excitation which he receives would lead him to prefer one decision over the other. Viewed in this light, the corresponding "fundamental inspection problem" would require that the inspector report whether or not the perceptual information obtained in his inspection would lead him to accept or reject the given item. This analogy between the fundamental detection problem and the inspection problem served as the basis for his research as well as for other studies reported. Smith first identified the parameters associated with the fundamental detection problem. Then he identified the corresponding analogous parameters of the inspection process. Conditions of noise (signal background) and signal plus noise correspond to conditions of acceptable and defective product. A hit, miss and false alarm correspond to a correct detection of a defect, Type I decision error and Type II decision error. Using the parameter d' as a measure of signal detectability has intuitive appeal in the context of industrial inspection. The more nearly the defective product resembles the acceptable product, the weaker the signal strength, and the smaller the theoretical value of d'.

A decrease in d' necessarily results in an increase in the total probability of incorrect decisions. The concept of a movable decision criterion point provides another area of correspondence between the SDT and inspection. It is a matter of common experience in industry that a simple word from a supervisor or product engineer can result in a shift in the proportion of product passing inspection, with no change in the manufacturing process or variation in the incoming lot fraction defective. This change in performance is hypothesized to result from a shift in the inspector's decision criterion.

Smith (1972) further says that the analog between SDT and industrial inspection appears strong, with one notable exception. A potentially serious divergence rests in the definition of signal detection as a decision made during a fixed interval of time. Most individual inspection tasks are operator paced and hence this constraint is violated. The effect of self-pacing on the estimated value of d' could be substantial. When operating at conditions above threshold (most industrial inspection operations), operators can improve their accuracy if they sacrifice speed. Hence, a difficult detection problem, which would normally yield a small value of estimated d', could produce a fairly high estimated value if the observer simply slowed down and reduced his error rate.

Drury and Sheehan (1969) studied the ergonomic and economic factors in an industrial inspection task conducted in a factory which specializes in the manufacture of small precision metal parts (hooks). Six experienced inspectors between 30 to 62 years were their subjects. The subject's visual acuities ranged from 0.67 to 1.54 min^{-1}. The performance of these inspectors were assessed by having them inspect batches of 250 hooks, of which 200 were good and 50 defective. Twelve different batches were used, six having one type of defect only, five having 25 of each of two defects, and one having no defect at all. Four different levels of illumination were used (180, 90, and 65 f.c. and the fourth level was of the inspectors choice). There was a significant difference between the inspectors ($p < .01$) and no significant differences between illumination levels or defect types. From the study it appeared that an inspector, discovering a defect early in the task, expected a run of these and detected nearly all of them. However, when the defect is undetected early in the task, all subsequent similar defects were undetected. The studies showed that if the inspectors were acquainted with the type of defects to expect, the percentage of rejects they missed dropped from 18.5 to 7.5.

Sheehan and Drury (1971) re-studied this experiment in the light of signal detection theory, so the advantages and drawbacks of the SDT were brought forward. The major advantage of analysis by SDT was that it gave two measures, the detectability d' and the criterion β, which were more useful than either of the probabilities from which they were calculated. Detectability was found to be an unbiased measure of the difficulty of the inspection task to the inspector. The criterion measures the combined effect of the various pressures on the inspector. Receiving prior information about which faults a batch would contain improved the detectability of the faults and reduced Type I and Type II errors simultaneously.

The main drawbacks of the SDT analysis was that it did not explain the tendency of detecting all defects when an observer finds an error in the earlier part of the inspection.

From the previous researches cited, it can be concluded that quite a number of studies investigated relationships between detection performance, team size, and decision rules for signal detection or vigilance tasks. But virtually no study was undertaken to investigate the effect of combining inspectors on detection performance in an inspection task. Some of the researchers tried to study inspection task performance in the light of signal detection theory, but their investigations were restricted to one individual inspector, rather than to a group of inspectors. One aspect of inspection that has received little research attention is the nature of inspection task itself.

The present investigation examined inspection performance between four inspection situations with two simulated defects in each item and two inspectors in each team.

PROBLEM

This study compared inspection efficiencies for two methods of inspection. Two types of team organization were considered:

1) The division of labor organization where the items to be inspected were divided among the members of the team (i.e. Jane inspects items 1 through 10, Paula inspects items 11 to 20).

2) The redundancy organization where all items were inspected by each team member (i.e. Jane inspects items 1 to 20 and Paula inspects items 1 to 20.) What one team-member misses, the other team-member may detect. Hence the individual team-member's work-load will be greater under the redundancy design, if presentation speed is constant.

M E T H O D

TASK

Three random numbers, generated through a computer, with 10 digits each were typed on paper. This message was then reproduced onto a transparent plastic sheet. The sheet then was cut into 1-3/4 x 1-3/4 inches and placed in a 2 x 2 inch slide frame. The window of the slide was 1-5/16 x 7/8 inches as shown in Figure 1.

Occurrence of the digit '6' on the slide was designated as defect 1; occurrence of any two identical consecutive digits was designated as defect 2. Absence of these two conditions was considered as no defect.

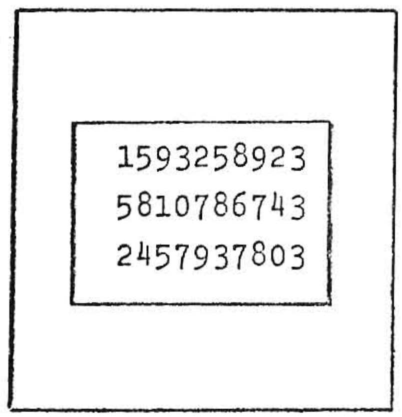
A model of visual inspection task (Drury, 1974) suits this task and is reproduced in Figure 2.

The numbers were typed with a type size of 12 letters to an inch. One hundred and sixty such slides were used for the experiment. A white cardboard with a reflectance factor of 0.7 was used as the screen. The typed numbers have a reflectance factor on the screen of 0.1. Thus the task has a brightness contrast of $(.7 - .1)/.7 = 0.85$.

The projector was placed 100 inches from the screen; this magnified the digits 24 times. Thus each projected digit was 2.4 inches in height and 2 inches in width, and each number series was 20 inches long. Each digit therefore subtended

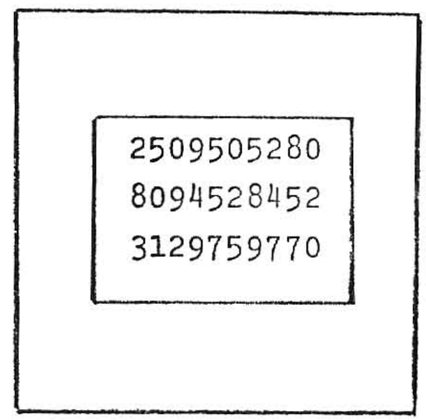
THIS BOOK CONTAINS NUMEROUS PAGES WITH DIAGRAMS THAT ARE CROOKED COMPARED TO THE REST OF THE INFORMATION ON THE PAGE.

THIS IS AS RECEIVED FROM CUSTOMER.



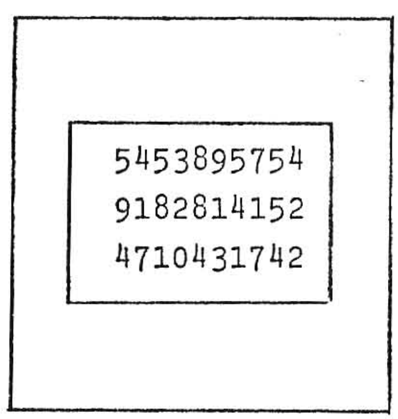
defect # 1

(presence of digit 6)



defect # 2

(presence of consecutive identical digits in a row)



no defect

Figure 1 Example of defective and non-defective slides.

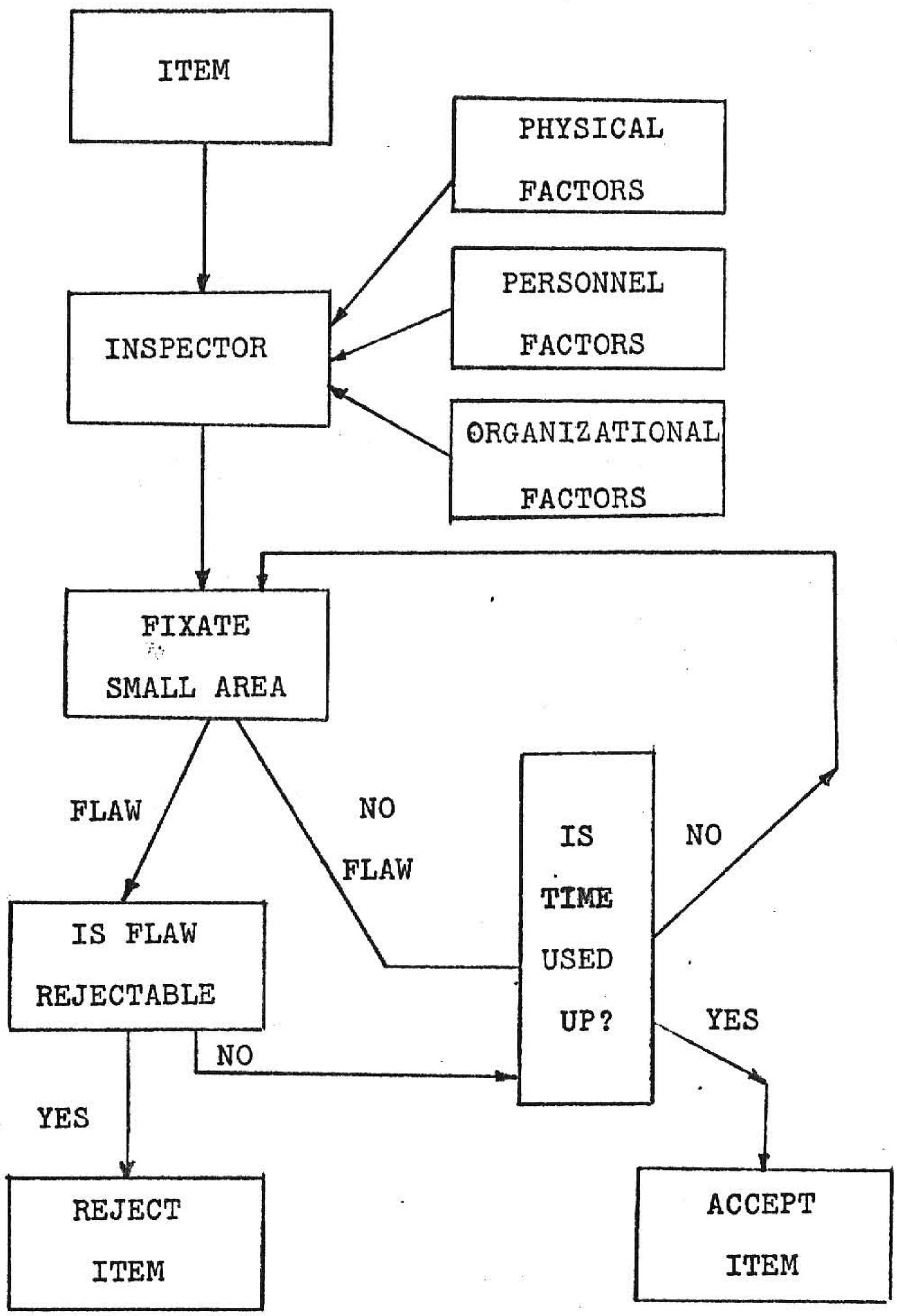


Figure 2 Model of visual inspection task.

$\theta = \tan^{-1}(2.4/100) = 82.5$ minutes of arc. The ambient noise from the projectors was 67-70 dbA.

Ten percent of the slides had defect 1 and another ten percent had defect 2. The 480 numbers used are listed in Appendix 6. The subjects were not informed of the percentages of defects. Two slide projectors with time settings of 5 and 10 seconds were used. All situations were machine-paced. Another 20 similar slides were used for a practice session for each subject. The subjects were given the results of this practice session to ensure that they had understood the task throughly. Both subjects in the team were seated in the I.E. Lab. I. They recorded on a score sheet (Appendix 3) the type of defect on each slide inspected.

A Work Factor analysis (Quick, Duncan, and Malcolm, 1962) was made for the task. Table 3 shows there are 0.1 Work-Factors per digit for the message with defect 1, 0.2 Work-Factors per digit for message with defect 2, and 0.0 Work-Factors per digit for a no-error message. Table 4 shows the Work-Factor time for the 'Action Read' is 267 Work-Factor time units when there is no error. Table 5, 6 and 7 show the total Work-Factor time units for the task of inspecting one slide with each type of message, while Table 8 shows the overall total Work-Factor time units for an overall task of 'inspecting one slide'.

Table 3. Determination of information content of an action Read Span for each type of message.

