MULTI-DEFECT INSPECTION

by

JINN-YEN SU

B.S. (Industrial), Tunghai University
Taiwan, Republic of China, 1976

A MASTER'S THESIS

submitted in partial fulfillment of the
requirements for the degree

MASTER OF SCIENCE

Department of Industrial Engineering

KANSAS STATE UNIVERSITY
Manhattan, Kansas

1981

Approved by:

Stephen Kay
Major Professor

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

I wish to express my sincere thanks to my advisor, Dr. Stephan Konz, for his continuing guidance and advice.

Thanks are due to Dr. Corwin Bennett and Dr. Frederick Rohles for their help in designing the experiment.

Special thanks are also due to my family and friends for their everlasting love and encouragement.

LIST OF TABLES

Page

LIST OF FIGURES

Page

INTRODUCTION

A common task, at least the one most people seem to think of with respect to inspection operations, is the scanning task. In this type of inspection the inspector searches for defects by making a point by point examination. Figure 1 (Williams, 1973) illustrates the general search process. Scanning tasks are included in a wide variety of inspection jobs. Included are the examination of electronic hardware for improper connections, wiring, component placement, and contamination; the inspection of fruit, vegetables, peanuts, and so on for the purpose of culling out items or classifying by grades; the inspection of automobile body surfaces for blemishes in paint; and the examination, by lot, of machine-produced items for purposes of identifying and enumerating defective items. In general, tasks of this type are difficult to perform even under the best of conditions and the lowest level of task complexity. As a consequence, some attention has been given to improving inspection performance by overcoming the inspector's human limitations and enhancing his human capabilities. Colquhoun (1957) and Megaw (1979) listed a number of variables affecting visual inspection performance. They are summarized in Table 1. We are going to introduce only four task factors closely related to this study (pacing, defect discriminability, inspection method, and defect rate).

Pacing

Williges and Streeter (1971) investigated the effect of

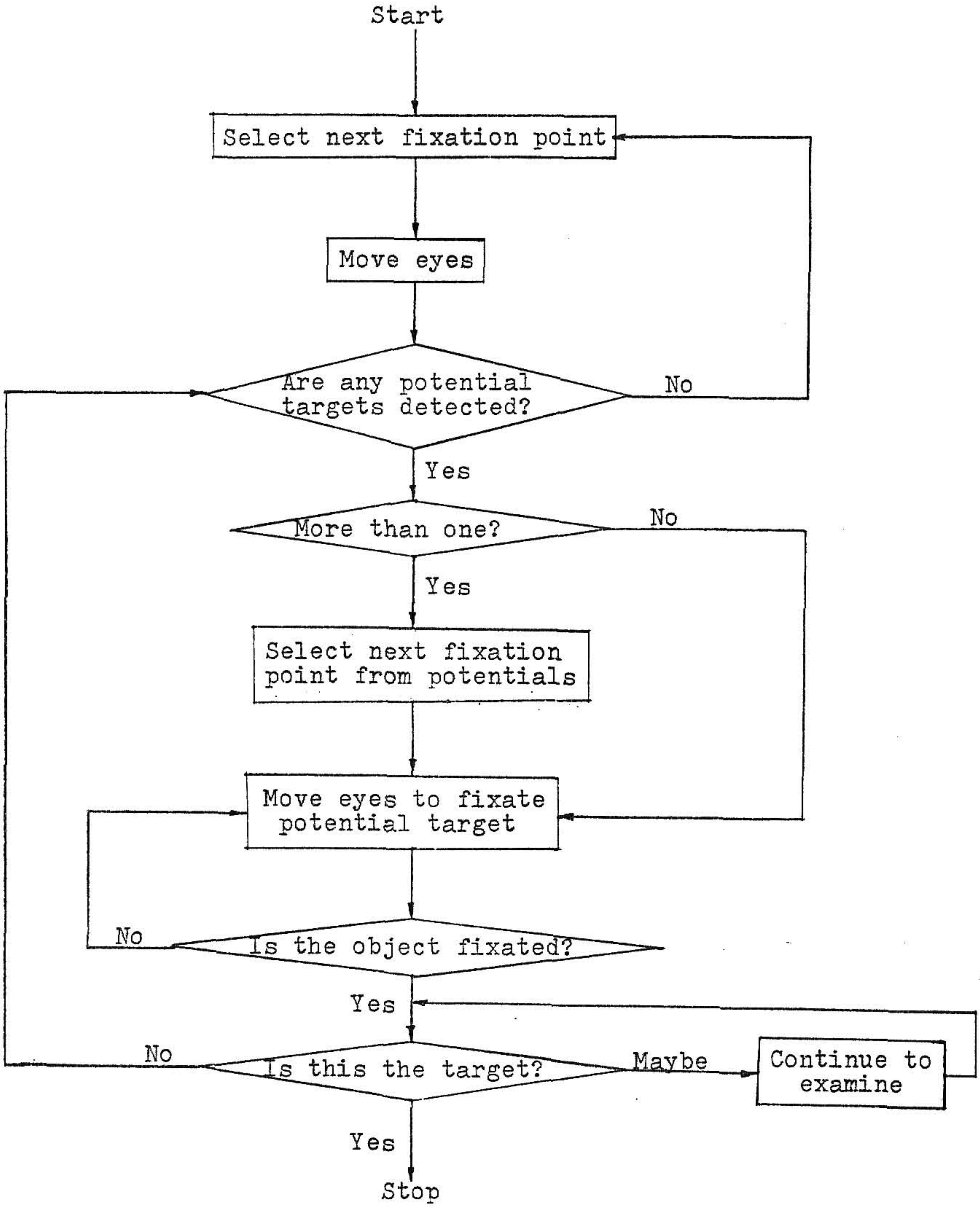

Figure 1. The general search process (from Williams, 1973).

TABLE 1

Factors Affecting Visual Inspection Performance (Adapted from Colquhoun, 1957 and Megaw, 1979)

Individual Factors	Environmental Factors
a. Visual acuity	a. Illumination
b. Color vision	b. Inspection aids
c. Scanning patterns	c. Background music
d. Age	d. Background noise
e. Experience	e. Work station design
f. Personality	**Organizational Factors**
g. Sex	a. Instructions
h. Intelligence	b. Feedback
Task Factors	c. Feedforward
a. Inspection time	d. Training
b. Pacing	e. Selection
c. Direction of movement	f. Standards
d. Viewing area	g. Time-on-task
e. Density of items	h. Rest pauses
f. Distribution of items	i. Shift
g. Defect rate	j. Sleep deprivation
h. Defect mix	k. Social factors
i. Defect discriminability	l. Motivation
j. Item complexity	m. Grouping of inspectors
k. Inspection method	n. Job rotation

pacing on visual inspection. The subjects were asked to detect pin-hole defects in 1/4 inch diameter transparent yellow plastic discs under machine-paced and self-paced procedures. They found that the self-paced inspection was performed at a much lower rate but yielded more defect detections and fewer false alarm errors.

McFarling and Heimstra (1975) had subjects inspect simulated printed circuits varying in circuit complexity under machine-paced and self-paced conditions. At the end of the experiment, the subject completed a task descriptor rating list where each descriptor was rated on a four-point scale from "very adequate descriptor" to "very inaccurate descriptor". For all levels of circuit complexity, self-paced inspection took longer but resulted in a higher level of defect detection. The false alarm rate was maintained at a high level for both pacing conditions. In addition, the task descriptor rating indicated that both self-paced and machine-paced subjects perceived the inspection task as being relatively unpleasant, but the machine-paced subjects tended to have more negative responses to the task.

Although machine-paced procedures are used widely in inspection tasks today, the self-paced procedure also is used. From the above studies, we concluded that a self-paced procedure is needed when the inspection is complex and the accuracy is emphasized.

Defect Discriminability

When searching for targets, inspectors tend to fixate on objects having the defect specifications. In many inspection situations, defects can be difficult to detect for one of the following reasons: (1) The defect does not emerge perceptually from its immediate background because the contrast difference between the defect and its background is very low. (2) The defect does not emerge perceptually from its immediate background because the patterning of the background and the defect combine to obscure the defect. (3) The defect is clearly distinguishable from its immediate background, but it remains hard to detect because it is very small relative to the total area that has to be inspected. (4) The defect is clearly distinguishable from its immediate background, but the presence of many non-defects in close proximity to the defect confuses the inspectors with excessive information so that he is unable to isolate the defects from within the search area.

Howarth and Bloomfield (1969) derived a relationship between search time and defect discriminability:

$$\bar{t} \propto 1/D^2 \qquad (1)$$

where $\bar{t}$ is mean search time and D is a measure of discriminability. Then they verified the relationship by conducting an experiment where targets and nontargets differing in size both were presented within the display area. Figure 2 shows the results. The discriminability measure used here was the difference in diameter of the target and background discs.

The relation was verified by a number of studies (Miller and Ludvigh, 1960; Bloomfield, 1972; Bloomfield, Marmurek, and Traub, 1974) where the measure of discriminability was defined

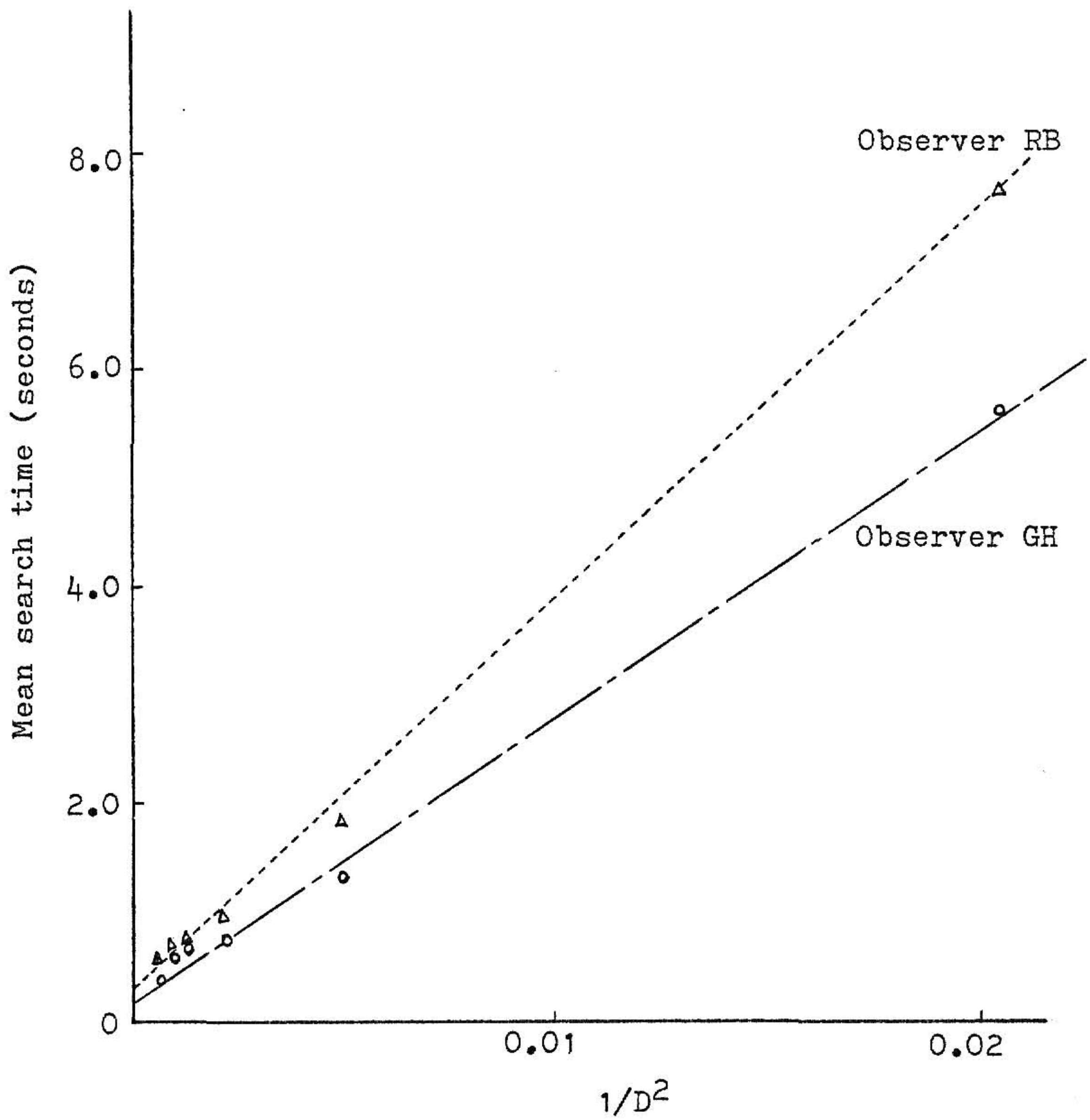

Figure 2. Relationship between mean search time and discriminability for two observers searching for disc targets among larger nontarget discs (from Howarth and Bloomfield, 1969).

differently each time. However, the relation did not hold in the study by Drury (1975) where t was linear to the reciprocal of D^3 instead of D^2.

From the studies cited here, we concluded that a quantitative description of the defect discriminability can be used to predict the search time. However, additional experimentation to determine the validity of this relationship under a wide range of display conditions is needed before it can be applied to an operational situation.

Inspection Method

Inspectors have certain inherent limitations and capabilities with respect to inspection operations. Harris and Chaney (1969) reported that even under relatively ideal working conditions, the average percentage of defects detected ranged from 30% for the most complex items to 80% for the least complex items. Harris (1966) investigated the effect of the complexity of electronic equipment on inspection performance. The inspectors were given as long as they wanted to complete their inspections. Figure 3 shows that equipment complexity had a significant detrimental effect on inspection performance. This effect could not be overcome by extending the amount of inspection time allotted.

When inspecting complex products, the inspector frequently is required to scan the product for relatively large number of different types of defects. He must be aware of more standards upon which to base his judgments of conformance. Also, confusion

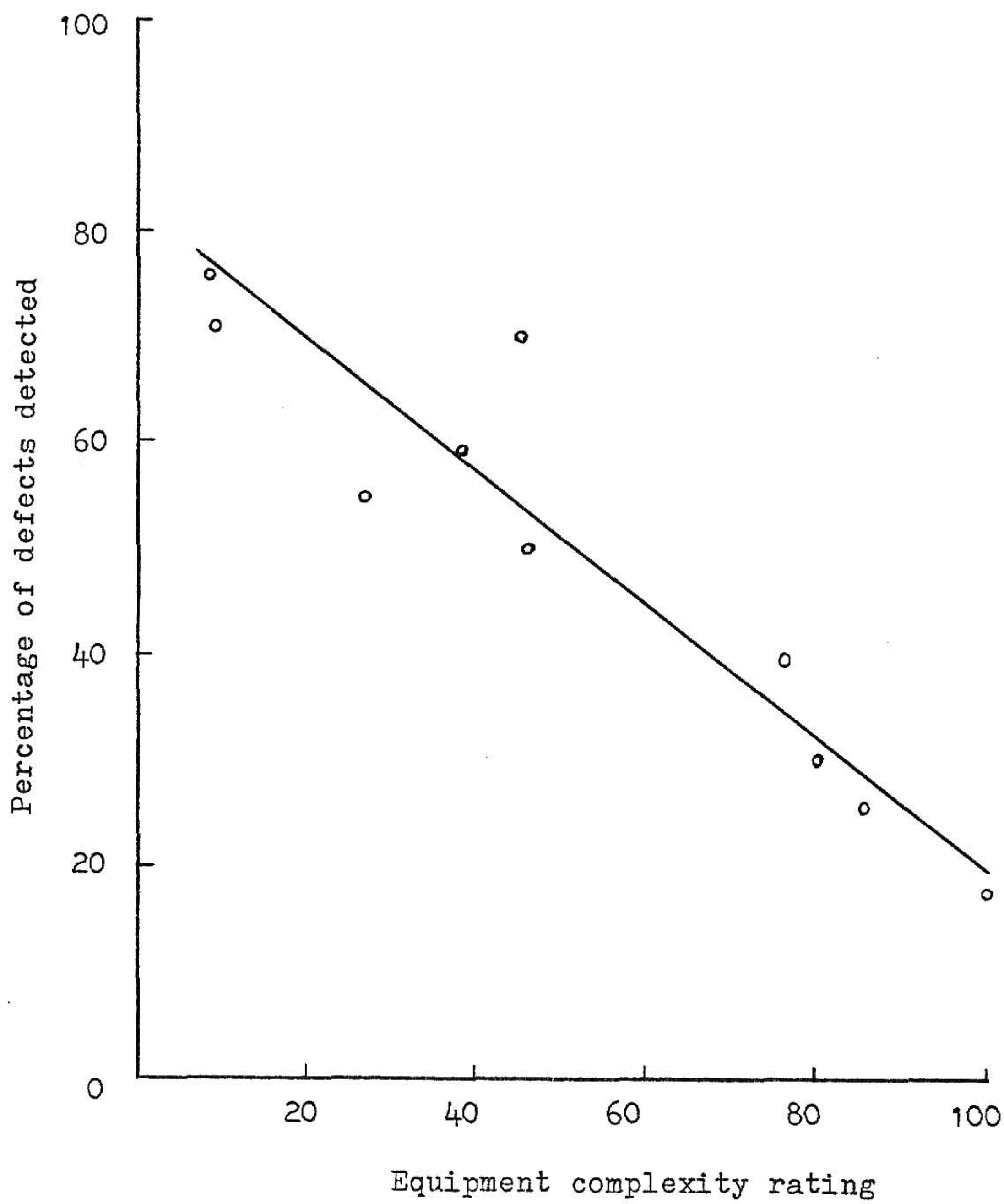

Figure 3. Inspection performance as a function of equipment complexity (from Harris, 1966).

may be more likely to occur when he is searching a more complex item for defects. How should the inspector approach this task? Should he inspect for all possible defects in one item before going on to the next (simultaneous scanning method) or should he inspect all items for one type of defect and then the next type of defect (sequential scanning method)? Figure 4 compares the flow diagrams of these two approaches. Figure 5 shows a symbolic comparison of these two methods. The simultaneous scanning method reduces the length of eye movement and material handling while the sequential scanning method has less mental "gear shifting" and less possible omission of using a criterion.

Harris and Chaney (1969) had experienced inspectors inspect two complex electronic chassis of the type regularly inspected by the inspectors under these two inspection methods. They found that the sequential scanning method was 90% more effective than the simultaneous scanning method in detecting the critical defects. In addition, when the simultaneous scanning method was used, certain types of defects were completely overlooked and some inspectors reported much higher percentages of certain defect categories than they did of others. In other words, being unable to attend to all of the quality characteristics and their associated standards, the inspector's attention was reduced to relatively few types of characteristics.

Neisser, Novick, and Lazar (1963) had subjects search for letter targets through lists of printed symbols. Each item was represented by a string of six letters or numbers (for example, NBJ3SG). Each list consisted of 50 items. Among them, one and

Simultaneous scanning method — Sequential scanning method

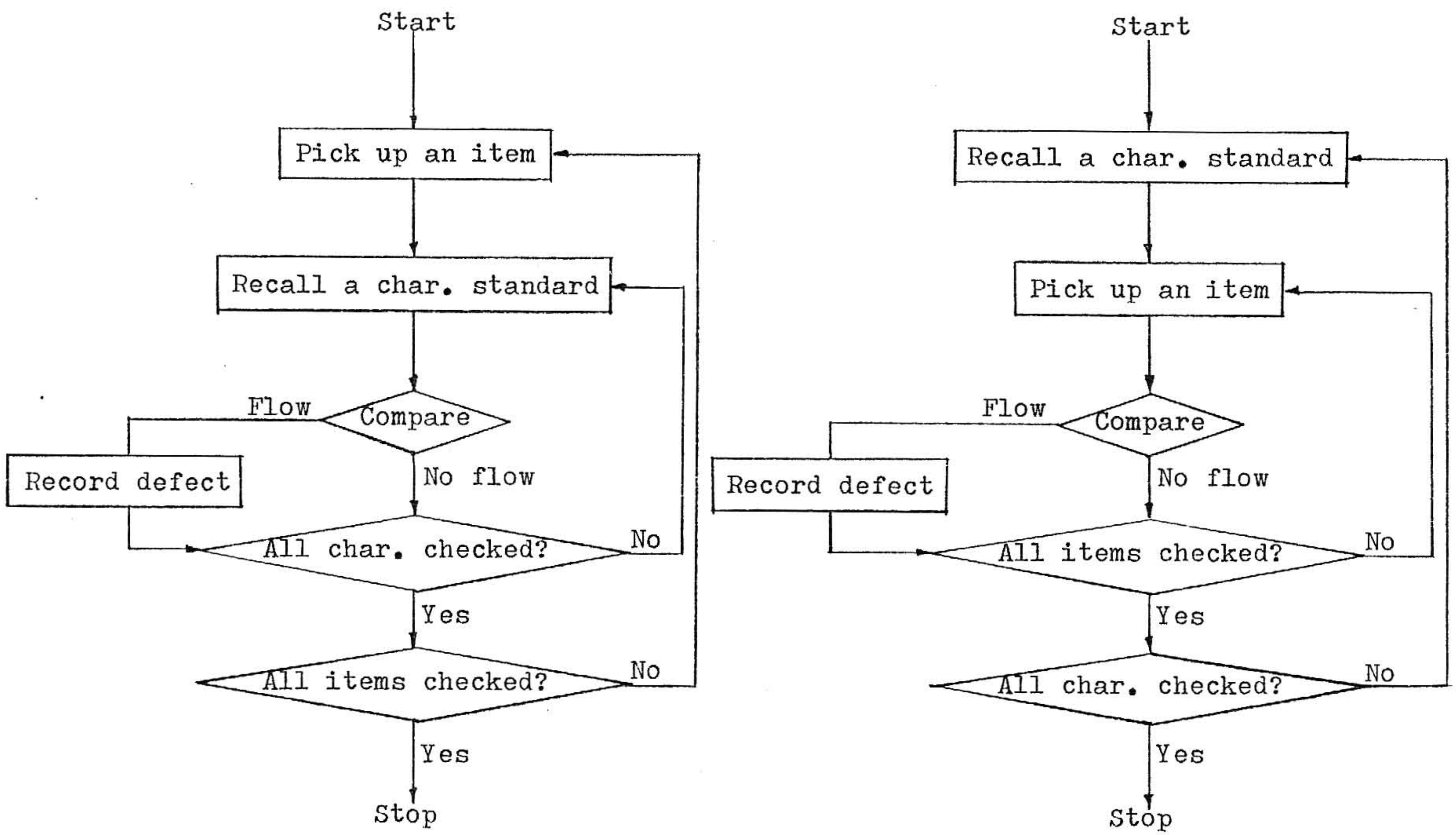

Figure 4. Flow diagrams of simultaneous and sequential scanning methods.