Example of defect #1				
Reading Material	3851423518	1082138903	2365304818	Total
Read Segment	1	1	1	3
Digit Groups	1	1	1	3
Digits	10	10	10	30
Information units	-	-	1	1
Program Work Factors	-	-	3	3
Work-Factors per Digit			3/30 = 0.1	

Example of defect #2				
Reading Material	0142385095	1745425987	8897139341	Total
Read Segments	1	1	1	3
Digit Groups	1	1	1	3
Digits	10	10	10	30
Information units	-	-	2	2
Program Work Factors	-	-	3	6
Work-Factors per Digit			6/30 = 0.2	

Example of no defect				
Reading Material	2120978023	1802809035	1708797975	Total
Read Segments	1	1	1	3
Digit Groups	1	1	1	3
Digits	10	10	10	30
Information units	-	-	-	-
Program Work Factors	-	-	-	-
Work-Factors per Digit			0/30 = 0.0	

Table 4. Work Factor analysis of the Read operation.

Description	Number of Letters	Information Units	Program Work-Factors	Work-Factors Per Digit	Work-Factor Time Units	Percentage Occurrence	Total Work-Factor Time Units
Action Read message of defect #1	30	1	3	0.1 Class 21A	187		
Read Segment Add-on				F-3	80		
Total (message 1)					267	10	26.7
Action Read message of defect #2	30	1	3	0.2 Class 21A	187		
Read Segment Add-on				F-3	80		
Total (message 2)					267	10	26.7
Action Read message of no defect	30	0	0	0.0 Class 21	0		
Read Segment Add-on				F-3	80		
Total (message 3)					80	80	64.0
TOTAL							117.4

Table 5. Total Work-Factor time analysis for the task "Inspect slide with no defect".

Description	Analysis	Work-Factor time units	Reference tables in Quick, Duncan & Malcolm
Eye Focus	FoH2	55.0	13.1
Eye Shift	ES10	-	13.1
Inspect intervals	I_1B_8	16.0	13.6
Action Read		80.0	Table 4 this study
Identify	IdB_3	24.0	13.2
React to one of 3 signals	DeB_3	24.0	13.2
TOTAL		199.0	

Table 6. Total Work-Factor time analysis for the task "Inspect slide with defect #1".

Description	Analysis	Work-Factor time units	Reference tables in Quick, Duncan & Malcolm
Eye Focus	FoH2	55.0	13.1
Eye Shift	ES10	-	13.1
Inspect intervals	I_1B_8	16.0	13.6
Action Read	-	267.0	Table 4 this study
Identify	IdB_3	24.0	13.2
React to one of 3 signals	DeB_3	24.0	13.2
TOTAL		386.0	

Table 7. Total Work-Factor time analysis for the task "Inspect slide with defect #2".

Description	Analysis	Work-Factor time units	Reference tables in Quick, Duncan & Malcolm.
Eye Focus	FoH2	55.0	13.1
Eye Shift	ES10	-	13.1
Inspect intervals	I_1B_8	16.0	13.6
Action Read		267.0	Table 4 this study
Identify	IdB_3	24.0	13.2
React to one of 3 signals	DeB_3	24.0	13.2
TOTAL		386.0	

Table 8. Total Work-Factor time analysis for the task "Inspect one slide for defect".

Description	Analysis	Work-Factor time units	Reference tables in Quick, Duncan & Malcolm
Eye Focus	FoH2	55.0	13.1
Eye Shift	ES10	-	13.1
Inspect intervals	I_1B_8	16.0	13.6
Action Read		117.4	Table 4 this study
Identify	IdB_3	24.0	13.2
React to one of 3 signals	DeB_3	24.0	13.2
TOTAL		172.4	

SUBJECTS

Twelve teams of 2 students in each team served as subjects. The students were American females from freshman Home Economics classes. Each condition took 800 seconds; thus the entire experimental run of 4 conditions took 53.4 minutes plus time to change between conditions. A 5 minute break after each condition was given. Each subject was paid $3. Thus a subject budget of $72 served the experiment.

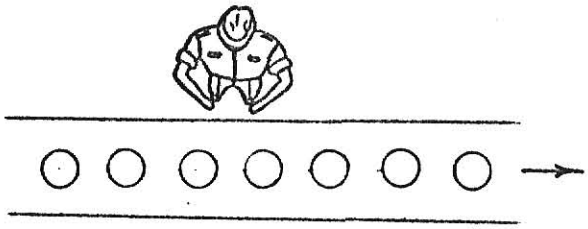
PROCEDURE

CONDITION 1

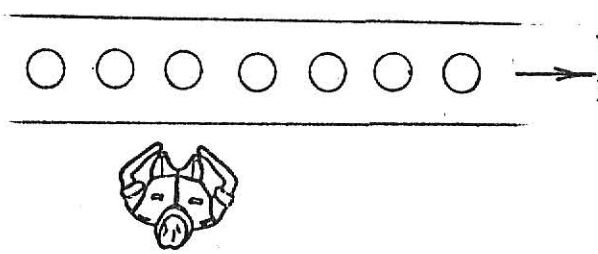
Both subjects were seated in the I.E. Lab I facing each other. The entire lot of 160 slides was divided into two sub-lots; 1-80 and 81-160. Each team member inspected 80 slides at a speed of 1 slide per 10 seconds. Each subject inspected for both types of defects in their respective lots. Two slide projectors were used. The experimental layout is shown in Figure 3.

CONDITION 2

The subjects were seated in the I.E. Lab I facing each other. Each subject inspected 80 slides for defect #1 (subject 1 on slides 1 through 80, subject 2 on slides 81 through 160). Then they inspected the same lot for defect #2. Time for each



Task: Inspect for defect 1 and defect 2 in slides numbered 1 through 80.



Task: Inspect for defect 1 and defect 2 in slides numbered 81 through 160.

Figure 3 Experimental layout for condition 1.
Both subjects seated in the I.E. Lab. I facing each other. Two slide projectors were used.

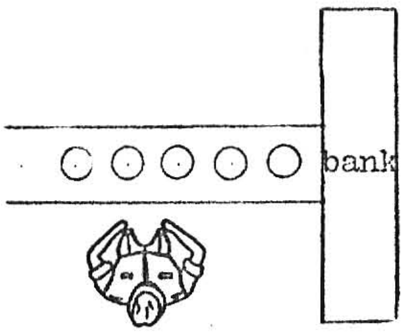
inspection was 5 seconds per slide for each defect. Two slide projectors were used. The experimental layout is shown in Figure 4.

CONDITION 3

The subjects were seated in the I.E. Lab. I. in the same row, beside each other, both facing the same screen. Each subject looked for one defect; on slides 1-80 subject 1 looked for defect 1 and subject 2 looked for defect 2. After 80 slides the task of both the subjects was reversed to counter-balance the situation. For slides 81-160 subject 1 inspected for defect 2 and subject 2 inspected for defect 1. Each slide was projected for 5 seconds. Only one slide projector was used. The experimental layout is shown in Figure 5.

CONDITION 4

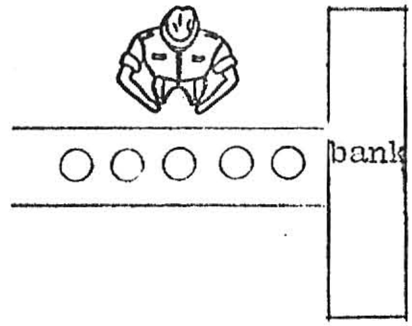
Both subjects were seated in the I.E. Lab. I facing each other. Each subject looked for both the defects. Subject 2 could detect the error that subject 1 missed. In stage 1, subject 1 examined slides 1 through 80 for defects 1 and 2, while at the same time subject 2 examined slides 81 through 160 for defects 1 and 2. They noted the defective slides on a score sheet. In stage 2, subject 1 inspected the lot 81 through 160; similarly subject 2 inspected the lot 1 through 80. The entire cumulative inspection task was treated as the



v = 1 slide per 5 seconds

N = 80 slides

subject 1

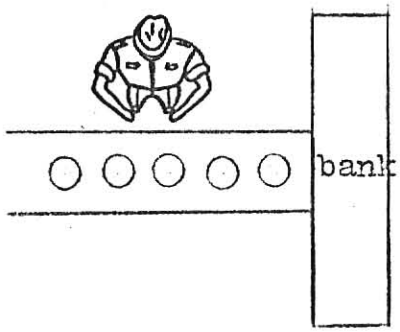


v = 1 slide per 5 seconds

N = 80 slides

subject 1

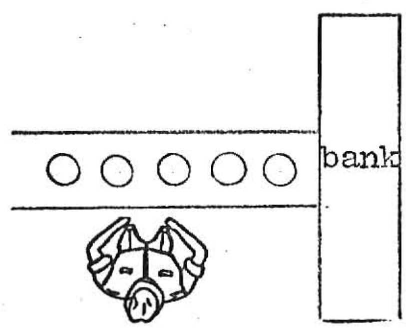
Task: Inspect slides 1 through 80 for defect 1; then inspect the same slides for defect 2.



v = 1 slide per 5 seconds

N = 80 slides

subject 2



v = 1 slide per 5 seconds

N = 80 slides

subject 2

Task: Inspect slides 81 through 160 for defect 1; then inspect the same slides for defect 2.

Figure 4. Experimental layout for condition 2. Both subjects seated in the I.E. Lab. I facing each other. Two slide projectors were used.

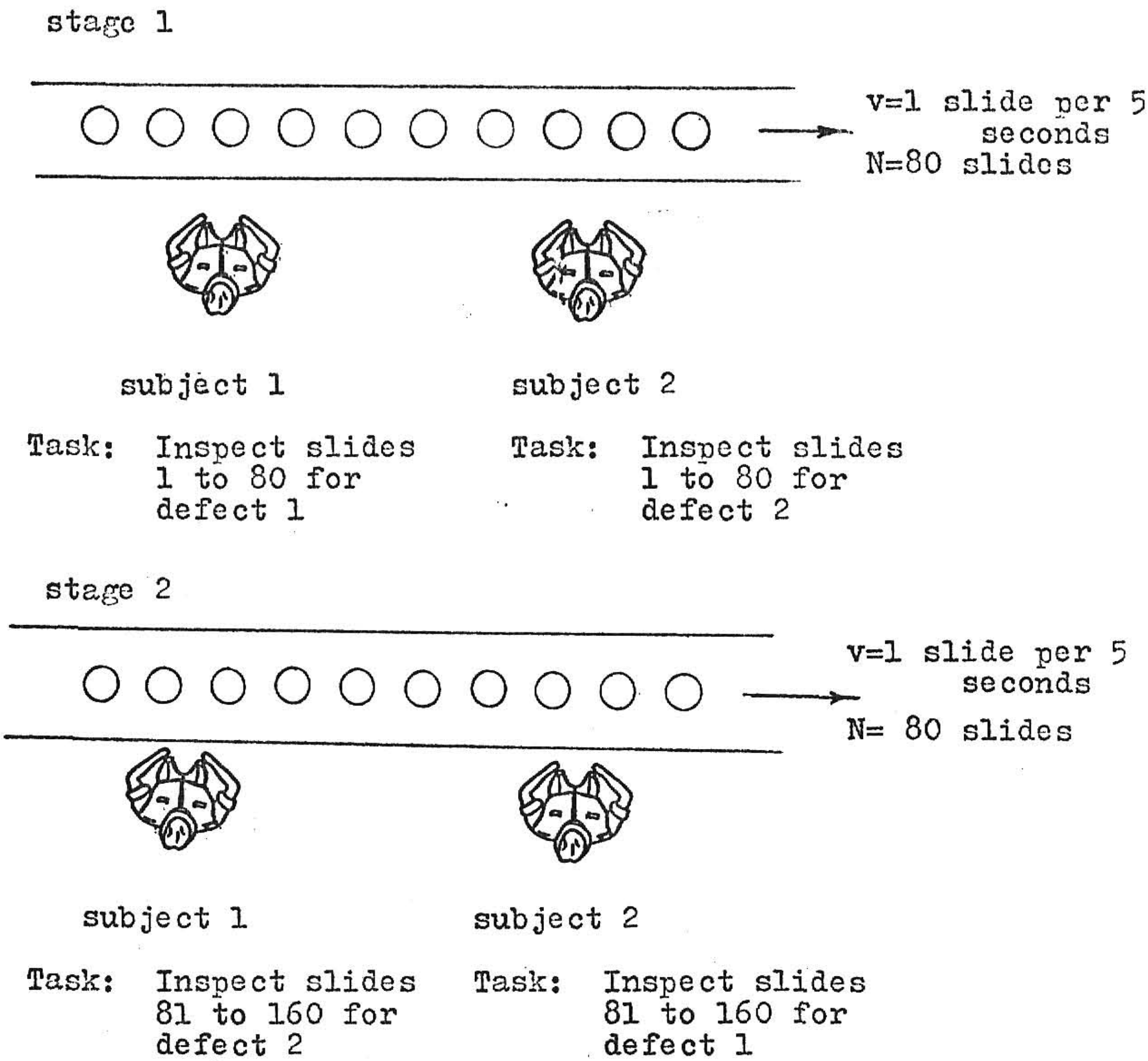


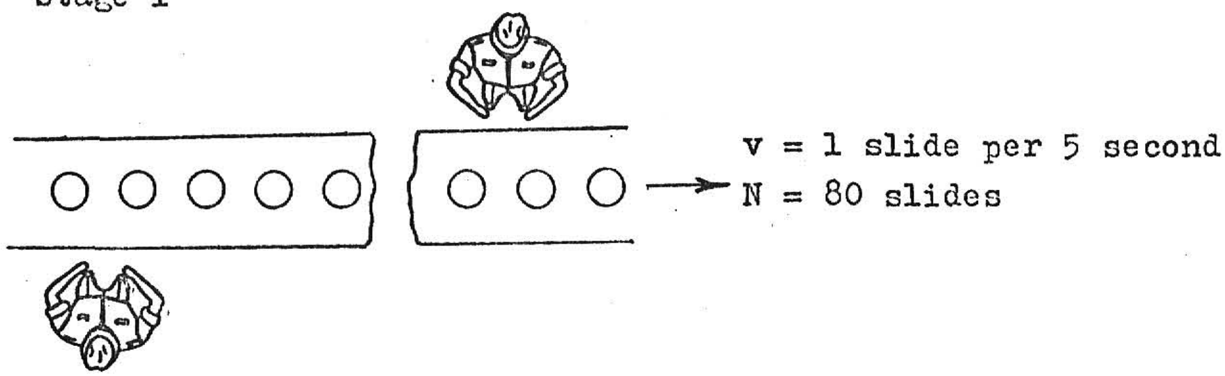
Figure 5 Experimental layout for Condition 3.
Both subjects seated in the I.E. Lab. I. beside each other. One slide projector was used.

team performance. Each slide was projected for 5 seconds. Two slide projectors were used. The performance of the team was analyzed: 1) in an integration mode, 2) in a series mode (where all members are required to detect a defect) and 3) in a parallel mode (where only one member is required to detect a defect). The experimental layout is shown in Figure 6.