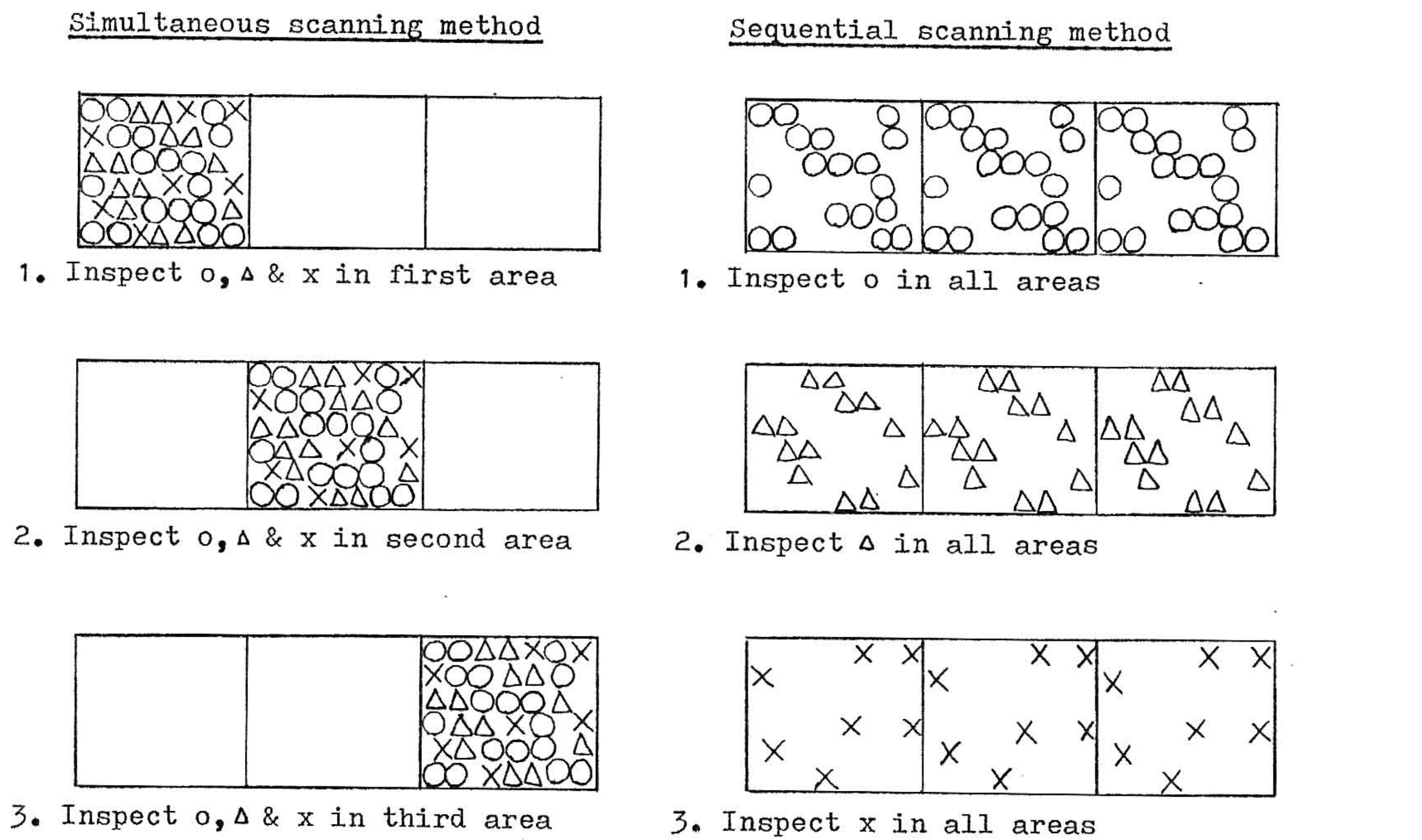

Figure 5. Symbolic comparison of simultaneous scanning method and sequential scanning method (from Harris and Chaney, 1969).

only one item contained a target. Once finding the target, subjects recorded it and started a new list. In Condition "K", the target was the letter K. In Condition "AFKU9", the target was equally likely to be A, F, K, U, or 9. In Condition "HMPZ4", the target was equally likely to be H, M, P, Z, or 4. In Condition "Ten", the target was equally likely to be A, F, K, U, 9, H, M, P, Z, or 4. Only speed was emphasized. The probability of target occurring and the length of eye movement remained the same in the four conditions while the total number of targets inspected was varied. The less-target-searching conditions seemed to be more beneficial than the more-target-searching conditions. But, despite this, after practice, it took no longer to search for five or ten targets simultaneously than just one target. The percentages of missed signals were 30.2, 26.0, 21.4, and 25.8 for Conditions K, AFKU9, HMPZ4, and Ten respectively. There were no significant differences of error rates among the four conditions except that Condition K (30.2) had more errors than Condition HMPZ4 (21.4).

Kristofferson (1972) replicated the Neisser et al. (1963) study and allowed for measurement of both missed signals and false alarms. He had suggested that the Neisser et al. results were confounded by low error rates and that indeed multiple target search did take longer than single target search. However, there were no significant differences of error rates (3.0, 3.8, 5.6, and 3.6 percent for Conditions K, AFKU9, HMPZ4, and Ten respectively) among the four conditions.

Konz and Osman (1977) had subjects search, in a paced situation, for defects by the simultaneous scanning method and the sequential scanning method. Each 3x10 digit matrix represented an item. The occurrence of the digit "6" was called defect A while the occurrence of two identical consecutive digits (for example, 77, 22, 33) was called defect B. There was no overlap; that is, an item might contain either a type A or B defect but not both. Total inspection time was constant. Inspecting for one defect at a time was found significantly better for defect detection than inspecting for two defects simultaneously.

Ebeling and Orr (1978) compared the two inspection methods in potato grading operations. There were four grades of potatoes: premium, number 1, number 2 and cull. The premium potatoes were defined as good items while the other three were defined as three types of defective items. The inspectors inspected three quality levels of potatoes on a conveyor. Their results suggested that the best inspection method depends on the percent of defectives. Simultaneous scanning is best at lower defective percents while sequential scanning is best at higher defective percents (Figure 6).

From the previous studies cited, we concluded that a number of studies investigated the influence of defect discriminability on inspection time or the efficiency of the inspection method. But virtually no study investigated the possible effect on the optimal inspection method for differing discriminability of defects. In addition, a third inspection method (inspecting for one group of similar defects at a time)

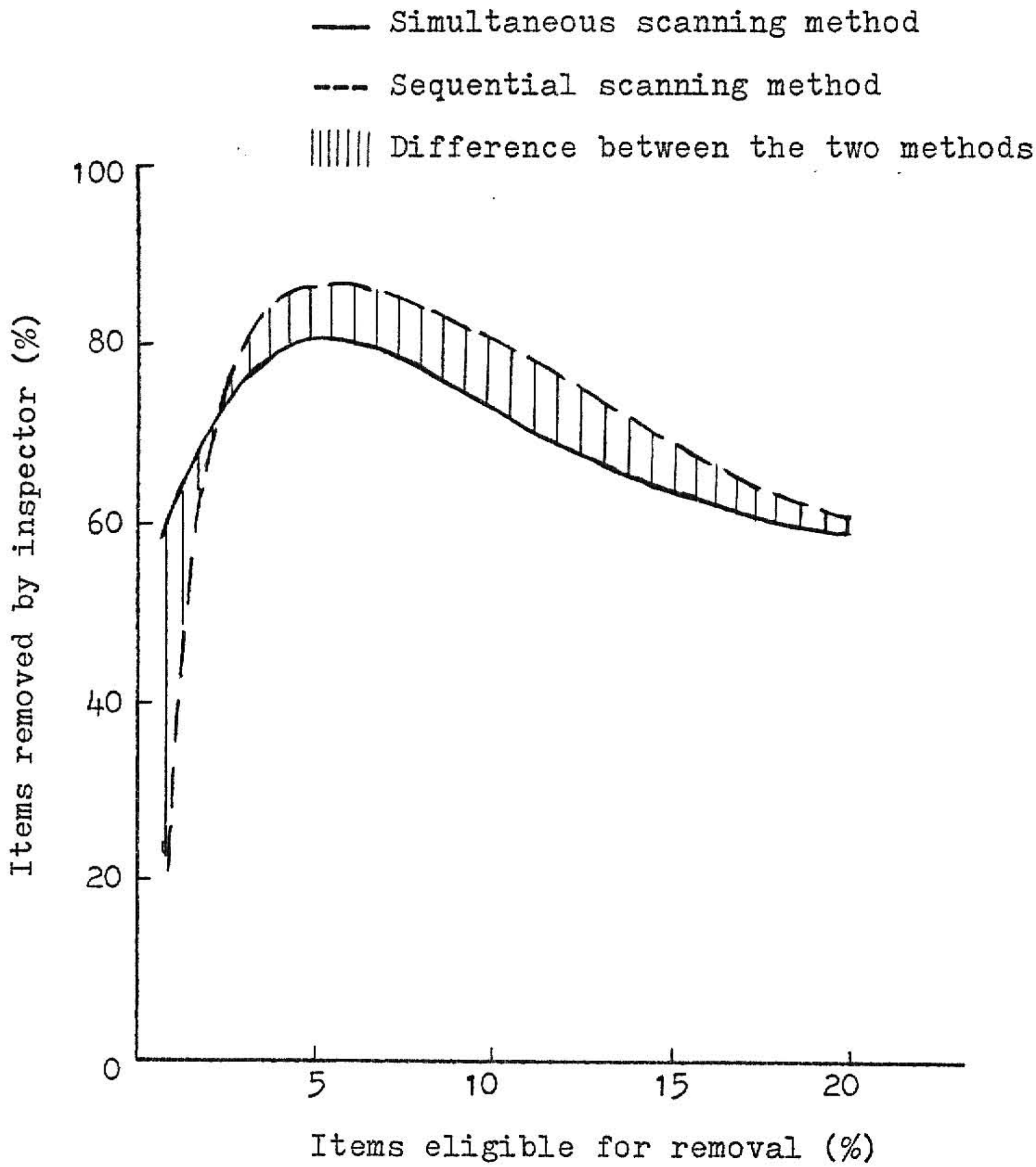

Figure 6. Grading accuracy of simultaneous and sequential scanning methods (from Ebeling and Orr, 1978).

was evaluated in this study.

Defect Rate

Jenkins (1958) had subjects monitor the periodic movements of a pointer and detect the occurrence, at random time-intervals, of a movement of greater than the standard amplitude that constituted the signal. The signals were programmed at average rates of 480, 60, 30, or 7.5 signals per hour. The percentage of signals detected declined in the course of the watch. The extent of the decline depended upon the signal-rate, with lower rates producing greater declines. The percentage of false alarms increased sharply as the rate decreased.

Harris (1969) had subjects inspect for defects among symbols on printed sheets. Inspection accuracy decreased with reductions in defect rate (range from 16% to 0.25%). See Figure 7. The percentage of missed signals increased slightly between defect rates of 16% and 1%, but dropped sharply between 1% and 0.25%. The percentage of false alarms increased with reductions in defect rate and at an accelerated rate as the defect rate approached zero.