Table 9 summarizes the conditions.

All four conditions were randomized so as to balance the effect of learning and fatigue. Table 10 shows the sequence in which the conditions were presented to the subjects. Four teams (numbers 1, 2, 9, and 11) were run in the morning, four (numbers 3, 4, 7 and 8) in the afternoon and four (numbers 5, 6, 10 and 12) in the evening.

stage 1



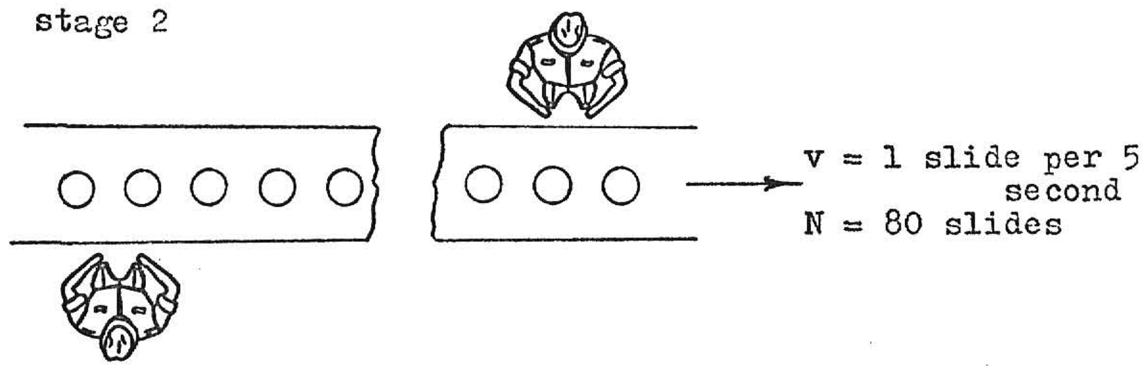
subject 1

Task: Inspect slides 1 to 80 for defect 1 and 2

subject 2

Task: Inspect slides 1 to 80 for defect 1 and 2

stage 2



subject 2

Task: Inspect slides 81 to 160 for defect 1 and 2

subject 1

Task: Inspect slides 81 to 160 for defect 1 and 2

Figure 6 Experimental layout for Condition 4.
Both subjects seated in I.E. Lab.I facing each other. Two slide projectors were used.

Table 9. Task breakdown. Total time is 800 seconds in all four conditions.

Condition	Subject	Stage	Pace (slide/sec)	Serial No. of Slides	Task (Inspects for defect nos.)	Task (type)
1	1	1	1/10	1 to 80	1 & 2	Division of Labor
	2	1	1/10	81 to 160	1 & 2	
2	1	1	1/5	1 to 80	1	Division of Labor
	1	2	1/5	1 to 80	2	
	2	1	1/5	81 to 160	1	
	2	2	1/5	81 to 160	2	
3	1	1	1/5	1 to 80	1	Division of Labor
	1	2	1/5	81 to 160	2	
	2	1	1/5	1 to 80	2	
	2	2	1/5	81 to 160	1	
4	1	1	1/5	1 to 80	1 & 2	Redundancy
	1	2	1/5	81 to 160	1 & 2	
	2	1	1/5	81 to 160	1 & 2	
	2	2	1/5	1 to 80	1 & 2	

Table 10. Experimental design of randomized sequences.

TEAMS	SEQUENCE			
1	C1	C2	C3	C4
2	C2	C1	C4	C3
3	C3	C4	C2	C1
4	C4	C3	C1	C2
5	C2	C3	C4	C1
6	C3	C2	C1	C4
7	C4	C1	C3	C2
8	C1	C4	C2	C3
9	C3	C1	C4	C2
10	C1	C3	C2	C4
11	C4	C2	C1	C3
12	C2	C4	C3	C1

RESULTS

The four conditions were compared to determine which was the best. For the purpose of analysis the detection efficiency of the inspector was calculated as (Juran, 1974):

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

where:

A = detection efficiency

d = defects reported by the inspector

k = number of good units rejected by the inspector

$d - k$ = true defects found by the inspector

b = defects missed by the inspector

$d - k + b$ = true defects originally in the process

Speed of working vs. Work Factor Time Standards

Sufficient time was given in each condition in this thesis. From Tables 5, 6 and 7 we can see that inspection of slides with no defect, defect 1, and defect 2 need 199, 386 and 386 Work Factor Time Units respectively. These correspond to 1.19, 2.32 and 2.32 seconds respectively. In condition 1, where each subject was required to inspect for both defect 1 and defect 2, they were allotted 10 seconds, whereas the maximum time needed was 2.32 seconds. In the extra time available, they could review the slide again or relax and be ready for the next slide.

In conditions 2 and 3 (each subject was looking for only one defect at a time), they were allotted 5 seconds, whereas the maximum time required was 2.32 seconds.

In condition 4 (each subject was required to inspect for both defects), they were allotted 5 seconds, while the maximum time needed was 2.32 seconds.

All the subjects were instructed that defect 1 and defect 2 would not appear simultaneously on any slide, and if any defect occurred on a slide, they were not required to inspect on that slide any further. Thus the maximum time required for inspection of any slide was 2.32 seconds - if the Work Factor times are correct.

Analysis of Type I error

The efficiency scores 'A' were tabulated as shown in Tables 11 and 12, for defect 1 and defect 2 respectively. Team efficiencies, computed by taking the mean performance of its members for all four conditions on both defects respectively, are presented in Tables 13 and 14.

The means for each condition were tested for significance by 1) Sign test (non-parametric), 2) Wilcoxon signed-rank test (non-parametric), and 3) Duncan's multiple range

Table 11. Inspection efficiencies for detection of defect 1.

		DEFECT 1			
TEAM	SUBJECT	C1	C2	C3	C4
1	1	0.2500	0.8750	1.0000	0.8750
	2	1.0000	1.0000	1.0000	1.0000
2	3	0.8750	0.8750	1.0000	0.8125
	4	0.8750	1.0000	1.0000	0.9375
3	5	0.8750	1.0000	1.0000	0.8750
	6	1.0000	1.0000	1.0000	0.9375
4	7	1.0000	1.0000	0.7500	0.5625
	8	0.8750	0.8750	0.8750	0.7500
5	9	0.8750	1.0000	0.8750	0.8750
	10	0.7500	1.0000	1.0000	0.7500
6	11	0.6250	0.7500	1.0000	0.6250
	12	1.0000	1.0000	1.0000	0.8750
7	13	0.3750	0.8750	0.8750	0.6875
	14	1.0000	1.0000	1.0000	0.7500
8	15	0.7500	0.8750	0.7500	0.8750
	16	0.8750	1.0000	0.8750	0.8750
9	17	0.8750	1.0000	1.0000	0.8125
	18	0.6250	1.0000	0.8750	0.5625
10	19	1.0000	1.0000	1.0000	0.6250
	20	1.0000	1.0000	1.0000	0.8125
11	21	1.0000	1.0000	0.8750	0.7500
	22	1.0000	1.0000	0.8750	0.6875
12	23	1.0000	1.0000	1.0000	1.0000
	24	0.6250	0.7500	0.8750	0.3750
	Mean	0.8385	0.9531	0.9375	0.7786
STD. DEVIATION		0.2068	0.0809	0.0825	0.1520

Table 12. Inspection efficiencies for detection of defect 2.

		DEFECT 2			
TEAM	SUBJECT	C1	C2	C3	C4
1	1	1.0000	1.0000	1.0000	1.0000
	2	0.8750	1.0000	1.0000	0.9375
2	3	1.0000	0.7500	0.7500	0.6250
	4	1.0000	1.0000	1.0000	1.0000
3	5	0.7500	1.0000	1.0000	1.0000
	6	1.0000	1.0000	1.0000	1.0000
4	7	0.8750	1.0000	1.0000	0.8125
	8	1.0000	0.8750	0.6250	0.8750
5	9	0.8750	0.8750	0.8750	0.7500
	10	0.8750	1.0000	0.8750	0.8750
6	11	1.0000	1.0000	0.8750	0.6875
	12	1.0000	1.0000	1.0000	0.8750
7	13	0.7500	0.7500	0.8750	0.6875
	14	1.0000	0.8750	0.8750	0.8125
8	15	0.7500	1.0000	1.0000	0.9375
	16	1.0000	0.8750	0.8750	0.8750
9	17	1.0000	0.8750	1.0000	0.9375
	18	0.7500	0.7500	1.0000	0.7500
10	19	0.8750	0.7500	1.0000	0.7500
	20	1.0000	1.0000	0.8750	0.8750
11	21	1.0000	1.0000	1.0000	0.7500
	22	0.8750	1.0000	0.8750	0.9375
12	23	1.0000	1.0000	0.8750	0.8750
	24	0.6250	1.0000	1.0000	0.8750
	MEAN	0.9115	0.9323	0.9271	0.8542
STD. DEVIATION		0.1135	0.0974	0.0969	0.1085

Table 13. Team efficiency for inspection of defect 1.

TEAM	C1	C2	C3	C4
1	0.6250	0.9375	1.0000	0.9375
2	0.8750	0.9375	1.0000	0.8750
3	0.9375	1.0000	1.0000	0.9062
4	0.9375	0.9375	0.8125	0.6562
5	0.8125	1.0000	0.9375	0.8125
6	0.8125	0.8750	1.0000	0.7500
7	0.6875	0.9375	0.9375	0.7187
8	0.8125	0.9375	0.8125	0.8750
9	0.7500	1.0000	0.9375	0.6875
10	1.0000	1.0000	1.0000	0.7187
11	1.0000	1.0000	0.8750	0.7187
12	0.8125	0.8750	0.9375	0.6875
MEAN	0.8385	0.9531	0.9375	0.7786
STD. DEVIATION	0.1176	0.0471	0.0705	0.0974

Table 14. Team efficiency for inspection of defect 2.

TEAM	C1	C2	C3	C4
1	0.9375	1.0000	1.0000	0.9687
2	1.0000	0.8750	0.8750	0.8125
3	0.8750	1.0000	1.0000	1.0000
4	0.9375	0.9375	0.8125	0.8437
5	0.8750	0.9375	0.8750	0.8125
6	1.0000	1.0000	0.9375	0.7812
7	0.8750	0.8125	0.8750	0.7500
8	0.8750	0.9375	0.9375	0.9062
9	0.8750	0.8125	1.0000	0.8437
10	0.9375	0.8750	0.9375	0.8125
11	0.9375	1.0000	0.9375	0.8437
12	0.8125	1.0000	0.9375	0.8750
MEAN	0.9115	0.9323	0.9271	0.8542
STD. DEVIATION	0.0563	0.0728	0.0586	0.0734

test (parametric), and are presented in Table 15. From this note that for detection of defect 1, performance in conditions 2 and 3 was significantly better (.95 and .94) than in conditions 1 and 4 (.84 and .78). For detection of defect 2, performance in conditions 2 and 3 was significantly better (.93 and .93) than in condition 4 (.85), while condition 1 (.91) lies in the intermediate position. Thus conditions 2 and 3 (which required only one mental image of a defect at a time) were better than conditions 1 and 4 (which required two mental images at a time). Time of day had no significant effect for either kind of defects.