Fox and Haslegrave (1969) tested the effect of defect rate (range from 1% to 5%) for machine-paced and self-paced inspection. Subjects inspected for defective screws on a moving belt and on an inspection table. The defect rate effect was found in the self-paced situation. Raising the defect rate increased the level of detection of defects significantly. But the effect was not shown in the machine-paced situation.

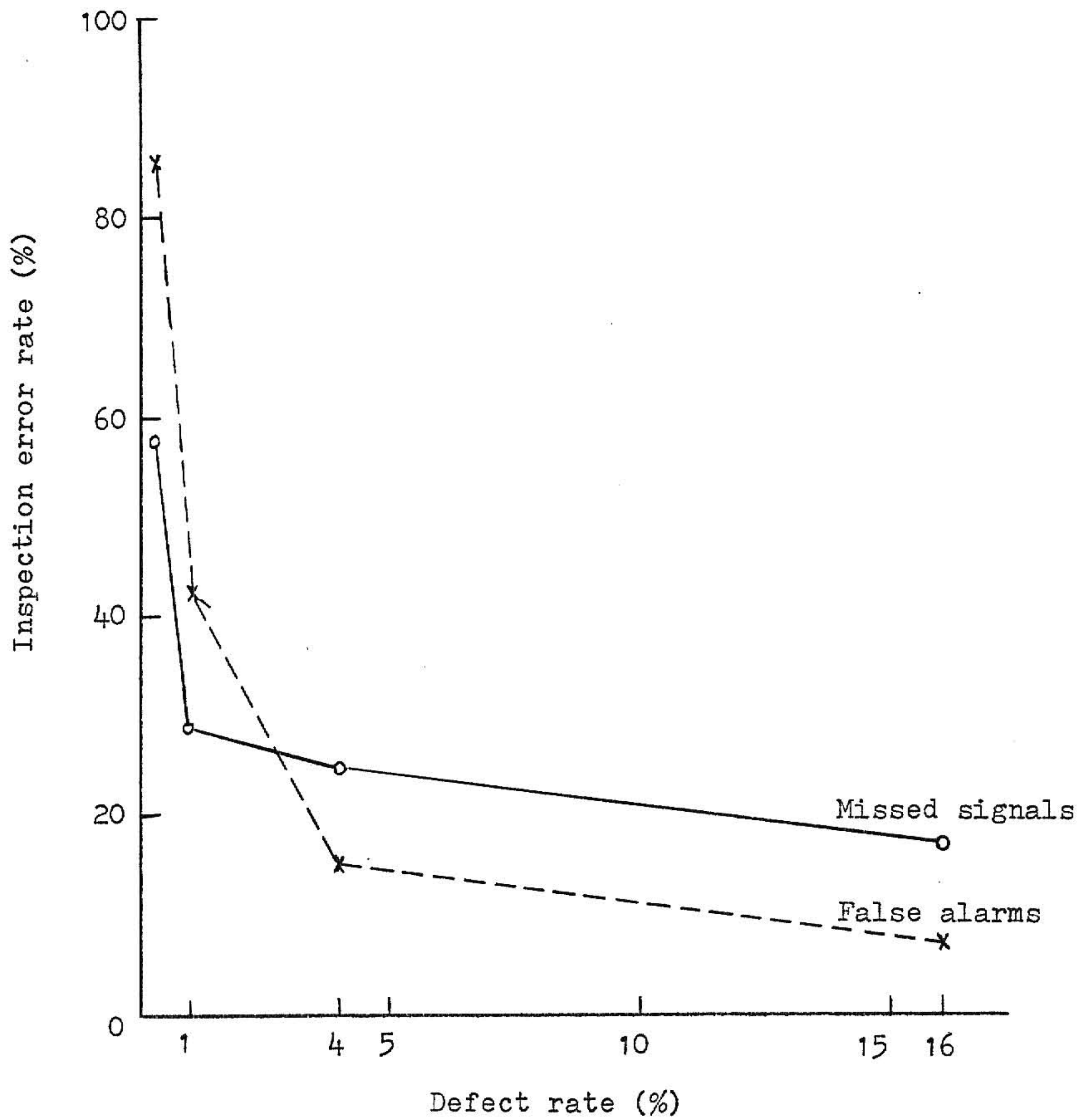

Figure 7. The relationship between defect rate and inspection accuracy (adapted from Harris, 1968).

In multiple defect searching, the defect rate varies with the number of defects inspected at a time. When the inspector inspects for one defect at a time, the defect rate is the percentage of that type of defects in the lot. When the inspector inspects for five defects at a time, the defect rate is the percentage of the five types of defects in the lot. The above studies showed that defect rate had a significant effect on inspection accuracy. Thus defect rate is a factor which needs to be considered when comparing the methods for multi-defect inspection.

PROBLEM

This study compared three inspection methods under three difficulty levels of defects.

The three inspection methods were: (1) Inspect for one defect at a time, (2) Inspect for all defects at a time, (3) Inspect for one group of similar defects at a time.

The three difficulty levels were:

1) Mixed level: Five types of easy and difficult defects needed to be inspected. Compare the three methods.
2) Easy level: Three types of easy defects needed to be inspected. Compare Method 1 and Method 2.
3) Difficult level: Two types of difficult defects needed to be inspected. Compare Method 1 and Method 2.

METHOD

Task and Enviroment

This experiment simulated a self-paced inspection situation where inspection accuracy was emphasized although the task was dull and difficult.

The inspected items in the present experiment were developed from the designs of Neisser et al. (1963) and Konz and Osman (1977). Each item was represented by a 5x5 matrix. First, a computer program randomly generated 360 5x5 matrices where the element of the matrix was equally likely to be 2, 3, 4, 5, 6, 7, 8, 9, F, or R. They had been checked for not having any defect patterns (such as 345 or 678 in a row). Next, the computer program randomly selected 115 matrices to contain the defects. The defective matrices and their defect types are listed in Table 2. Then the original element or elements on the defect position (chosen at random) were replaced by the appropriate defect pattern.

There were five types of defects (A, B, C, D, and E) varying in discriminability. See Figure 8. Occurrence of the letter "A" in a matrix was designated as defect A. There were 25 (6.9%) matrices with defect A. Occurrence of the letter "B" in a matrix was designated as defect B. There were 25 (6.9%) matrices with defect B. Occurrence of the symbol "$" in a matrix was designated as defect $ (or C). There were 25 (6.9%) matrices with defect C. Occurrence of the number "345" in a row of a matrix was designated as defect 345 (or D). There

TABLE 2

List of Defective Matrices

M	D	M	D	M	D	M	D
7	A	182	B	305	C	57	E
37	A	183	B	313	C	68	E
47	A	192	B	314	C	72	E
51	A	211	B	340	C	95	E
66	A	237	B	344	C	101	E
85	A	190	B	6	D	126	E
123	A	322	B	9	D	144	E
127	A	325	B	45	D	160	E
147	A	328	B	60	D	171	E
157	A	333	B	67	D	206	E
186	A	343	B	99	D	208	E
233	A	346	B	129	D	221	E
249	A	349	B	136	D	239	E
250	A	23	C	180	D	256	E
251	A	34	C	181	D	274	E
255	A	35	C	188	D	294	E
260	A	63	C	204	D	302	E
269	A	70	C	231	D	360	E
297	A	91	C	254	D	345	A,B
321	A	97	C	273	D	10	A,C
331	A	111	C	276	D	32	A,D
2	B	114	C	279	D	86	A,E
11	B	151	C	286	D	106	B,C
14	B	161	C	335	D	38	B,D
58	B	167	C	341	D	354	B,E
80	B	200	C	352	D	158	C,D
104	B	224	C	15	E	154	C,E
119	B	230	C	46	E	59	B,E
121	B	288	C	55	E		

M = Matrix number; D = Defect type.

Defect A: Occurrence of the letter "A" in a matrix.

```
F 7 6 9 5
4 3 A 8 2
6 3 8 6 7
2 F R 5 4
4 8 6 7 3
```

Defect B: Occurrence of the letter "B" in a matrix.

```
8 5 3 4 B
3 7 R 6 9
7 5 6 4 9
F F R 6 3
R 3 5 2 7
```

Defect C: Occurrence of the symbol "$" in a matrix.

```
8 F 8 3 R
4 8 7 5 9
2 8 5 9 F
7 4 7 3 3
3 R F 2 $
```

Defect D: Occurrence of the number "345" in a row of a matrix.

```
9 R 3 5 F
3 2 7 5 6
4 3 4 5 4
F R 3 8 5
7 3 5 4 F
```

Defect E: Occurrence of the number "678" in a row of a matrix.

```
R 3 8 5 6
8 6 7 6 5
4 F R 4 7
6 7 8 7 6
5 F 3 6 R
```

Figure 8. An example of the five types of defects.

were 25 (6.9%) matrices with defect D. Occurrence of the number "678" in a row of a matrix was designated as defect 678 (or E). There were 25 (6.9%) matrices with defect E. A matrix could contain more than one type of defect.

The matrices were typed with a type size of 10 letters to an inch. Twelve matrices, four in a row, were typed on a letter size (8.5"x11") sheet of paper. See Figure 9. The distance between two adjacent matrices was about 1.5 inches. The 360 matrices were typed on 30 original sheets. Then they were duplicated by a copy machine to make test sheets. Black ink and white paper were used. Upon finding a defect, the subject was asked to underline it with a blue pen.

Each subject performed under three conditions twice in a balanced sequence. The 30 test sheets were provided in random orders for the subject in each condition. See Table 3.

Condition 1: Inspect for one defect at a time. The subject inspected all matrices for one type of defect, then for the next type of defect, and so on. For example, the subject inspected all matrices for defect B, next for defect E, next for defect A, next for defect C, and then for defect D. The order of defects was arranged randomly (Table 4).

Condition 2: Inspect for all five defects at a time. The subject inspected each matrix for all five defects simultaneously. That is, the subject inspected one matrix for all possible defects A, B, C, D, and E simultaneously before going on to the next matrix.

Condition 3: Inspect for one group of similar defects at

R 2 8 2 4
3 R 7 9 9
R 5 5 F R
3 R 2 4 9
9 8 2 4 7

3 F 9 5 5
2 B 6 5 6
5 F 9 4 8
6 4 5 7 4
9 R 9 9 5

6 8 5 5 8
9 F 3 R 3
5 R 6 2 5
9 9 2 7 3
F 5 R 3 5

9 7 2 6 F
8 6 7 2 R
5 6 5 6 5
8 5 4 7 F
9 4 9 2 R

9 F R 2 4
6 R 6 8 6
9 R 6 3 8
7 4 5 7 3
8 7 9 9 8

3 3 5 R 7
6 3 4 5 5
9 R 4 6 2
7 6 5 3 2
4 7 4 2 4

5 5 R R 3
8 8 A 8 9
8 7 3 4 4
4 4 F 9 6
7 F 6 6 5

7 2 5 4 8
F 7 9 3 7
3 4 3 7 2
R 2 2 F R
9 4 3 3 2

R R 6 8 R
F 2 R 3 3
3 4 4 F F
4 9 5 F 5
5 3 4 5 2

3 $ 5 3 R
R 4 3 5 5
8 2 8 4 5
6 F 2 8 6
A 5 7 5 8

2 2 4 R F
9 3 7 8 5
4 5 R B 3
4 4 7 4 2
2 5 8 F 4

R F 9 4 8
7 R 6 2 8
8 8 6 R 8
F F 3 R 9
7 6 R 5 2

Figure 9. An example of the test sheet.