Analysis of Type II error

A similar statistical analysis was run for Type II errors. The error scores are tabulated in Tables 18 and 19, for defect 1 and defect 2 respectively. The means for each conditions were tested for significance with all three tests and are presented in Table 20. For errors made in inspection of defect 1, there were no significant differences between performances in conditions 2, 3, 1 and 4 (0.00000, 0.00000, 0.00058, and 0.00173 respectively). Similarly, for inspection of defect 2, there were no significant differences between performances in conditions 2, 4, 3 and 1 (0.00000, 0.00029, 0.00058, and 0.00116 respectively). Time of day was not significant for either kind of defect.

Table 15. Comparison of statistical significance ($p < .05$) between the means of the four conditions for the probability of detection.

Defect No.	Means	Sign Test	Wilcoxon Sign-Rank Test	Duncan's MRT
1	0.9531	C2	C2	C2
	0.9375	C3	C3	C3
	0.8385	C1	C1	C1
	0.7786	C4	C4	C4
2	0.9323	C2	C2	C2
	0.9271	C3	C3	C3
	0.9115	C1	C1	C1
	0.8542	C4	C4	C4

Table 16. Analysis of variance of inspection efficiency of defect 1.

Sources	Sum of Squares	Degrees of Freedom	Mean Sum of Squares	F	Alpha Hat
Condition	0.4946	3	0.1649	6.87	0.00
Period	0.0039	2	0.0020	0.08	0.92
Team	0.0340	3	0.0113	0.47	0.70
C x P	0.0534	6	0.0089	0.37	0.89
P x T	0.1341	6	0.0224	0.93	0.48
C x T	0.1553	9	0.0172	0.72	0.68
C x P x T	0.2904	18	0.0161	0.67	0.81
Error	1.1504	48	0.0240		
Total	2.3160	95			

Table 17. Analysis of variance of inspection efficiency of defect 2.

Sources	Sum of Squares	Degrees of Freedom	Mean Sum of Squares	F	Alpha Hat
Condition	0.1131	3	0.0377	2.94	0.04
Period	0.0119	2	0.0060	0.46	0.62
Team	0.0557	3	0.0186	1.45	0.23
C x P	0.0410	6	0.0068	0.53	0.77
P x T	0.0974	6	0.0162	1.26	0.29
C x T	0.1296	9	0.0144	1.12	0.36
C x P x T	0.0760	18	0.0042	0.32	0.99
Error	0.6152	48	0.0128		
Total	1.1403	95			

Table 18. Inspection data for Type II errors on defect 1.

TEAM	SUBJECT	DEFECT 1 C1	C2	C3	C4
1	1	0.0000	0.0000	0.0000	0.0000
	2	0.0000	0.0000	0.0000	0.0000
2	3	0.0000	0.0000	0.0000	0.0000
	4	0.0000	0.0000	0.0000	0.0000
3	5	0.0000	0.0000	0.0000	0.0000
	6	0.0000	0.0000	0.0000	0.0000
4	7	0.0000	0.0000	0.0000	0.0000
	8	0.0000	0.0000	0.0000	0.0069
5	9	0.0000	0.0000	0.0000	0.0000
	10	0.0000	0.0000	0.0000	0.0000
6	11	0.0000	0.0000	0.0000	0.0069
	12	0.0000	0.0000	0.0000	0.0000
7	13	0.0139	0.0000	0.0000	0.0000
	14	0.0000	0.0000	0.0000	0.0000
8	15	0.0000	0.0000	0.0000	0.0000
	16	0.0000	0.0000	0.0000	0.0000
9	17	0.0000	0.0000	0.0000	0.0000
	18	0.0000	0.0000	0.0000	0.0000
10	19	0.0000	0.0000	0.0000	0.0069
	20	0.0000	0.0000	0.0000	0.0000
11	21	0.0000	0.0000	0.0000	0.0208
	22	0.0000	0.0000	0.0000	0.0000
12	23	0.0000	0.0000	0.0000	0.0000
	24	0.0000	0.0000	0.0000	0.0000
	MEAN	0.00058	0.0000	0.0000	0.00173
STD. DEVIATION		0.0000	0.0000	0.0000	0.00002

Table 19. Inspection data for Type II errors on defect 2.

TEAM	SUBJECT	DEFECT 2 C1	C2	C3	C4
1	1	0.0000	0.0000	0.0000	0.0000
	2	0.0139	0.0000	0.0000	0.0000
2	3	0.0000	0.0000	0.0000	0.0000
	4	0.0000	0.0000	0.0000	0.0000
3	5	0.0139	0.0000	0.0000	0.0000
	6	0.0000	0.0000	0.0000	0.0000
4	7	0.0000	0.0000	0.0000	0.0000
	8	0.0000	0.0000	0.0000	0.0000
5	9	0.0000	0.0000	0.0000	0.0000
	10	0.0000	0.0000	0.0000	0.0000
6	11	0.0000	0.0000	0.0000	0.0000
	12	0.0000	0.0000	0.0000	0.0000
7	13	0.0000	0.0000	0.0000	0.0000
	14	0.0000	0.0000	0.0000	0.0000
8	15	0.0000	0.0000	0.0000	0.0000
	16	0.0000	0.0000	0.0000	0.0000
9	17	0.0000	0.0000	0.0000	0.0000
	18	0.0000	0.0000	0.0139	0.0000
10	19	0.0000	0.0000	0.0000	0.0000
	20	0.0000	0.0000	0.0000	0.0000
11	21	0.0000	0.0000	0.0000	0.0000
	22	0.0000	0.0000	0.0000	0.0000
12	23	0.0000	0.0000	0.0000	0.0069
	24	0.0000	0.0000	0.0000	0.0000
	MEAN	0.00116	0.0000	0.00058	0.00029
STD. DEVIATION		0.00002	0.0000	0.0000	0.0000

Table 20. Comparison of statistical significance ($p < .05$) between the means of the four conditions on Type II errors.

Defect No.	Means	Sign Test	Wilcoxon Sign-Rank Test	Duncan's MRT
1	0.00173	C4	C4	C4
	0.00058	C1	C1	C1
	0.00000	C3	C3	C3
	0.00000	C2	C2	C2
2	0.00116	C1	C1	C1
	0.00058	C3	C3	C3
	0.00029	C4	C4	C4
	0.00000	C2	C2	C2

Table 21. Analysis of variance of Type II errors on defect 1.

Source	Sum of Squares	Degrees of Freedom	Mean Sum of Squares	F	Alpha Hat
Condition	0.000047	3	0.000016	1.99	0.12
Period	0.000001	2	0.000000	0.06	0.93
Team	0.000012	3	0.000004	0.50	0.68
C x P	0.000027	6	0.000005	0.56	0.75
P x T	0.000051	6	0.000008	1.06	0.39
C x T	0.000052	9	0.000006	0.72	0.68
C x P x T	0.000161	18	0.000009	1.11	0.36
Error	0.000384	48	0.000008		
Total	0.000736	95			

Table 22. Analysis of variance of Type II errors on defect 2.

Source	Sum of Squares	Degrees of Freedom	Mean Sum of Squares	F	Alpha Hat
Condition	0.000017	3	0.000006	0.89	0.44
Period	0.000007	2	0.000004	0.54	0.58
Team	0.000017	3	0.000006	0.89	0.44
C x P	0.000029	6	0.000005	0.74	0.61
P x T	0.000029	6	0.000005	0.74	0.61
C x T	0.000109	9	0.000012	1.85	0.08
C x P x T	0.000079	18	0.000004	0.67	0.81
Error	0.000313	48	0.000007		
Total	0.000602	95			

Analysis for the multi-monitor treatment

Condition 4 also can be analyzed as team inspection. Table 23 and 24 show the data for this treatment for defects 1 and 2 respectively.

The efficiencies for the integration mode were calculated by taking the mean performance of the two members in each team.

The efficiencies for the parallel mode were obtained from the (n,n) decision rule in that every inspector must respond "no signal", otherwise the combination of inspectors responds "signal". For the predicted value of the parallel mode, the probability of detection of a signal by the team is given by:

$$P = 1 - \prod_{i=1}^{2} (1 - P_i) \qquad (1)$$

where P_i = the probability of detection of the ith inspector.

The efficiencies for the series mode were obtained from the (y,y) decision rule in that every inspector must respond "yes, a signal has been received" for the combination of inspectors to respond "signal" otherwise, the combination reaponds "no signal". For the predicted value, the probability of detection of a signal by the team is given by

$$P = \prod_{i=1}^{2} P_i \qquad (2)$$

where P_i = probability of detection of the ith inspector.

From Tables 23 and 24 it can be seen that the models have predicted slightly higher values than the observed in the

Table 23. Efficiencies in various modes of team formation for inspection of defect 1.

Team	Integration mode	Parallel mode			Series mode		
	(observed)	(observed)	(predicted) $P=1-\prod_{i=1}^{2}(1-P_i)$	(obs.-pred.)	(observed)	(predicted) $P=\prod_{i=1}^{2}P_i$	(obs.-pred.)
1	0.9375	0.9375	1.0000	-0.0625	0.8750	0.8750	0.0000
2	0.8750	1.0000	0.9883	0.0117	0.7500	0.7617	-0.0117
3	0.9062	1.0000	0.9922	0.0078	0.8125	0.8203	-0.0078
4	0.6562	0.8125	0.8906	-0.0781	0.5000	0.4219	0.0781
5	0.8125	1.0000	0.9687	0.0313	0.6250	0.6563	-0.0313
6	0.7500	0.8750	0.9531	-0.0781	0.6250	0.5469	0.0781
7	0.7187	0.8750	0.9219	-0.0469	0.5625	0.5156	0.0469
8	0.8750	1.0000	0.9844	0.0156	0.7500	0.7656	-0.0156
9	0.6875	0.9375	0.9180	0.0195	0.5000	0.4570	0.0430
10	0.7187	0.9375	0.9297	0.0078	0.5000	0.5078	-0.0078
11	0.7187	0.8750	0.9219	-0.0469	0.5625	0.5156	0.0469
12	0.6875	1.0000	1.0000	0.0000	0.3750	0.3750	0.0000
Mean	0.7786	0.9375	0.9557	-0.0182	0.6197	0.6015	0.0182
Std. Deviation	0.0974	0.0653	0.0381	0.0408	0.1495	0.1677	0.0381

Table 24. Efficiencies in various modes of team formation for inspection of defect 2.

Team	Integration mode	Parallel mode			Series mode		
	(observed)	(observed)	(predicted) $P=1-\prod_{i=1}^{2}(1-P_i)$	(obs.-pred.)	(observed)	(predicted) $P=\prod_{i=1}^{2} P_i$	(obs.-pred.)
1	0.9687	1.0000	1.0000	0.0000	0.9375	0.9375	0.0000
2	0.8125	1.0000	1.0000	0.0000	0.6250	0.6250	0.0000
3	1.0000	1.0000	1.0000	0.0000	1.0000	1.0000	0.0000
4	0.8437	0.9375	0.9766	-0.0391	0.6875	0.7109	-0.0234
5	0.8125	0.9375	0.9687	-0.0312	0.6875	0.6563	0.0392
6	0.7812	0.9375	0.9609	-0.0234	0.6250	0.6016	0.0234
7	0.7500	0.8750	0.9414	-0.0664	0.6250	0.5586	0.0664
8	0.9062	1.0000	0.9922	0.0078	0.8125	0.8203	-0.0078
9	0.8437	0.8750	0.9844	-0.1094	0.6875	0.7031	-0.0156
10	0.8125	1.0000	0.9688	0.0312	0.6250	0.6563	-0.0313
11	0.8437	1.0000	0.9844	0.0156	0.6875	0.7031	-0.0156
12	0.8750	1.0000	0.9688	0.0312	0.7500	0.7656	-0.0156
Mean	0.8542	0.9635	0.9788	-0.0153	0.7292	0.7282	-0.0010
Std. Deviation	0.0734	0.0496	0.0181	0.0414	0.1259	0.1331	0.0282

parallel mode and slightly lower values in the series mode. For inspection of defect 1, the mean error for the parallel mode was -1.8% and for the series mode it was +1.8%. For inspection of defect 2, the mean error for the parallel mode was -1.5% and for the series mode it was +.1%.

The Pearson product-moment correlation coefficients for the expected and observed values were all significant at the 5% level (see Table 25).

Signal Detection Theory Analysis

One purpose in this study is to analyze inspection performance with the help of conventional measures of the percentages of detected signals and with the help of measures derived from the theory of signal detection.

Of the 24 subjects in this study, only nine committed false alarms (Type II errors) in condition 4. The remaining 15 subjects did not commit any false alarms, indicating an infinite value of d' for them. Other workers in this field have circumvented this difficulty by calling a zero frequency of a particular error 'less than ½ an error' and thus replacing each zero by ½ to obtain a lower bound of d' (Jerison et. al., 1965).

Table 25. Pearson product-moment correlation coefficients ($p < .05$) between the observed and expected inspection efficiencies on condition 4.