TABLE 3

Test Sheet Orders

Subject	Condition	Test Sheet Order
1	2	21 23 30 11 8 1 15 29 13 14 24 9 7 2 28 16 19 17 10 27 20 12 26 4 18 22 25 3 5 6
	3	28 30 4 27 22 1 13 24 16 5 2 15 3 25 7 8 17 19 26 9 21 12 14 29 10 23 6 20 18 11
	1	2 24 11 7 17 6 14 1 30 4 28 26 13 25 15 21 3 16 23 20 10 29 27 5 19 8 22 12 18 9
	1	28 12 21 17 14 4 27 5 23 25 11 3 22 30 13 1 10 24 15 6 2 20 9 7 16 18 26 8 29 19
	3	22 25 5 14 17 12 19 10 28 26 9 30 24 7 4 8 11 13 21 1 20 29 18 16 6 2 27 23 3 15
	2	29 5 7 30 16 25 19 4 13 10 1 17 27 15 24 28 22 3 23 20 2 6 12 11 9 18 14 8 21 26
2	3	6 9 3 12 10 27 24 22 4 25 11 30 14 23 1 16 8 15 13 29 2 26 7 20 18 19 5 28 21 17
	1	23 9 12 20 17 8 18 16 2 3 7 10 29 11 30 15 22 13 6 21 26 27 14 4 24 25 1 19 28 5
	2	30 15 16 14 23 27 12 6 9 10 5 4 22 8 13 21 28 3 1 19 26 29 2 20 24 25 11 18 17 7
	2	7 25 1 6 5 13 27 21 29 11 30 15 14 19 24 28 2 26 3 4 22 23 12 18 9 10 17 8 16 20
	1	22 26 23 2 19 24 5 12 25 28 6 9 21 1 3 17 8 20 10 13 16 14 4 27 1 18 29 30 7 15
	3	2 7 6 26 1 18 24 14 19 22 10 20 16 25 15 17 27 5 28 29 23 8 13 11 9 30 12 21 4 3
3	1	23 3 28 4 8 17 12 29 15 1 18 11 9 14 27 22 7 21 20 19 6 25 10 24 16 2 30 16 26 5
	2	18 24 28 13 8 25 20 4 21 2 22 12 17 14 30 23 26 9 3 15 1 7 16 19 29 5 11 6 10 27
	3	6 20 30 13 10 26 23 25 12 11 21 7 17 1 4 14 9 27 18 19 29 16 22 15 5 3 2 28 8 25

TABLE 3 (cont.)

Test Sheet Orders

Subject	Condition	Test Sheet Order
3	3	22 7 17 29 1 8 16 2 24 20 3 14 26 27 10 15 25 12 30 6 9 13 4 18 11 28 5 21 19 23
	2	10 1 29 15 8 26 17 13 5 3 23 20 18 22 19 27 25 30 24 12 11 7 28 9 16 6 21 14 2 4
	1	22 4 10 6 9 14 5 23 21 28 19 15 27 17 8 11 7 3 24 30 2 18 29 25 16 20 12 1 13 26
4	3	15 19 17 16 2 23 7 5 9 21 30 28 3 8 25 27 1 14 11 24 26 22 29 12 10 18 6 13 4 20
	2	23 16 12 15 26 13 14 8 1 27 3 22 19 11 28 24 6 21 30 7 18 4 5 9 2 29 25 17 10 20
	1	21 1 2 20 23 16 28 4 29 30 19 6 3 9 15 12 26 13 5 27 22 14 7 11 17 18 10 24 25 8
	1	23 4 12 13 9 21 8 3 2 11 5 16 14 15 22 18 10 1 7 29 24 28 25 26 30 17 6 20 19 27
	2	22 28 30 29 19 11 16 9 7 4 21 12 15 13 5 24 27 8 18 23 3 2 25 1 14 20 6 26 17 10
	3	30 6 14 24 15 28 29 5 16 21 4 19 13 8 17 9 22 7 2 3 26 23 20 1 12 18 25 10 27 11
5	1	28 2 15 19 3 11 9 8 14 5 17 24 30 22 6 20 23 27 10 18 25 1 16 12 29 4 21 7 26 13
	3	7 14 20 15 8 18 16 6 12 11 5 22 26 27 10 4 19 17 13 30 25 23 21 29 24 9 2 28 1 3
	2	15 17 11 25 3 30 5 9 23 13 22 7 1 24 2 4 27 20 29 26 28 14 12 21 16 8 6 10 19 18
	2	23 29 18 14 22 5 10 2 11 6 21 7 20 27 26 1 19 16 8 13 3 12 15 28 4 30 17 25 24 9
	3	13 5 20 7 23 14 2 6 8 4 10 12 29 11 22 25 15 18 17 19 27 30 9 3 21 16 1 28 26 24
	1	22 30 15 29 26 8 7 19 3 28 9 17 21 10 18 20 14 12 5 13 24 11 27 16 1 26 6 4 23 2

TABLE 3 (cont.)

Test Sheet Orders

Subject	Condition	Test Sheet Order
6	2	12 10 11 23 1 27 18 5 26 22 24 16 28 14 4 19 21 2 15 13 8 7 20 3 30 25 17 6 29 9
	1	17 1 13 23 3 16 15 22 21 19 25 18 14 11 30 6 29 8 26 7 9 12 10 5 20 4 28 27 2 24
	3	7 17 30 4 11 13 26 24 2 27 14 6 12 9 21 23 20 15 16 18 8 10 22 5 19 3 28 29 25 1
	3	8 24 29 12 14 7 13 23 11 25 3 22 9 2 15 1 10 17 5 21 28 4 19 27 6 20 26 30 18 16
	1	1 22 23 24 19 16 20 30 5 13 11 17 26 12 28 9 25 18 10 6 3 21 15 7 4 8 14 29 2 27
	2	24 23 8 9 17 3 30 15 16 28 6 19 25 20 13 4 29 7 27 22 14 12 10 1 26 2 21 5 18 11

TABLE 4

Defect Order in Condition 1

Subject	Trial	Defect Order
1	1	C A E B D
	2	D A C E B
2	1	A C D B E
	2	D E B C A
3	1	C D A E B
	2	A D C E B
4	1	B D E A C
	2	A D C E B
5	1	C E B A D
	2	E C A B D
6	1	B D E C A
	2	A B E C D

a time. The five defects were divided into two groups, group ABC and group DE, according to their similar characteristics. Group ABC contained defects A, B, and C while group DE contained defects D and E. The subject inspected all matrices for one group of defects simultaneously; then for another group of defects simultaneously. For example, the subject inspected all matrices for defects D and E simultaneously; then for defects A, B, and C simultaneously. The order of groups was arranged randomly (Table 5).

Before the inspection, the subjects were asked to memorize the characteristics of the five defects and were given a practice session of 48 matrices for each of the three conditions to reduce the learning effect on their performances. Then they were informed of the results of this practice session to ensure that they had understood the tasks throughly. All conditions were self-paced with a five minute break after each condition and a thirty minute lunch break after finishing three conditions. The subjects were instructed to use their favored scanning patterns and be consistent. They were told to inspect the matrices as fast as they could but not at the cost of missing any defects. Figure 10 gives the instructions.

After finishing the inspection, the subjects were asked to rated the degree of discriminability for each defect on two scales, the Adjective Rating Scale and the Semantic Differential Scale (Figure 11). When using the Adjective Rating Scale, the subjects were asked to state how hard they felt it was to discriminate the defect from its background objects. Then they

TABLE 5

Defect Group Order in Condition 3

Subject	Trial	Defect Group Order
1	1	ABC DE
	2	DE ABC
2	1	DE ABC
	2	ABC DE
3	1	ABC DE
	2	DE ABC
4	1	DE ABC
	2	ABC DE
5	1	DE ABC
	2	ABC DE
6	1	ABC DE
	2	DE ABC

Informed Consent and Instructions

The purpose of this experiment is to investigate how an inspection task can be improved by inspection methods.

The inspected item is a 5x5 matrix which is composed of numbers, letters, or symbols. There are twelve matrices on each sheet. Five types of defects are defined. A matrix can contain more than one defect. Your job is to identify the defects and underline them with a blue pen. Please memorize the characteristics of these defects before the beginning of inspection.

Defect A: Occurrence of the letter "A" in a matrix. For example:

F	7	6	9	5
4	3	A	8	2
6	3	8	6	7
2	F	R	5	4
4	8	6	7	3

Defect B: Occurrence of the letter "B" in a matrix. For example:

8	5	3	4	B
3	7	R	6	9
7	5	6	4	9
F	F	R	6	3
R	3	5	2	7

Defect $: Occurrence of the symbol "$" in a matrix. For example:

8	F	8	3	R
4	8	7	5	9
2	8	5	9	F
7	4	7	3	3
3	R	F	2	$

Figure 10. Instructions for subjects.

<u>Defect 345</u>: Occurrence of the number "345" in a row of a matrix.

For example:

$$\begin{matrix} 9 & R & 3 & 5 & F \\ 3 & 2 & 7 & 5 & 6 \\ 4 & \underline{3} & \underline{4} & \underline{5} & 4 \\ F & R & 3 & 8 & 5 \\ 7 & 3 & 5 & 4 & F \end{matrix}$$

<u>Defect 678</u>: Occurrence of the number "678" in a row of a matrix.

For example:

$$\begin{matrix} R & 3 & 8 & 5 & 6 \\ 8 & 6 & 7 & 6 & 5 \\ 4 & F & R & 4 & 7 \\ \underline{6} & \underline{7} & \underline{8} & 7 & 6 \\ 5 & F & 3 & 6 & R \end{matrix}$$

You will inspect 360 matrices (30 sheets) in each of the following three conditions twice in an assigned order.

<u>Condition 1</u>: Here you will inspect all matrices for one type of defect, then for the next type of defect, and so on.

<u>Condition 2</u>: Here you will inspect each matrix for all five defects simultaneously before going on to the next matrix.

<u>Condition 3</u>: Here you will inspect all matrices for a few defects simultaneously then for the rest of the defects simultaneously.

There will be a practice session of 48 matrices for each of the three conditions before the experimental run. Then you are informed of the results of this practice session to ensure that you have understood the tasks throughly. A thirty minute lunch break and five five minute short breaks will be given, during which you can have lunch or relax. You can use your

<u>Figure 10 (cont.). Instructions for subjects.</u>

favored scanning pattern in searching but please be consistent throughout the experiment. Since your performance is evaluated by both speed and accuracy, please inspect the matrices as fast as you can but not at the cost of missing any defects. After finishing the inspection, you will be asked to fill out the rating sheets.

The experiment takes about 6 hours. You will be paid $25 to complete it. Your name will be kept confidential. There is no risk or discomfort involved. However, you are free to stop at any time. I will appreciate if you will complete the experiment.

If you agree to participate in this experiment, please sign the consent form. Thank you for your cooperation.

____

(Signature)

____

(Date)

Figure 10 (cont.). Instructions for subjects.

Instructions and Ratings

You are now going to state how hard you felt to discriminate the defects from their background objects. According to the following scale, please place each defect name (A, B, $, 345, and 678) on the scale which best describes its ease of detection. Two defects can be placed on the same position.

very easy

easy

somewhat easy

neither easy nor hard

somewhat hard

hard

very hard

Figure 11. Instructions for the Adjective Rating Scale and the Semantic Differential Scale.

You are now going to describe how you feel about the defects by the following four scales. Each scale has two adjectives opposite in meaning to each other. The scale is divided into seven segments. If you would describe the defect as extremely clear, place an 'X' on the scale as shown below:

clear _X_:___:___:___:___:___:___ confused

The meannings of the segments in order from left to right are: extremely clear, very clear, slightly clear, neutral, slightly confused, very confused, extremely confused. Please place the X on the position which best describes your impression.

Defect A

clear	___:___:___:___:___:___:___	confused
difficult	___:___:___:___:___:___:___	easy
undiscriminable	___:___:___:___:___:___:___	discriminable
apparent	___:___:___:___:___:___:___	hidden

Defect B

clear	___:___:___:___:___:___:___	confused
difficult	___:___:___:___:___:___:___	easy
undiscriminable	___:___:___:___:___:___:___	discriminable
apparent	___:___:___:___:___:___:___	hidden

Figure 11 (cont.). Instructions for the Adjective Rating Scale and the Semantic Differential Scale.

Defect $

clear ___:___:___:___:___:___:___ confused

difficult ___:___:___:___:___:___:___ easy

undiscriminable ___:___:___:___:___:___:___ discriminable

apparent ___:___:___:___:___:___:___ hidden

Defect 345

clear ___:___:___:___:___:___:___ confused

difficult ___:___:___:___:___:___:___ easy

undiscriminable ___:___:___:___:___:___:___ discriminable

apparent ___:___:___:___:___:___:___ hidden

Defect 678

clear ___:___:___:___:___:___:___ confused

difficult ___:___:___:___:___:___:___ easy

undiscriminable ___:___:___:___:___:___:___ discriminable

apparent ___:___:___:___:___:___:___ hidden

Figure 11 (cont.). Instructions for the Adjective Rating Scale and the Semantic Differential Scale.

quoted the appropriate ratings on the scale. When using the Semantic Differential Scale, the subjects were asked to state how they felt about the defect by placing the X's on the segments of the scales which best described their impressions. The points for scales were assigned after the experiment as in Figure 12.

The experiment was held in IE Lab III in the Industrial Engineering Department at Kansas State University. A white table was used as the work station. It was 5 feet by 2.5 feet in area and 2.4 feet in height. The subject was seated on an adjustable chair. Figure 13 shows the layout. The data were obtained during the day. The lighting level at work level was 1080 lux. The noise level was 43 dBA. The room temperature was 72° F.

Subjects

Subjects were 6 non-student females. Women were utilized because large numbers of women currently are employed as inspectors in industry. Age ranged from 20 to 53 years, with a median age of 33. They were tested by the Titmus Vision Tester to have at least normal or corrected normal 20/20 vision. Each of them performed under all of three conditions twice. All conditions were self-paced. Each subject was paid $25 for her service. The subjects knew that the experiment was a study of inspection methods and their performances were evaluated by the criteria of speed and accuracy but not the details. For convenience, the subjects were identified Subject 1 to Subject 6 according to the order of their experimental runs.

Adjective Rating Scale:

13	very easy
12	
11	easy
10	
9	somewhat easy
8	
7	neither easy nor hard
6	
5	somewhat hard
4	
3	hard
2	
1	very hard

Semantic Differential Scale:

clear	7	6	5	4	3	2	1	confused
difficult	1	2	3	4	5	6	7	easy
undiscriminable	1	2	3	4	5	6	7	discriminable
apparent	7	6	5	4	3	2	1	hidden

Figure 12. Points assigned to the Adjective Rating Scale and the Semantic Differential Scale.

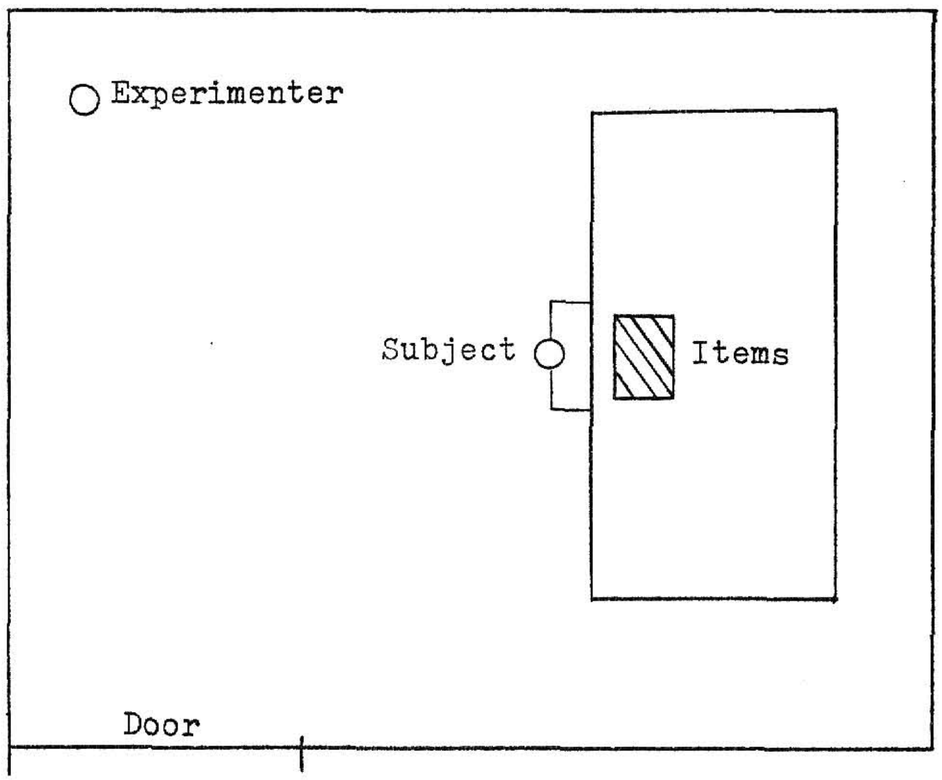

Figure 13. Layout of the work station.

Procedure

The procedure was as follows:

1) The subject was tested by the Titmus Vision Tester to have at least normal or corrected normal 20/20 vision.
2) The subject read the Instructions and signed the Consent Form (Figure 10).
3) The subject was asked to memorize the characteristics of the five defects.
4) The subject had a practice session of 48 matrices per condition. Then she was informed of the results to ensure that she had understood the tasks throughly.
5) A five minute break was given after the practice session.
6) The subject performed under the first assigned condition.
7) A five minute break was given after the condition.
8) The subject performed under the second assigned condition.
9) A five minute break was given after the condition.
10) The subject performed under the third assigned condition.
11) A thirty minute lunch break was given after the condition.
12) The subject performed under the fourth assigned condition.
13) A five minute break was given after the condition.
14) The subject performed under the fifth assigned condition.
15) A five minute break was given after the condition.
16) The subject performed under the sixth assigned condition.
17) The subject made the ratings (Figure 11).

RESULTS

Inspection performance was evaluated by inspection time and error percentages. In inspection, the inspector can make two types of errors: type I error (false alarm-- classifying an acceptable item as rejectable) and type II error (missed signal-- classifying a rejectable item as acceptable). Error percentages were: 1) percentage of type I errors, which is the ratio of false alarms to acceptable items in the lot, 2) percentage of type II errors, which is the ratio of missed signals to rejectable items in the lot, 3) percentage of incorrect decisions, which is the ratio of incorrect decisions (false alarms + missed signals) to total decisions in the lot. Tables 6 through 9 show the data of inspection times and error percentages.

Analysis of Variance (ANOVA) and Duncan's Multiple Range Test (MRT) methods were used to analyze the data. A significance level of 5% was used throughout the analyses.

Defect Discriminability

Table 10 lists the inspection times, the inspection error rates, the Adjective ratings (ADJ), and the Semantic ratings (SEM) of the five defects. Table 11 lists their correlation coefficients. The two ratings were very consistent ($r = 0.84$). Tables 12 and 13 show the ANOVA for ratings of the five defects. Defect and subject were significant. Tables 14 and 15 show the mean ratings. E and D (5.5 and 6.5 for ADJ; 13.2 and 13.7 for SEM) were the least discriminable defects; C (12.3 for ADJ; 26.3 for SEM) was the most; B and A (9.7 and 10.5 for ADJ; 16.7 and 19.8 for SEM)

TABLE 6

Inspection Time (min.) in The Three Conditions

Subject	Trial	Condition	TABC	TDE	TALL
1	1	1	14.85	23.57	38.42
		2			36.94
		3	8.93	19.58	28.51
	2	1	12.89	22.59	35.48
		2			20.45
		3	8.14	22.25	30.39
2	1	1	16.83	22.90	39.73
		2			29.82
		3	13.26	30.16	43.42
	2	1	18.32	18.28	36.60
		2			19.41
		3	9.28	22.95	32.23
3	1	1	28.60	39.43	68.03
		2			57.22
		3	10.94	29.11	40.05
	2	1	27.39	34.66	62.05
		2			44.99
		3	17.16	24.17	41.33
4	1	1	60.79	62.22	123.01
		2			54.48
		3	28.58	45.52	74.10
	2	1	27.19	35.77	62.96
		2			51.09
		3	16.11	32.11	48.22
5	1	1	46.57	52.87	99.44
		2			74.49
		3	21.66	39.40	61.06
	2	1	32.99	54.71	87.70
		2			72.28
		3	19.88	40.28	60.16
6	1	1	35.04	39.92	74.96
		2			42.49
		3	14.46	31.17	45.63
	2	1	29.22	37.77	66.99
		2			34.08
		3	18.59	26.89	45.48
Mean		1	29.22	37.06	66.28
		2			44.81
		3	15.58	30.30	45.88

TABC = Inspection time for easy defects A, B, and C.

TDE = Inspection time for difficult defects D and E.

TALL = Inspection time for all defects A, B, C, D, and E.

TABLE 7

Type I Error Rates (%) in The Three Conditions

Subject	Trial	Condition	EABC	EDE	EALL
1	1	1	0	0	0
		2			0
		3	0	0	0
	2	1	0	0	0
		2			0
		3	0	0	0
2	1	1	0	0	0
		2			0
		3	0	0	0
	2	1	0	0	0
		2			0
		3	0	0	0
3	1	1	0	0	0
		2			0.12
		3	0	0.30	0.12
	2	1	0	0	0
		2			0.12
		3	0	0	0
4	1	1	0	0	0
		2			0
		3	0	0	0
	2	1	0	0	0
		2			0
		3	0	0	0
5	1	1	0	0	0
		2			0
		3	0	0	0
	2	1	0	0	0
		2			0
		3	0	0	0
6	1	1	0	0	0
		2			0
		3	0	0	0
	2	1	0	0	0
		2			0
		3	0	0	0
Mean		1	0	0	0
		2			0.02
		3	0	0.02	0.01

EABC = Inspection error rate for easy defects A, B, and C.

EDE = Inspection error rate for difficult defects D and E.

EALL = Inspection error rate for all defects A, B, C, D, and E.