Mode / Defect No.	Series	Parallel
1	0.96	0.81
2	0.98	0.60

Using a similar kind of treatment, let us treat the detection of all 8 signals (10% of the 80 slides were defective) as being at least as great as 7.5 our of 8 correct reports. This would give the following lower limit on the probability of detection:

$$P\ (sn,Y) = 7.5/8 = 0.9375$$

Similarly, the emission of no false alarms may be treated as the emission of fewer than 0.5 false alarms, and the maximum false alarm probability is therefore:

$$P\ (n,Y) = 0.5/(80-8) = 0.00694$$

From normal probability distribution tables, the value of P(n,Y) corresponds to $z_n = 2.46$ and the value of P(sn,Y) corresponds to $z_{sn} = -1.53$. Therefore $d' = z_n - z_{sn} = 2.46 - (-1.53) = 3.99$. The ordinate for z_{sn} is 0.1238 and the ordinate for z_n is 0.0196; therefore $\beta = \frac{.1238}{.0196} = 6.31$.

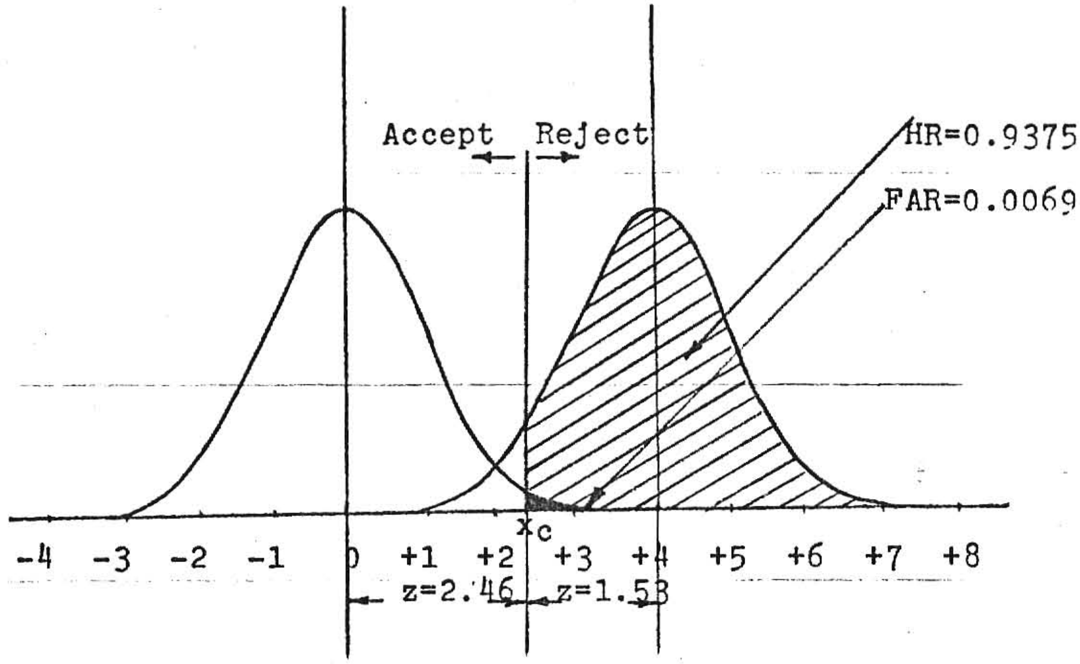
Hence at these limiting probabilities we obtain (Figure 7)

$$d' = 3.99$$

$$\beta = 6.31$$

Thus, for our best subjects we have a minimum value of d' of about 4. This is a numerical statement of a "readily detectable signal". For the other subjects the minimum value of d' was 3.73.

Table 26 and 27 give the values of d' (the SDT measures of index of detectability) on defect 1 and defect 2 respectively for all four conditions. Tables 28 and 29 give the values



$$d' = z_{FAR} - z_{HR} = 2.46 - (-1.53) = 3.99$$

$$\beta = \frac{f_{d1}(x_c)}{f_{d_2}(x_c)} = \frac{.1238}{.0196} = 6.31$$

Figure 7 Calculation of d' and β for the 15 best subjects.

Table 26. SDT measure of the index of detectability d' for inspection of defect 1.

TEAM	SUBJECT	DEFECT 1 C1	C2	C3	C4
1	1	1.78	3.61	3.99	3.85
	2	3.99	3.99	3.99	4.56
2	3	3.61	3.61	3.99	3.59
	4	3.61	3.99	3.99	4.23
3	5	3.61	3.99	3.99	3.85
	6	3.99	3.99	3.99	4.23
4	7	3.99	3.99	3.13	2.86
	8	3.61	3.61	3.61	3.13
5	9	3.61	3.99	3.61	3.85
	10	3.13	3.99	3.99	3.38
6	11	2.78	3.13	3.99	2.77
	12	3.99	3.99	3.99	3.85
7	13	1.88	3.61	3.61	3.19
	14	3.99	3.99	3.99	3.38
8	15	3.13	3.61	3.13	3.85
	16	3.61	3.99	3.61	3.85
9	17	3.61	3.99	3.99	3.59
	18	2.78	3.99	3.61	2.86
10	19	3.99	3.99	3.99	2.77
	20	3.99	3.99	3.99	3.59
11	21	3.99	3.99	3.61	3.03
	22	3.99	3.99	3.61	3.19
12	23	3.99	3.99	3.99	4.56
	24	2.78	3.13	3.61	2.38
	MEAN	3.48	3.84	3.79	3.52
STD. DEVIATION		0.66	0.27	0.27	0.58

Table 27. SDT measure of the index of detectability d' for inspection of defect 2.

		DEFECT 2			
TEAM	SUBJECT	C1	C2	C3	C4
1	1	3.99	3.99	3.99	4.56
	2	3.35	3.99	3.99	4.23
2	3	3.99	3.13	3.13	3.02
	4	3.99	3.99	3.99	4.56
3	5	2.87	3.99	3.99	4.56
	6	3.99	3.99	3.99	4.56
4	7	3.61	3.99	3.99	3.59
	8	3.99	3.61	2.78	3.85
5	9	3.61	3.61	3.61	3.38
	10	3.61	3.99	3.61	3.85
6	11	3.99	3.99	3.61	3.19
	12	3.99	3.99	3.99	3.85
7	13	3.13	3.13	3.61	3.19
	14	3.99	3.61	3.61	3.59
8	15	3.13	3.99	3.99	4.23
	16	3.99	3.61	3.61	3.85
9	17	3.99	3.61	3.99	4.23
	18	3.13	3.13	3.73	3.38
10	19	3.61	3.13	3.99	3.38
	20	3.99	3.99	3.61	3.85
11	21	3.99	3.99	3.99	3.38
	22	3.61	3.99	3.61	4.23
12	23	3.99	3.99	3.61	3.61
	24	2.78	3.99	3.99	3.85
	MEAN	3.68	3,77	3.75	3.83
STD. DEVIATION		0.40	0.33	0.31	0.48

Table 28. SDT measure of the decision criterion point β inspection of defect 1.

		DEFECT 1			
TEAM	SUBJECT	C1	C2	C3	C4
1	1	16.29	10.56	6.31	19.80
	2	6.31	6.31	6.31	6.80
2	3	10.56	10.56	6.31	25.89
	4	10.56	6.31	6.31	11.83
3	5	10.56	6.31	6.31	19.80
	6	6.31	6.31	6.31	11.83
4	7	6.31	6.31	16.29	37.89
	8	10.56	10.56	10.56	16.21
5	9	10.56	6.31	10.56	19.80
	10	16.29	6.31	6.31	30.55
6	11	19.44	16.29	6.31	19.35
	12	6.31	6.31	6.31	19.80
7	13	10.68	10.56	10.56	34.04
	14	6.31	6.31	6.31	30.55
8	15	16.29	10.56	16.29	19.80
	16	10.56	6.31	10.56	19.80
9	17	10.56	6.31	6.31	25.89
	18	19.44	6.31	10.56	37.89
10	19	6.31	6.31	6.31	19.35
	20	6.31	6.31	6.31	25.89
11	21	6.31	6.31	10.56	12.91
	22	6.31	6.31	10.56	34.04
12	23	6.31	6.31	6.31	6.80
	24	19.44	16.29	10.56	36.47
	MEAN	10.62	8.03	8.56	22.62
STD. DEVIATION		4.75	3.08	3.11	9.36

Table 29. SDT measure of the decision criterion point β for inspection of defect 2.

TEAM	SUBJECT	DEFECT 2 C1	C2	C3	C4
1	1	6.31	6.31	6.31	6.80
	2	5.80	6.31	6.31	11.83
2	3	6.31	16.29	16.29	36.47
	4	6.31	6.31	6.31	6.80
3	5	8.95	6.31	6.31	6.80
	6	6.31	6.31	6.31	6.80
4	7	10.56	6.31	6.31	25.89
	8	6.31	10.56	19.44	19.80
5	9	10.56	10.56	10.56	30.55
	10	10.56	6.31	10.56	19.80
6	11	6.31	6.31	10.56	34.04
	12	6.31	6.31	6.31	19.80
7	13	16.29	16.29	10.56	34.04
	14	6.31	10.56	10.56	25.89
8	15	16.29	6.31	6.31	11.83
	16	6.31	10.56	10.56	19.80
9	17	6.31	10.56	6.31	11.83
	18	16.29	16.29	3.46	30.55
10	19	10.56	16.29	6.31	30.55
	20	6.31	6.31	10.56	19.80
11	21	6.31	6.31	6.31	30.55
	22	10.56	6.31	10.56	11.83
12	23	6.31	6.31	10.56	10.51
	24	19.44	6.31	6.31	19.80
	MEAN	9.08	8.86	8.75	20.10
STD. DEVIATION		4.09	3.80	3.61	9.86

of β (the SDT measure of the decision criterion) on defect 1 and defect 2 respectively for all four conditions.

We may note some general attributes of the SDT measures over all four conditions. The range of d' was from 5.80 to 37.89.

The means of d' for inspection of defect 1 were 3.48, 3.84, 3.79 and 3.52 for condition 1, 2, 3 and 4 respectively; for defect 2, they were 3.68, 3.77, 3.75 and 3.88 respectively.

The means of β for inspection of defect 1 were 10.62, 8.03, 8.56 and 22.62 for conditions 1, 2, 3 and 4 respectively; for defect 2, they were 9.08, 8.86, 8.75 and 20.10 respectively.

The reason for not being able to perform an exact SDT analysis with the observers data was their failure to give false alarms. This was true for 15 of our 24 observers. Our inability to obtain exact SDT measures probably biased our data towards values of both d' and β that are low.

Table 30 shows the Pearson product moment correlations of d' with the inspection efficiency 'A'. For inspection of defect 1, the correlations were 0.997, 0.998, 0.997 and 0.961 for conditions 1, 2, 3 and 4 respectively. For inspection of defect 2, the correlations were 0.983, 0.998, 0.983 and 0.978 for conditions 1, 2, 3 and 4 respectively. All these correlations are significant at the 1% level.

There was no significant correlation found between the

Table 30. Pearson product-moment correlation coefficients ($p < .01$) between the detectability index d' and inspection efficiencies.

Correlation Coefficient	Defect 1	Defect 2
$r_{d',C1}$	0.997	0.983
$r_{d',C2}$	0.998	0.998
$r_{d',C3}$	0.997	0.983
$r_{d',C4}$	0.961	0.978

false alarms and the decision criterion point β . Only 9 out of our 24 subjects made a false alarm in one of the four conditions that they were inspecting. Hence only in these 9 cases do we have the exact values of β . For the remaining 15 situations, the β values were obtained by treating 'no false alarms' as 'maximum of half a false alarm'. Hence there was no correlation found.

C O N C L U S I O N S   A N D   D I S C U S S I O N

From Table 11, for inspection of defect 1, conditions 2 and 3 are better than conditions 1 and 4 using inspection efficiencies as the criterion. No statistical significance was observed for the false alarms. Since conditions 2 and 3 showed no false alarms (whereas conditions 1 and 4 had some false alarms) we conclude that, for detection of defect 1, conditions 2 and 3 are better than conditions 1 and 4.

For inspection of defect 2, conditions 2, 3 and 1 were found better than condition 4 using inspection efficiencies as the criterion. No statistical significance was observed for the false alarms. Since condition 2 showed no false alarms (whereas conditions 1, 3 and 4 had a few false alarms), we conclude that, for detection of defect 2, condition 2 is better than conditions 1, 3 and 4.

Condition 4 was found to be the worst of all four conditions in terms of probability of detection (criterion A) as well as for Type II errors. But, when it was analyzed in the parallel mode (where at least one inspector was required to respond), the means improved considerably. In this parallel mode the means were not significantly different than the means for condition 2. (For inspection of defect 1, the means (of criterion A) for condition 2 and the parallel mode of condition 4 respectively were .95 and .94; for inspection of defect 2, the means were .93 and .98). However, the parallel mode also had the maximum false alarms.

On the other hand, the series mode (which has no practical applications but is only of theoretical interest) showed no false alarms, but had the lowest probability of detection. (Its means were .60 and .73 for defect 1 and defect 2 respectively).

A very high correlation was found between the detectability index d' and the probability of detection 'A'.