TABLE 8

Type II Error Rates (%) in The Three Conditions

Subject	Trial	Condition	EABC	EDE	EALL
1	1	1	12.00	46.00	25.60
		2			25.60
		3	9.33	44.00	23.20
	2	1	20.00	52.00	32.80
		2			24.00
		3	24.00	46.00	32.80
2	1	1	22.66	50.00	33.60
		2			28.00
		3	21.33	30.00	24.80
	2	1	13.33	48.00	27.20
		2			40.00
		3	8.00	44.00	22.40
3	1	1	12.00	24.00	16.80
		2			9.60
		3	18.66	32.00	24.00
	2	1	24.00	20.00	22.40
		2			7.20
		3	9.33	36.00	20.00
4	1	1	0.00	6.00	2.40
		2			5.60
		3	1.33	16.00	7.20
	2	1	13.33	8.00	11.20
		2			13.60
		3	4.00	18.00	9.60
5	1	1	9.33	16.00	12.00
		2			14.40
		3	18.66	28.00	22.40
	2	1	13.33	28.00	19.20
		2			16.80
		3	14.66	54.00	30.40
6	1	1	5.33	10.00	7.20
		2			8.80
		3	10.66	18.00	13.60
	2	1	6.66	20.00	12.00
		2			10.40
		3	16.00	22.00	18.40
Mean		1	12.67	27.33	18.53
		2			17.00
		3	13.00	32.33	20.73

EABC = Inspection error rate for easy defects A, B, and C.

EDE = Inspection error rate for difficult defects D and E.

EALL = Inspection error rate for all defects A, B, C, D, and E.

TABLE 9

Incorrect Decision Rates (%) in The Three Conditions

Subject	Trial	Condition	EABC	EDE	EALL
1	1	1	0.83	3.19	1.77
		2			1.77
		3	0.64	3.05	1.61
	2	1	1.38	3.61	2.27
		2			1.66
		3	1.66	3.19	2.27
2	1	1	1.57	3.47	2.33
		2			1.94
		3	1.48	2.08	1.72
	2	1	0.92	3.33	1.88
		2			2.77
		3	0.55	3.05	1.55
3	1	1	0.83	1.66	1.16
		2			0.77
		3	1.29	2.50	1.77
	2	1	1.66	1.38	1.55
		2			0.61
		3	0.64	2.50	1.38
4	1	1	0.00	0.41	0.16
		2			0.38
		3	0.09	1.11	0.50
	2	1	0.92	0.55	0.77
		2			0.94
		3	0.27	1.25	0.66
5	1	1	0.64	1.11	0.83
		2			1.00
		3	1.29	1.94	1.55
	2	1	0.92	1.94	1.33
		2			1.16
		3	1.01	3.75	2.11
6	1	1	0.37	0.69	0.50
		2			0.61
		3	0.74	1.25	0.94
	2	1	0.46	1.38	0.83
		2			0.72
		3	1.11	1.52	1.27
Mean		1	0.88	1.90	1.29
		2			1.20
		3	0.90	2.27	1.45

EABC = Inspection error rate for easy defects A, B, and C.

EDE = Inspection error rate for difficult defects D and E.

EALL = Inspection error rate for all defects A, B, C, D, and E.

TABLE 10

Inspection Times, Error Rates, and Ratings for The Five Defects

Subject	Defect	ADJ (pt.)	SEM (pt.)	Time (min.)	Error Rate (%)
1	A	11	21	5.24	0.65
	B	11	24	5.99	2.50
	C	13	28	2.63	0
	D	6	12	12.48	2.92
	E	6	10	10.59	3.43
2	A	12	23	8.37	0.56
	B	12	19	5.79	2.92
	C	13	28	3.41	0
	D	11	23	10.69	3.38
	E	11	23	9.90	3.33
3	A	11	13	10.67	0.19
	B	9	13	11.57	2.18
	C	13	28	5.75	0.14
	D	5	6	17.65	1.67
	E	3	5	19.39	1.90
4	A	11	20	17.49	0.28
	B	9	20	16.43	0.60
	C	13	23	10.06	0
	D	3	10	25.17	1.34
	E	3	10	23.82	0.65
5	A	5	14	14.24	0.83
	B	7	9	15.93	1.99
	C	9	23	9.60	0.23
	D	7	9	25.76	1.57
	E	3	9	28.03	2.04
6	A	13	28	12.68	0.83
	B	10	15	12.22	0.83
	C	13	28	7.23	0.05
	D	7	22	20.64	1.20
	E	7	22	18.20	1.16
Mean	A	10.5	19.8	11.45	0.56
	B	9.7	16.7	11.32	1.84
	C	12.3	26.3	6.45	0.07
	D	6.5	13.7	18.73	2.01
	E	5.5	13.2	18.32	2.08

ADJ = Adjective rating (low number = low discriminability).

SEM = Semantic rating (low number = low discriminability).

TABLE 11

Correlations Between Inspection Time, Error Rate, ADJ, and SEM

Pearson Correlation Coefficients / N = 30

	ADJ	SEM	Time	Error
ADJ	1.000 0.0000	0.839 0.0001	-0.806 0.0001	-0.326 0.0788
SEM	0.839 0.0001	1.000 0.0000	-0.671 0.0001	-0.388 0.0340
Time	-0.806 0.0001	-0.671 0.0001	1.000 0.0000	0.129 0.4956
Error	-0.326 0.0788	-0.388 0.0340	0.129 0.4956	1.000 0.0000

Spearman Correlation Coefficients / N = 30

	ADJ	SEM	Time	Error
ADJ	1.000 0.0000	0.848 0.0001	-0.785 0.0001	-0.500 0.0049
SEM	0.848 0.0001	1.000 0.0000	-0.701 0.0001	-0.485 0.0066
Time	-0.785 0.0001	-0.701 0.0001	1.000 0.0000	0.307 0.0988
Error	-0.500 0.0049	-0.485 0.0066	0.307 0.0988	1.000 0.0000

TABLE 12

ANOVA for Adjective Rating (points) of The Five Defects

Source of Variance	DF	MS	F	Alpha Hat
Defect (D)	4	48.38	15.90	.0001
Subject (S)	5	18.86	6.20	.0013
Error	20	3.04		
Total	29			

TABLE 13

ANOVA for Semantic Rating (points) of The Five Defects

Source of Variance	DF	MS	F	Alpha Hat
Defect (D)	4	175.05	10.46	.0001
Subject (S)	5	107.01	6.40	.0011
Error	20	16.73		
Total	29			

TABLE 14

Duncan's MRT for Adjective Rating (points) of The Five Defects

Means with the same letter are not significantly different.

Alpha level = .05 DF = 20 MS = 3.04

Defect	Mean	N	Grouping		
C	12.3	6	A		
A	10.5	6	A	B	
B	9.7	6		B	
D	6.5	6			C
E	5.5	6			C

TABLE 15

Duncan's MRT for Semantic Rating (points) of The Five Defects

Means with the same letter are not significantly different.

Alpha level = .05 DF = 20 MS = 16.73

Defect	Mean	N	Grouping	
C	26.3	6	A		
A	19.8	6	B	
B	16.7	6	B	C
D	13.7	6		C
E	13.2	6		C

were in the middle.

Inspection time and ratings were highly correlated (r = -0.81 for time vs ADJ; r = -0.67 for time vs SEM). Figures 14 and 15 illustrate their relationships. Table 16 shows the ANOVA for inspection time of the five defects. Defect and subject were significant. Table 17 shows the mean times. Defects D (18.73 min.) and E (18.32 min.) had the longest time; C (6.45 min.) had the least; A (11.45 min.) and B (11.32 min.) were in the middle.

Inspection error rate, using the Pearson coefficient, was not significantly related to time. The correlation vs ADJ was -0.33 (p = .08) and vs SEM was -0.39 (p = .03). Using the Spearman coefficient, error rate still was not related to time but the correlations increased for both ADJ (-0.50; p = .005) and SEM (-0.49; p = .007). Table 18 shows the ANOVA for inspection error rate of the five defects. Defect and subject were significant. Table 19 shows the mean error rates. There were significantly more inspection errors on defects E, D, and B (2.08, 2.01, and 1.84%) than on defects A and C (0.56 and 0.07%).

<u>Inspection Time</u>

Table 20 shows the ANOVA for the mixed level. Method, subject, and trial were significant. Table 21 shows the mean times. Inspection time (66.3 min.) in Method 1 (inspecting for one defect at a time) was significantly longer than in the other two methods (44.8 min. in inspecting for all defects at a time and 45.9 min. in inspecting for one group of similar defects at a time).

Table 22 shows the ANOVA for the easy level. Method and

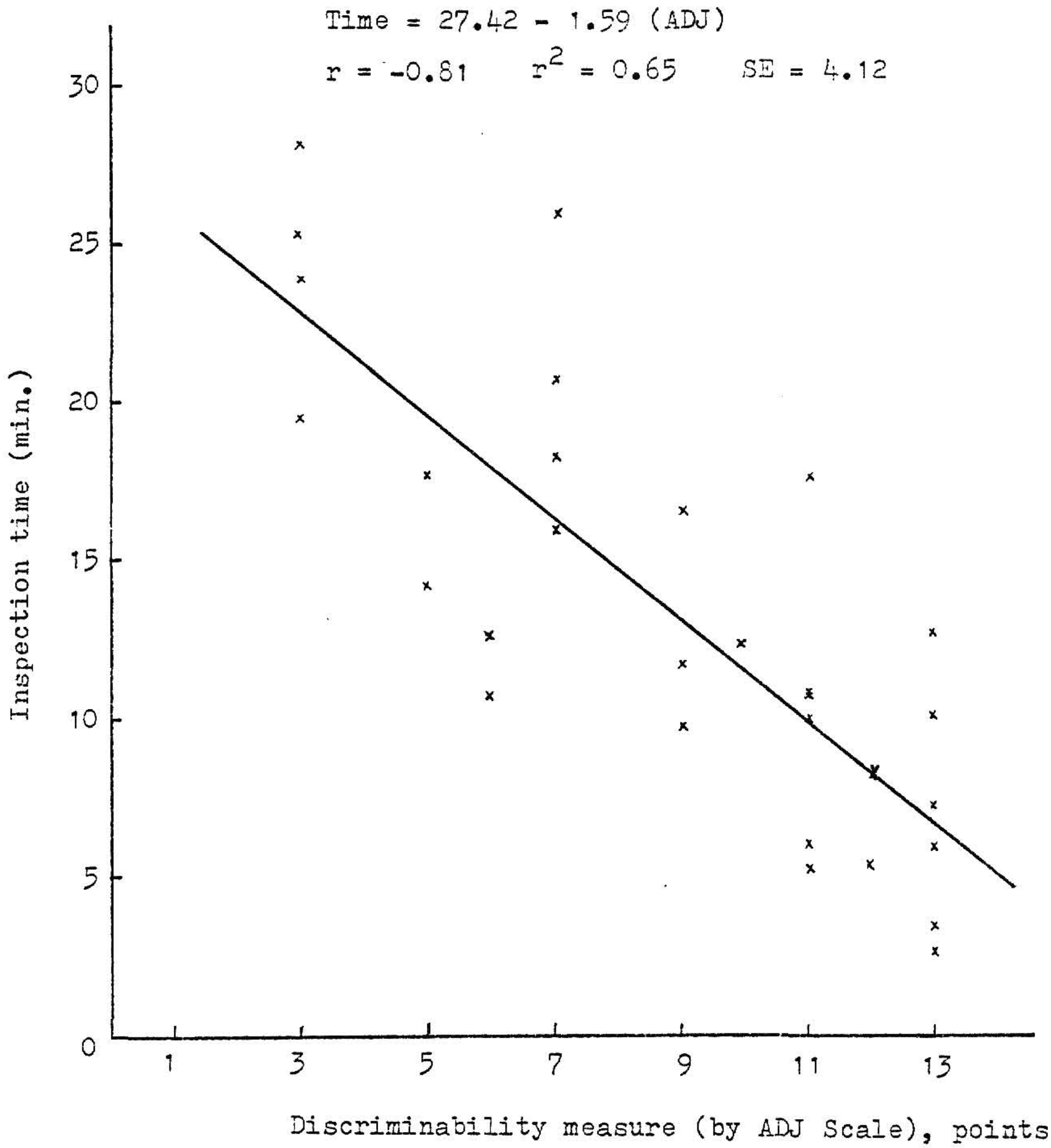

Figure 14. Relationship between inspection time and discriminability (measured by Adjective Rating Scale).