By allotting a set of payoffs consisting of the value of making correct decisions relative to the costs of making errors, Sheehan and Drury (1971) were able to discern deviant inspector behavior during a hook inspection task. One of the inspectors had a 'hit rate' similar to those of the other inspectors, but had a false alarm rate that was much higher (and d' that was much lower) than those of the other inspectors. An interview indicated that the inspector was rejecting acceptable hooks on the basis of an irrelevant characteristic. Payoffs could be explicit (such as actual penalties for falsely rejecting too many conforming items) or implicit (such as potential safety hazards to consumers if defective items are not detected).

Since in this study no payoffs were set, the advantages of SDT could not be brought out.

R E F E R E N C E S

Baker, E.M., Signal detection theory analysis of quality control inspector performance. Journal of Quality Technology, 1975, Vol. 7, No. 2, 62-71.

Baker, R.A., Rogerware, J., and Sipowicz, R., Signal detection by multiple monitors. Psychology Record, 1962, Vol. 12, 113-137.

Bergum, B.O., and Lehr, D.J., Vigilance performance as a function of paired monitoring. Journal of Applied Psychology, 1962, Vol. 46, No. 5, 341-343.

Colquhoun, W.P., Vigilance and the inspection problem. Personnel Management, 1957, Vol. 39, No. 341, 163-171.

Drury, C.G., The human factor in industrial inspection. Quality Progress, 1974, Vol. 7, No. 12, 14-19.

Drury, C.G., and Sheehan, J.J., Ergonomic and economic factors in an industrial inspection task. The International Journal of Production Research, 1969, Vol. 7, No. 4, 333-341.

Egan, J.P., Signal Detection Theory and R.O.C. Analysis. Academic Press, 1975.

Green, D.M., and Swets, J.A., Signal Detection Theory and Psychophysics. Wiley, 1966.

Hornseth, J.P., and Davis, J.H., Individual and two-man team target finding performance. Human Factors, 1967, Vol. 9, No. 1, 39-43.

Jackson, J.E., Effects of inspection errors on waste and quality of outgoing products. Industrial Quality Control, 1957, Vol. 14, No. 6, 5-7.

Jerison, H.J., Pickett, R.M., and Stenson, H.H., The elicited observing rate and decision processes in vigilance. Human Factors, 1965, Vol. 7, No. 2, 107-128.

Juran, J.M., Gryna, F.M., and Bingham, R.S., Quality Control Handbook. McGraw-Hill, 1974, 3rd Edition, 1261-1263.

Lion, J.S., Richardson, E., Weightman, D., and Brown, R.C., The influence of the visual arrangement of material, and of working singly or in pairs, upon performance at simulated industrial inspection. Ergonomics, 1975, Vol. 18, No. 2, 195-204.

McCormick, E.J., Human Factors Engineering. McGraw-Hill, 1964.

McFarling, L.H., and Heimstra, N.W., Pacing, product complexity, and task perception in simulated inspection. Human Factors, 1975, Vol. 17, No. 4, 361-367.

McKenzie, R.M., On the accuracy of inspectors, Ergonomics, 1958, Vol. 1, 250-272.

Morgan, B.B., and Alluisi, E.A., On the inferred independence of paired watchkeepers. Psychonomic Science, 1965, Vol. 2, 161-162.

Morrissette, J.O., Hornseth, J.P., and Shellar, K., Team organization and monitoring performance. Human Factors. 1975, Vol. 17, No. 3, 296-300.

Quick, J.H., Duncan, J.H., and Malcolm, J.A., Work-Factor Time Standards. McGraw-Hill, 1962.

Schafer, T.H., Detection of a signal by several observers. (Report No. 101), U.S. Naval Electronic Laboratory, 1949.

Schlegel, R.E., Boardman, D.W., and Purswell, J.L., A comparison of single and multiple inspection systems. Proceedings, 17th Human Factors Society meeting, 1973, 429-432.

Sheehan, J.J., and Drury, C.G., The analysis of industrial inspection. Applied Ergonomics, 1971, Vol. 2, No. 2, 74-78.

Smith, G.L., Signal detection theory and industrial inspection. Proceedings, 16th Human Factors Society meeting, 1972, 284-290.

Swets, J.A., Decision processes in perception. Psychological Review, 1961, Vol. 68, No. 5, 301-340.

Waag, W.L., The prediction of team monitoring performance under conditions of varied team size and decision rules. Proceedings, 16th Human Factors Society meeting, 1972, 140-143.

Waiker, A.M., Quality improvement using multi-detector inspection systems. Unpublished M.S. thesis, University of Oklahoma, 1973.

Wald, A., Statistical Decision Functions. Wiley, 1950.

Wiener, E.L., Note on the probability of detection of signal by multiple monitors. Psychological Records, 1963, Vol. 13, 79-81.

Wiener, E.L., The performance of multi-man monitoring teams. Human Factors, 1964, Vol. 6, 179-184.

APPENDIX 1

Signal Detection Theory

Signal detection theory blends into a single model many dissimilar factors that affect the detection and recognition of signals embedded in noise. This model emphasizes that the decision of an observer may depend upon

(1) the information available to the observer before the presentation of the stimulus,

(2) the information content of the stimulus,

(3) the properties of the sensory analyzer, and

(4) the consequences of each decision.

When detection performance is imperfect, it is never assumed that the observer "detects the signal". Rather, it is assumed that the observer receives an input, and this input corresponds to, a unique value of liklihood ratio. Then, given other factors, such as the prior probability of signal existence, the observer makes a decision "Yes, the odds favour the event <u>signal plus noise</u>" or "No, the odds favour the event <u>noise alone</u>".

In the fundamental detection problems with only two events and two responses, there are four event-response conjunctions. Let sn and n stand for the two events, and let Y and N stand, respectively, for the responses of "Yes" and "No". Then, the conjunction (sn,Y) is called a correct acceptance, or a hit, and its complement (sn,N) is an incorrect rejection, a miss, or a false dismissal. The conjunction (n,N) is called a correct rejection, and its complement (n,Y) is called an incorrect acceptance, a false alarm, or a false report.

With only two events, the sum of the probabilities is unity.

$$P(sn) + P(n) = 1 \qquad (A.1.1)$$

where P(sn) = the probability that a signal is present.
and P(n) = the probability of noise alone.

The hit rate, P(Y/sn) is the conditional probability of the response of "yes" given that the signal occurred, and its complement is P(N.sn) = 1- P(Y/sn). The false alarm rate, P(Y/n), is the conditional probability of the response of "yes", given that the signal did not occur, and its complement is P(N/n) = 1 - P(Y/n). Thus, we have

$$P(Y/sn) + P(N/sn) = 1 \qquad (A.1.2)$$

$$P(Y/n) + P(N/n) = 1 \qquad (A.1.3)$$

The two conditional probabilities, the hit rate and the false-alarm rate, taken together specify the degree to which the observer can distinguish between the two types of events, sn and n.

The performance of an observer in the single-interval experiment is evaluated in terms of the probabilities of the event-response conjunctions. These probabilities are based upon the prior probability of signal existence, P(sn), and the two conditional probabilities, P(Y/sn) and P(Y/n). Thus we have:

(1) Correct acceptance : $P(sn,Y) = P(Y/sn).P(sn)$

(2) Incorrect rejection : $P(sn,N) = \left[1-P(Y/sn)\right].P(sn)$

(3) Correct rejection : $P(n,N) = \left[1-P(Y/n)\right].P(n)$

(4) Incorrect acceptance : $P(n,Y) = P(Y/n).P(n)$

(A.1.4)

Here we should note that P(sn,Y) is the probability of a "hit" whereas P(Y/sn) is the "hit rate". The four event-response probabilities sum to one, and they are called descriptive probabilities, because they describe the behavior of the detector.

The three descriptive probabilities, P(sn), P(Y/sn), and P(Y/n), are used in the definitions of the probability of a correct decision, P(C), and of the probability of a response of "yes", P(Y). Thus

$$P(C) = P(Y/sn).P(sn) + \left[1-P(Y/n)\right]\left[1-P(sn)\right] \quad (A.1.5)$$

$$P(Y) = P(Y/sn).P(sn) + P(Y/n)\left[1-P(sn)\right] \quad (A.1.6)$$

Our discussion will be concentrated upon simple cases in which the event can be represented to an observer by a single number x, which is a specific value of the random variable X. On a given trial, the value x constitutes the evidence upon which the decision is based; therefore x is sometimes referred to as a test-statistic. It is assumed that each of the two events, sn and n, has associated with it a univariate probability distribution, discrete or continuous.

If the random variable is discrete, then P(x/sn) is the conditional probability of x, given the event sn; P(x/n) is the corresponding conditional probability of this same value of x, given the event n. If the random variable is continuous then the corresponding probability elements are f(x/sn)dx and f(x/n)dx, where f(x/sn) and f(x/n) are the probability densities associated, respectively, with the two distributions of X.

For discrete distributions, the likelihood ratio, $L_{sn}(x)$ is : P(x/sn)/P(x/n), where x is a specific value of the discrete random variable X. Similarly $L_{sn}(x) = 1/L_{n}(x)$; for further discussion let the symbol L(x) stand for $L_{sn}(x)$. Note that, even though the values of x may be discrete, the numerical values of P(x/sn), P(x/n) and L(x), are, in general continuous variables. Also, L(x) is not a probability, and it may range from zero to infinity.

For continuous distributions, the likelihood ratio, L(x), is f(x/sn)/f(x/n), where x is a specific value of the continuous random variable X. The functional relationship L(x)=f(x/sn)/f(x/n) is called the likelihood-ratio function.

Two of the event-response conjunctions, (sn,Y) and (n,N), are defined as "correct" and the other two conjunctions are called "incorrect". A reasonable goal for a rational observer is to maximize the probability of a correct decision. We will call this observer a max-P(C) observer. The decision rule for the max-P(C) observer is : If $L(x) > P(n)/P(sn)$, say "yes", otherwise say "no". With only two events and two decisions, the max-P(C) observer also minimizes "decision errors". Thus a criterion is adopted such that $[P(N/sn).P(sn) + P(Y/n).P(n)]$ will be as small as possible. In the literature on detection theory, the max-P(C) observer is sometimes called Siegert's observer. The decision rule based upon likelihood ratio divides the x axis into two types of intervals, those over which $L(x) > L(x_o)$ and those over which $L(x) \leq L(x_o)$, where $L(x_o)$ is the cut-off point. Thus the stimuli associated with the two events, sn and n, are mapped onto two responses by the appropriate division of the x axis into two types of intervals.

We can now mark the similarities between the SDT model and the statistical test of the null hypothesis. The quality control inspector is conceptualized as a statistical hypothesis tester, gathering data from each target in order to decide between the following null (H_o) and alternative (H_1) hypothesis:

H_o: observation x sampled from internal distribution associated with conforming targets,

H_1: observation x sampled from internal distribution associated with defective targets.

The inspector's observation, x, is a summarization of the data from the characteristics of the target. Its value is compared to the criterion ("critical") value of x and H_o

is either rejected or accepted.

The probability P(n,N) of correctly rejecting a defective disk is analogous to $1-\beta$, the power of a statistical test in detecting H_1 when it is true. The probability of incorrect rejection P(sn,N) is analogous to the probability, α of the Type I error of falsely rejecting H_o. In the sampling plan operating characteristic curves, α is known as the producer's risk.

The probability of incorrect acceptance P(n,Y) in SDT has its counterpart in the Type II error in statistical hypothesis testing of failing to reject a false H_o and has probability β. It is also known as the consumer's risk in acceptance sampling inspection. The SDT criterion, x_c, is analogous to c in sampling inspection where c is the number of sampled defective items which must be exceeded in order to reject the entire lot as defective.

The criterion value of the likelihood ratio, x_c, reflects the inspector's own standard of what is acceptable and unacceptable product quality. Ideally the inspector would want all of his decisions to be correct. Figures 8 a, b, c show that it is unreasonable to expect perfect performance. There is no criterion that the inspector can set that would simultaneously allow HR to be maximized and FAR to be minimized. It can be seen that the HR and FAR covary directly, becoming smaller as x_c increases (Figure a). There has to be some

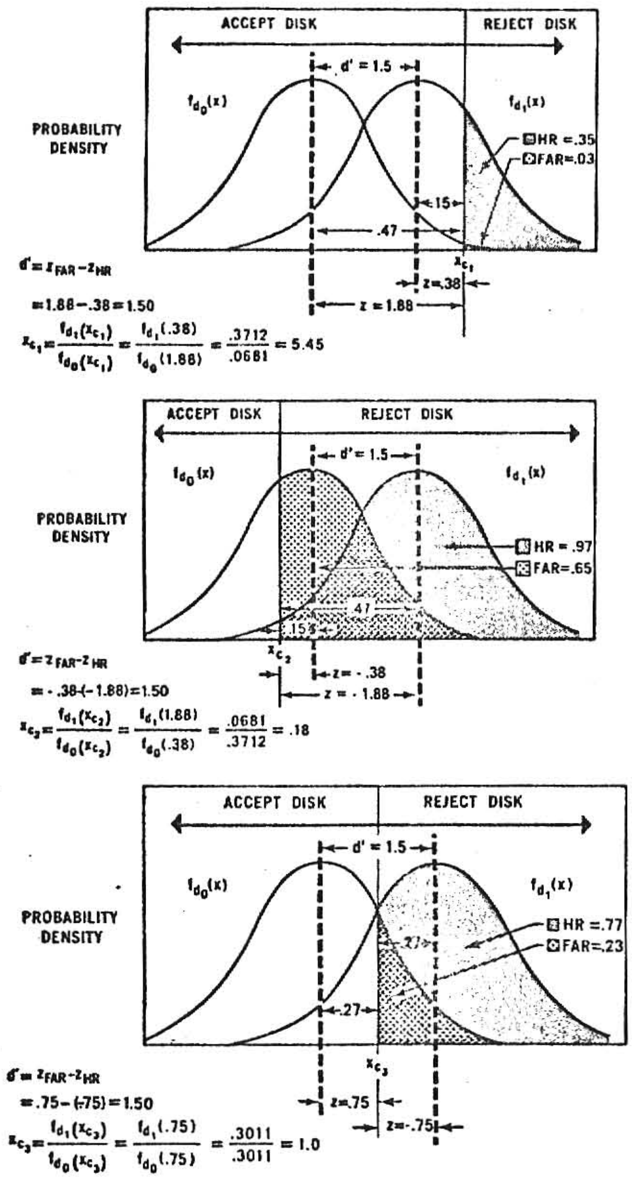


FIGURE 8a. Calculation of d' and x_c at a Strict Response Criterion.

FIGURE 8b. Calculation of d' and x_c at a Lax Response Criterion.

FIGURE 8c. Calculation of d' and x_c at a Moderate Response Criterion.

tradeoff of false alarms to achieve hits.

Figure 8 a, b, c illustrates the two probability density functions and d', the distance between their means, measured in standard deviation units. Increases in the inspector's ability to distinguish d_1 from d_o are indicated by larger values of d'.

Hit and false alarm rates co-vary directly as the inspector changes the value of the decision criterion. The plot of HR against FAR as x_c varies, for a given d', is known as an isosensitivity or receiver operating characteristic (ROC) curve. Since the scales for both the ordinate and abscissa range from 0 to 1, the sample space represented by these dimensions is called the unit square.

Every point on the positive diagonal of the unit square represents inspector inability to detect defective disks since HR=FAR, indicating that the two distributions overlap completely. Points below the positive diagonal could indicate deliberate inspector sabotage, or misunderstanding of what constitutes a defect.

If the two probability density functions are normal and have equal variances, the ROC curve will be symmetrical around the negative diagonal of the unit square.

An exhaustive coverage of the Signal Detection Theory is outlined in Swets (1964), Egan (1975) and Baker (1975).

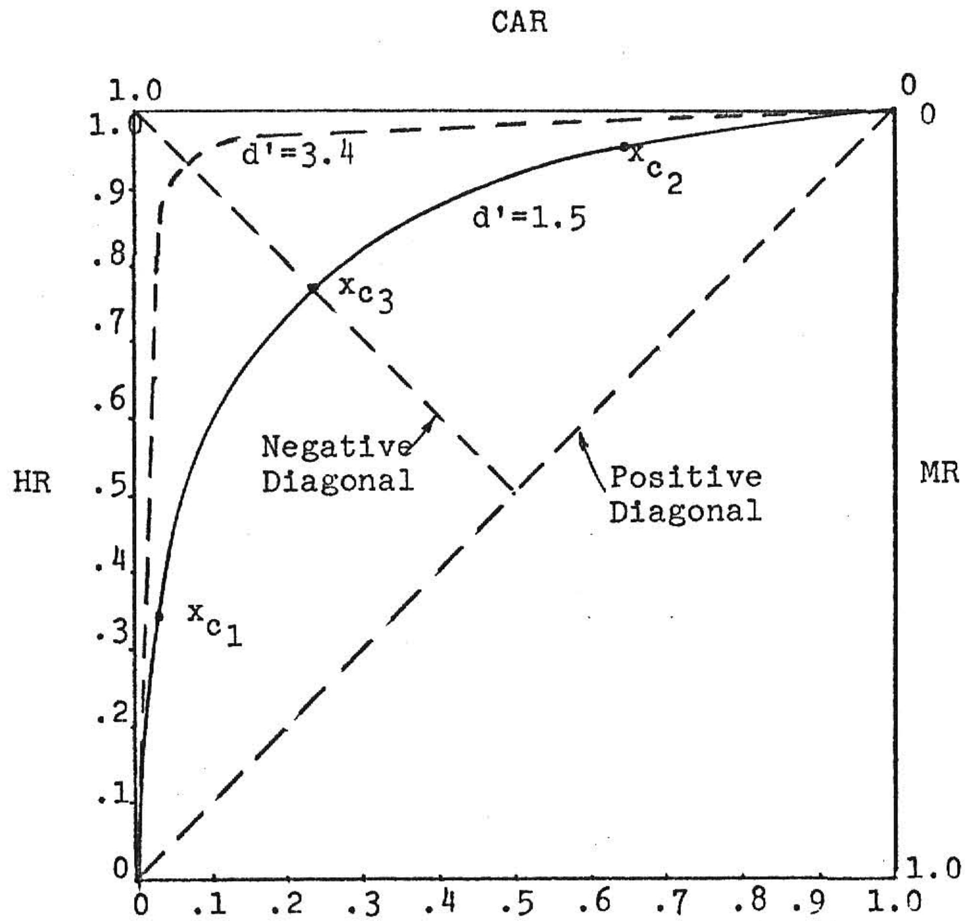


Figure 9 Receiver operating characteristic curves plotted in the unit square.

Let us consider the detection situation in which different sn-events are randomly mixed with the noise-alone event. Here, we will examine the case when there are only two sn-events as well as the n-event.

There are three types of trials: 1) On n-trials, the observer receives a value of X: $f(x/sn_1)$, 2) On sn_1-trials, the observer receives a value of X: $f(x/sn_1)$, 3) On sn_2-trials, the observer receives a value of X: $f(x/sn_2)$. After each observation interval, the observer is required only to state whether a signal was present. This situation restricts the possible decision goals of the observer. For example, the observer cannot simultaneously maximize P(C) for each of the two signal-plus-noise events. In general, the cutoff that will maximize P(C) for sn_1-trials will be different from the cutoff that will maximize P(C) from sn_2-trials. However, with known prior probabilities of the three events, the observer can maximize P(C) considered over all trials. Let us assume that the prior probabilities are: P(n), $P(sn_1)$ and $P(sn_2)$. The density of X for n-trials is f(x/n). The density of X for sn-trials is the weighted combination of the two densities of X, one for sn_1 and one for sn_2. Let $C_1= P(sn_1)/\ P(sn_1)+P(sn_2)$ and $C_2= 1-C_1$. Then the density of X on sn-trials, irrespective of sn_1 and sn_2, is:

$$g(x/sn)+ C_1f(x/sn_1)+ C_2f(x/sn_2) \qquad (A.1.7)$$

Then;

$$L(x) = \frac{g(x/sn)}{f(x/n)} = C_1 L(x/n, sn_1) + C_2 L(x/n, sn_2) \tag{A.1.8}$$

The likelihood ratio of the observation x for two signal-plus-noise events is the weighted combination of the two likelihood ratios, one for sn_1-event and one for sn_2-event. If x is monotonic increasing with each L(x) separately, then x is monotonic increasing with the weighted combination of the two L(x)'s. Finally the max-P(C) observer selects a cutoff C on the x axis such that,

$$L(C) = P(n)/\ P(sn_1) + P(sn_2) \tag{A.1.9}$$

APPENDIX 2

Derivation of Multiple Monitor Equations

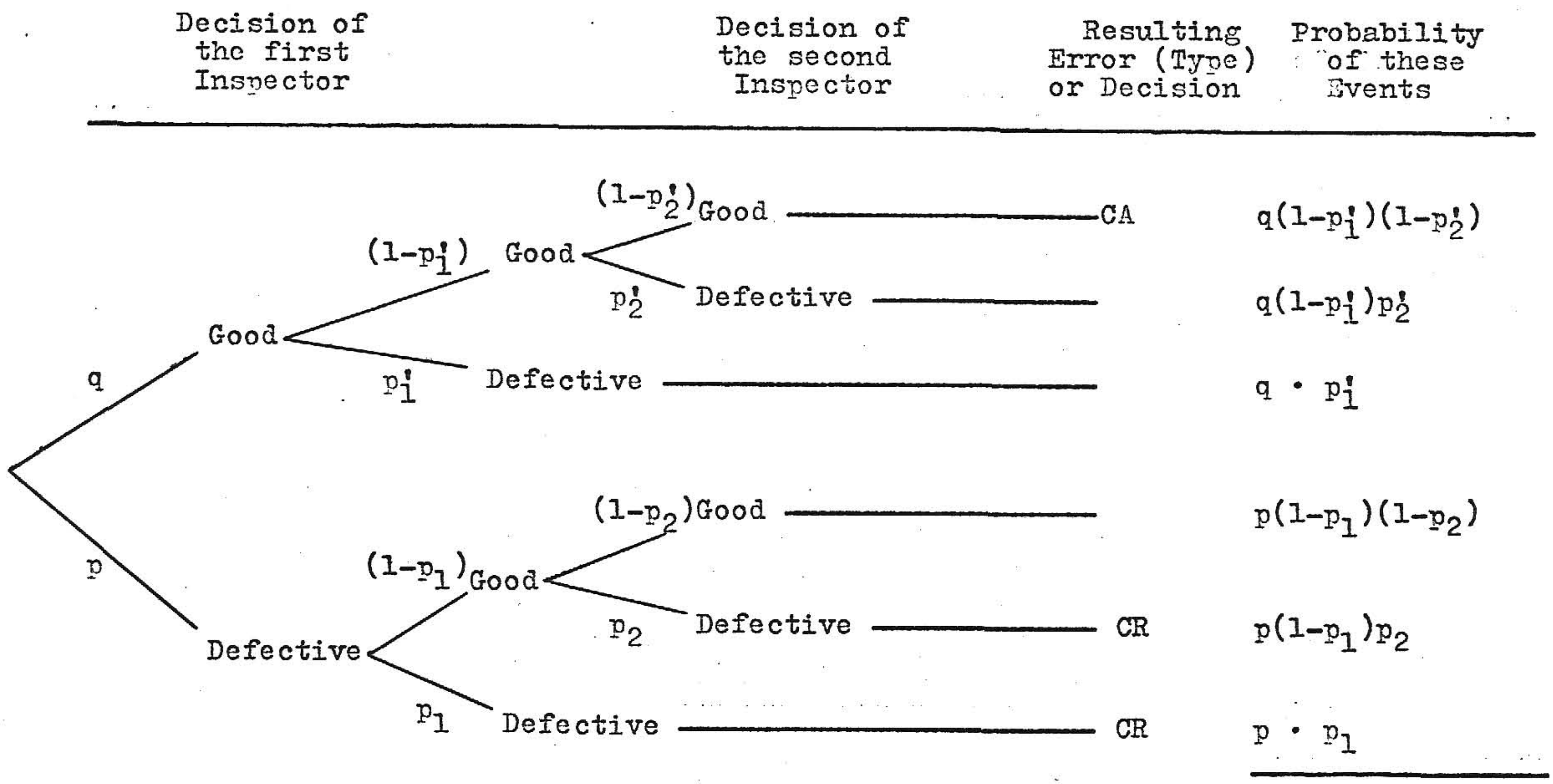
This chapter is devoted to development of a mathematical model for the problem of team inspection, based on the assumption of independence for individual responses as outlined by Waiker (1973).

Basic principles of probability theory will be used to determine the expected outcomes when a team is detecting defects, under the two decision rules, viz. series and parallel.

Two Inspector Team-Parallel Task

Condition 4 has also been analyzed in parallel mode.

Consider a tree diagram for a two-inspector group working on a parallel task, as shown in Figure 10. Referring to Figure 10, the true quality of the target can have two alternatives, good or defective. The first inspector can classify it as either a good or defective target under both alternatives. If he classifies the target as a defective, there is no further inspection and the result is either an α (Type II) error or a correct rejection (CR). If he classifies it as a good target, then the final result will depend on whether the second inspector classifies it as a defective or good item. All the errors and correct acceptances/rejections are listed in the tree diagram of Figure 10. From this it can be inferred that,



$P(\text{Events})=1.0$

p = percent defectives

$q = (1-p)$

p_i = probability of detection of i^{th} inspector

p_i' = probability of false alarm of i^{th} inspector

Figure 10. Tree Diagram Illustrating Probabilities of the Various Decisions of the Group of Inspectors

P(Event) = sum of the probabilities of the branches resulting in the same event.

$P(\text{Detection}) = P(CR) = p.p_1 + p(1-p_1)p_2 = p(p_1-p_1p_2+p_2)$

$P(\text{False Alarm}) = P(\alpha) = q.p_1' + q(1-p_1')p_2'$

$P(\text{Incorrect Acceptance}) = p(\beta) = p(1-p_1)(1-p_2)$

$P(\text{Correct Acceptance}) = P(CA) = q(1-p_1')(1-p_2')$

Thus, when the individual detection probabilities are found for the individual members of the group, their performance based on this model can be evaluated for a given level of incoming quality.

Two Inspector Team-Series Task.

Condition 4 has also been analyzed in series mode.

In this type of task, every detector is presented with all n defects. All detectors report the defect if detected. Thus, here, with the same input, the detectors seem to be working in a series system.

Consider the case of two detectors (inspectors) in the system. Let the probability that the first inspector detects a defect = p_1, and the probability that the second inspector detects a defect = p_2. Therefore, the probability that both detect the defect = $p_1 . p_2$. In a series type task the decision rule is used that every inspector must register a defect for the product presented to be rejected as a defective. Similarly, a good product is accepted only if it is accepted by all the inspectors in a group. Therefore, if there are r inspectors in the group, the probability that the defective is rejected is

$$P_r = p_1 \cdot p_2 \cdot p_3 \cdots \cdot p_r = \prod_{i=1}^{r} p_i$$

If there are N good and n defects in the lot, let D be the total detection of defectives and d be defectives passing in the outgoing lot,

$$D = \begin{bmatrix} \text{Probable number of} \\ \text{detections expected} \end{bmatrix} = n.P_r = n(p_1 \cdot p_2 \cdot p_3 \cdots \cdot p_r)$$

and the defect outgoing = $d = n - D = n - n(p_1 p_2 \cdots p_r)$

therefore $d = n(1 - \prod_{i=1}^{r} p_i)$

Similarly, the number of expected false alarms are

$$\text{Expected No. of False Alarms} = N.(p_1' \cdot p_2' \cdot p_3' \cdots \cdot p_r') = N(\prod_{i=1}^{r} p_i')$$

APPENDIX 3

Written instructions for the subjects

The purpose of this study is to investigate how the inspection of a quality control system can be improved. Method of inspection is one of the factors considered to improve outgoing quality. It is believed that the proposed investigation will be of significant importance in finding solutions to the problems of very costly inspection in cases where very high quality is required. Your help and co-operation is greatly appreciated.

The actual inspection task will be explained and demonstrated to you at the time of performing the experiment. You will have 4 different conditions and will examine 160 slides in each condition. After each condition you will have a small break of about 5 minutes, during which you can relax.

Your performance will be based on the number of defectives you correctly detect and on the number of good products you accept. Thus your aim will be to maximize correct acceptance and correct rejections and to minimize incorrect acceptances and incorrect rejections. Do not hesitate to ask if you have any questions.

This experiment will compare the efficiency of inspection tasks under different methods. In this experiment you will inspect the slides for two types of simulated defects.

Defect # 1: Occurrence of the digit '6' will be denoted as defect # 1. Whenever you see the digit '6' on any slide, stop inspecting that slide and mark the form as defect # 1.

Defect # 2: Any digit occurring consecutively will be denoted as defect # 2 e.g. '00', '11' etc. Whenever you see any two identical digit occurring consecutively on any slide, stop inspecting that slide and mark the form as defect # 2.

You will be inspecting the slides under four different conditions.

CONDITION 1

Here you will inspect 80 slides, Each slide will be projected for 10 seconds, you will have to inspect for both defect 1 and defect 2. When you detect any defect in a slide, stop inspecting that slide any further and mark an 'X' under appropriate column; then proceed to the next slide.

CONDITION 2

Here you will work in two stages. In the first stage you will inspect all the 80 slides for defect 1, each slide will be projected for 5 seconds. In the second stage, you will again inspect the same 80 slides but this time inspect for defect 2. Each slide will be projected for 5 seconds.

CONDITION 3

Here only one slide projector will be used. In the first stage, 80 slides will be run, subject # 1 will inspect for defect 1, while at the same time subject # 2 will inspect for defect 2. In second stage, the second set of 80 slides will be projected. This time the task is reversed, subject # 1 will inspect for defect 2 and subject # 2 for defect 1. Each slide will be projected for 5 seconds.

CONDITION 4

Here you will inspect all the 160 slides. In stage one, each subject takes a set of slides and inspects for both defect 1 and defect 2. In stage two, you will interchange the slides and inspect again for defect 1 and defect 2. Each slide will be projected for 5 seconds. Your performance will be combined and analyzed as a team effort.

Thank you for your help.

A P P E N D I X 4

List of defective slides (selected randomly)

Defect # 1	Defect # 2
26	11
36	18
41	29
51	44
56	48
65	60
68	62
72	69
93	85
98	87
114	109
132	112
147	117
149	126
152	141
159	154

APPENDIX 5

Score Sheet

Name: Date: Time:

Condition No: Stage: Subject:

Color of Rack: Speed: Major:

Task: Inspect for defect 1 and 2

defect 1 = Occurrence of digit '6'

defect 2 = Occurrence of two identical consecutive digits, e.g. '00'

	D1	D2		D1	D2		D1	D2		D1	D2
1			21			41			61		
2			22			42			62		
3			23			43			63		
4			24			44			64		
5			25			45			65		
6			26			46			66		
7			27			47			67		
8			28			48			68		
9			29			49			69		
10			30			50			70		
11			31			51			71		
12			32			52			72		
13			33			53			73		
14			34			54			74		
15			35			55			75		
16			36			56			76		
17			37			57			77		
18			38			58			78		
19			39			59			79		
20			40			60			80		

APPENDIX 6

List of numbers used

4839253292	5190594871	9491591327	1478707950
7917072525	0874150243	1057059149	1405205937
1312829043	1892583153	1318309472	7451714292
1320747318	5370154291	5801975289	2047974528
3247374018	7393435527	4854901971	2381248714
7340873080	2194591828	5234715953	8459575283
5908530807	2020370291	1249478082	2629795015
5130530807	1741259138	5191326127	1734282428
9853193804	9095075101	3090727450	7927921815
1742415943	9749835102	7098482834	8702827493
0948523859	5837050341	9725180985	0243720949
3848271528	0795871281	7810812302	2932378249
3385431913	2790431927	2537821504	5414728279
9092818372	9498294131	8483193580	1403570128
3241747939	2952384107	4713721081	5191454293
9480413242	4898454521	5419248504	5372351457
4015485730	7435242140	2094984514	3570575878
3089537297	3581897297	9180703431	5739840548

7050817019	1032198732	2707530513	0040914834
3108452420	2702303975	9098424255	7598791924
9583821808	4548142370	9425471057	7089517404
2925490827	5210517391	9243721298	3718354030
0471519784	0202423015	5842710347	7948095141
4542081403	1785975318	0817292393	0158298728
8749397093	8323781540	6124514930	2475053584
4235853802	5244287849	2403851780	1231358357
2891025245	1535293812	8487387109	4741352343
7107418405	5787043272	9391503218	5838535781
8087304714	5781454349	3915975814	7920314752
7139589158	9501405834	5328275758	9075453424
0713081814	4415281823	9185347510	9170834740
2457413728	8217072594	4290582853	5047197921
8358707849	3834272737	9843215840	7950178154
3585409259	2813194948	6234184121	8537083710
4873210409	0239079737	3042902378	7937243981
5414527459	9837089727	3027407497	9323523073
3141379295	7872782175	5757351492	5378717132
8301438949	0749345209	2349373854	0192741321
2385848528	3709787810	7125738919	0828101298

8170418571	0878379538	4743970248	1815875058
1805170478	0549037819	5789315234	2571908502
5789082717	4712302818	3725785258	0251240854
2895253920	8797982534	2747537951	4039152987
1459808158	9292481430	1878391543	1971092485
8439844327	9202480430	2140189814	2591348071
7851547732	8482917909	2353081590	1843973250
8189181205	1490127127	2757028457	5439830914
2327805275	2350378070	0209143597	8915472804
5947314312	0412871315	0739601249	7039718180
8703751309	3803769058	1324740304	9475175878
7175234798	4704903142	3840282814	4739393252
9234713535	0419842578	4585825310	8258501042
4353545345	4875415187	9407358271	7807010120
1941810725	4979152712	3193405429	7127905914
1937297540	1978594399	2487953024	4625072075
5059092482	9496571329	9755291384	0748479439
4803037109	7409374271	3189413831	0152908781
7189718478	9529139547	9389407542	8180715094
5091042183	1478383825	4317193485	5734153483
2347390421	2713810389	9128721075	8540913232

0127303505	1518930752	5187535398	5712435138
8547371505	5973934521	1784975905	9754149074
7352714381	0805430291	2735729308	1723215210
5728480528	7978389495	3740371878	2072115158
4982459197	9140412930	8329897087	7532125485
7202045234	3813841819	2083898283	2849210951
3107243149	3745715385	7230859218	9785325229
5981080232	0198524215	0738295142	0521531341
8090787347	9403587298	2398518093	3895874790
5723040510	8791385837	0898273103	3715980838
0147371425	5422158512	0172031209	4945943932
7038204293	4193193929	8795317920	2507382125
5019205404	7812473909	5407585294	0530407359
0757393803	7383089312	8454982785	8510182930
8931418232	7514124904	3483785825	5705795803
7428087575	1537531043	5013909548	8027472352
8074272547	8507098305	7029414092	4516897098
7217985838	6238504392	0432518205	8350948583
3510795798	8780375902	9232040451	9730291758
9519410457	5472192538	2031039583	7431752849
0789301879	4287325208	7038943813	0838532437

1385848413	1982430128	8924548347	3032745014
7583193278	1017247542	2501972848	7145072848
5020543494	1057385814	7545419072	5392714728
9170827320	0080785708	9012751870	9038328494
0531434318	5935989829	7910971823	4051710185
5725317828	2312040535	9871974874	1489241707
4741821832	1503438790	2937894749	9172723254
3480540939	4512358431	1354817205	4313904179
7431570742	8042179521	8104873153	0298181293
5715693987	9834040202	9140507545	9683851739
8274175939	9013190494	4396323098	1454894738
5092403471	1918489585	8091842357	1937849439
0895749737	2435078070	1737253432	0140378194
8235132153	3201832109	3045791592	4049058010
8379593459	7019349875	8075193015	8981843803
1871848542	0174585022	4751495237	7894348419
7979374724	4373208279	8343574254	7842841714
3519030845	7204794781	9573072532	2581085425
9123133425	7290745851	1217478028	7015742754
5208320313	1832732350	4730383425	9750172305
5297527978	2742351984	9580478284	7475751928

4807585049	4848318328	3923785185	3090351348
1372158409	4709414329	1820483879	8187521954
4257819735	7345139292	1391858275	5880780472
0270254341	8318972729	5348240151	2548308149
9450302824	7585474972	4745321303	5147402412
3184032140	2037240125	3798973209	3259798408
4142829523	2321834851	8903070320	3053914514
2708388204	2535715401	0547953878	2928371458
4587032473	4132859243	1495262739	1050470102
0309780109	5374753969	0905173839	8749703080
3589437145	8728529323	4843410421	4219217485
1805749150	5321949218	1374949825	1741405739
9495804291	9404059305	4914271319	9298352505
2984789237	9254978173	9397501085	5857932574
4703242432	9229293489	7195071412	4784381820
7043068531	8204041421	4101854042	3125010898
1828548405	8458712739	7473039835	2104794829
3205731424	2074810294	3525703481	8754391989

ACKNOWLEDGEMENTS

I would like to take this opportunity to extend my deep appreciation to my advisor, Dr. Stephan Konz, for his expert guidance and advice. But more importantly, I am indebted for his constant encouragement, prodding and support without which this work would have been abandoned long ago.

I wish to offer my sincere thanks to my committee members, Dr. Corwin Bennett and Dr. Frederick Rohles, for their cooperation, constant help and encouragement. Thanks are also due to Mrs. Deanna Nicodemus for her prompt typing. Last but not the least, I am grateful to my parents, wife, brother and sisters for their everlasting love and encouragement.

TEAM EFFICIENCIES ON A
PACED VISUAL INSPECTION TASK

by

KHALID OSMAN

B. Tech., Indian Institute of Technology, 1972
Bombay, India

AN ABSTRACT OF A MASTER'S THESIS

submitted in partial fulfillment of the
requirement for the degree

MASTER OF SCIENCE

Department of Industrial Engineering

KANSAS STATE UNIVERSITY
Manhattan, Kansas

1976

A B S T R A C T

Inspection efficiencies for two methods of inspection team organization (division of labor and redundancy) were considered. The inspection task consisted of detecting two types of visual defects (the digit 6 in a series of digits and two consecutive identical digits) under four different conditions. To simulate machine-pacing, 160 slides of numbers were projected on two slide projectors; 24 female students were the subjects. Inspector accuracy performances were analyzed by conventional measures as well as signal detection theory analysis. Inspection accuracies were .95, .94, .93 and .93 for the two conditions in which subjects inspected for only one defect. When they inspected for two different defects simultaneously, the accuracies was .91, .85, .84 and .78.