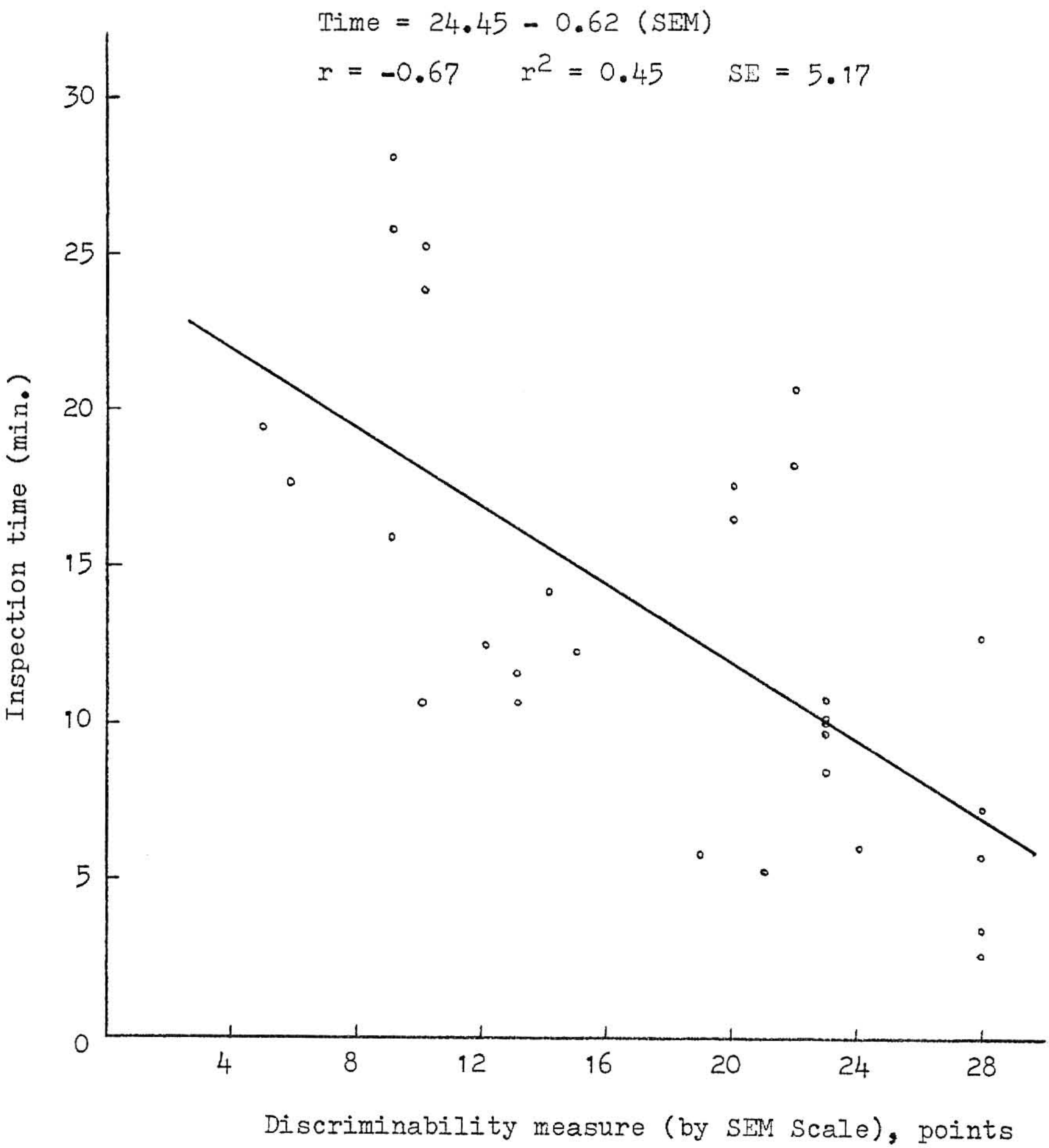

Figure 15. Relationship between inspection time and discriminability (measured by Semantic Differential Rating).

TABLE 16

ANOVA for Inspection Time (min.) of The Five Defects

Source of Variance	DF	MS	F	Alpha Hat
Defect (D)	4	163.53	41.48	.0001
Subject (S)	5	125.29	31.78	.0001
Error	20	3.94		
Total	29			

TABLE 17

Duncan's MRT for Inspection Time (min.) of The Five Defects

Means with the same letter are not significantly different.

Alpha level = .05 DF = 20 MS = 3.94

Defect	Mean	N	Grouping
D	18.73	6	A
E	18.32	6	A
A	11.45	6	B
B	11.32	6	B
C	6.45	6	C

TABLE 18

ANOVA for Inspection Error Rate (%) of The Five Defects

Source of Variance	DF	MS	F	Alpha Hat
Defect (D)	4	5.22	15.42	.0001
Subject (S)	5	1.68	4.94	.0041
Error	20	0.34		
Total	29			

TABLE 19

Duncan's MRT for Inspection Error Rate (%) of The Five Defects

Means with the same letter are not significantly different.

Alpha level = .05 DF = 20 MS = 0.34

Defect	Mean	N	Grouping	
E	2.08	6	A	
D	2.01	6	A	
B	1.84	6	A	
A	0.56	6		B
C	0.07	6		B

TABLE 20

ANOVA for Inspection Time (min.) in The Mixed Level

Source of Variance	DF	MS	F	Alpha Hat
Method (M)	2	1756.39	14.68	.0001
Subject (S)	5	1931.98	16.15	.0001
Trial (T)	1	899.10	7.52	.0107
Error	27	119.61		
Total	35			

TABLE 21

Duncan's MRT for Inspection Time (min.) in The Mixed Level

Means with the same letter are not significantly different.

Alpha level = .05 DF = 27 MS = 119.61

Method	Mean	N	Grouping
1	66.28	12	A
3	45.88	12	B
2	44.81	12	B

Method 1 = Inspect for one defect at a time.

Method 2 = Inspect for all defects at a time.

Method 3 = Inspect for one group of similar defects at a time.

TABLE 22

ANOVA for Inspection Time (min.) in The Easy Level

Source of Variance	DF	MS	F	Alpha Hat
Method (M)	1	1116.43	21.13	.0003
Subject (S)	5	298.07	5.64	.0035
Trial (T)	1	167.22	3.17	.0942
Error	16	52.83		
Total	23			

subject were significant. Table 23 shows the mean times. Inspection time (29.2 min.) in Method 1 (inspecting for one defect at a time) was significantly longer than the 15.6 min. in Method 2 (inspecting for all defects at a time).

Table 24 shows the ANOVA for the difficult level. Method, subject, trial, the interaction between method and subject, and the interaction between subject and trial were significant. Table 25 shows the mean times. Inspection time (37.1 min.) in Method 1 (inspecting for one defect at a time) was significantly longer than the 30.3 min. in Method 2 (inspecting for all defects at a time).

Method was found to have a significant effect on inspection time in all three levels. Inspecting for one defect at a time took a significantly longer time than inspecting for all defects at a time or inspecting for one group of similar defects at a time.

Type I Errors

Only Subject 3 committed type I errors. All of her errors were on defect D. She took "354" as "345" six times (four times in Condition 2 and two times in Condition 3).

Table 26 shows the ANOVA for the mixed level. Subject and the interaction between method and subject were significant.

There were no type I errors in the easy level.

Table 27 shows the ANOVA for the difficult level. Nothing was significant.

Method had no significant effect on type I errors in all three levels.

TABLE 23

Duncan's MRT for Inspection Time (min.) in The Easy Level

Means with the same letter are not significantly different.

Alpha level = .05   DF = 16   MS = 52.83

Method	Mean	N	Grouping
1	29.22	12	A
2	15.58	12	B

Method 1 = Inspect for one defect at a time.

Method 2 = Inspect for all defects at a time.

TABLE 24

ANOVA for Inspection Time (min.) in The Difficult Level

Source of Variance	DF	MS	F	Alpha Hat
Method (M)	1	270.68	32.82	.0012
Subject (S)	5	415.29	50.35	.0001
Trial (T)	1	164.96	20.00	.0042
M*S	5	53.51	6.49	.0207
S*T	5	60.43	7.33	.0155
Error	6	8.25		
Total	23			

TABLE 25

Duncan's MRT for Inspection Time (min.) in The Difficult Level

Means with the same letter are not significantly different.

Alpha level = .05 DF = 6 MS = 8.25

Method	Mean	N	Grouping
1	37.06	12	A
2	30.30	12	B

Method 1 = Inspect for one defect at a time.

Method 2 = Inspect for all defects at a time.

TABLE 26

ANOVA for Type I Errors (%) in The Mixed Level

Source of Variance	DF	MS	F	Alpha Hat
Method (M)	2	.0012	3.00	.0766
Subject (S)	5	.0036	9.00	.0003
Trial (T)	1	.0004	1.00	.3313
M*S	10	.0012	3.00	.0224
Error	17	.0004		
Total	35			

TABLE 27

ANOVA for Type I Errors (%) in The Difficult Level

Source of Variance	DF	MS	F	Alpha Hat
Method (M)	1	.0037	1.00	.3322
Subject (S)	5	.0037	1.00	.4489
Trial (T)	1	.0037	1.00	.3322
Error	16	.0037		
Total	23			

Type II Errors

Table 28 shows the ANOVA for the mixed level. Subject and trial were significant.

Table 29 shows the ANOVA for the easy level. Nothing was significant.

Table 30 shows the ANOVA for the difficult level. Subject, trial, and the interaction between method and subject were significant.

Method had no significant effect on type II errors in all three levels.

Incorrect Decisions

Table 31 shows the ANOVA for the mixed level. Subject and trial were significant.

Table 32 shows the ANOVA for the easy level. Nothing was significant.

Table 33 shows the ANOVA for the difficult level. Method, subject, trial, and the interaction between method and subject were significant. Table 34 shows the mean error rates. Incorrect decisions (1.90%) in Method 1 (inspecting for one defect at a time) were significantly less than the 2.27% in Method 2 (Inspecting for all defects at a time).

Method had a significant effect on incorrect decisions only in the difficult level where inspecting for one defect at a time made significantly fewer incorrect decisions than inspecting for all defects at a time.

TABLE 28

ANOVA for Type II Errors (%) in The Mixed Level

Source of Variance	DF	MS	F	Alpha Hat
Method (M)	2	42.26	1.60	.2195
Subject (S)	5	419.23	15.92	.0001
Trial (T)	1	119.54	4.54	.0424
Error	27	26.33		
Total	35			

TABLE 29

ANOVA for Type II Errors (%) in The Easy Level

Source of Variance	DF	MS	F	Alpha Hat
Method (M)	1	.67	.02	.8964
Subject (S)	5	90.09	2.36	.0868
Trial (T)	1	26.74	.70	.4145
Error	16	38.09		
Total	23			

TABLE 30

ANOVA for Type II Errors (%) in The Difficult Level

Source of Variance	DF	MS	F	Alpha Hat
Method (M)	1	150.00	4.49	.0576
Subject (S)	5	755.47	22.62	.0001
Trial (T)	1	240.67	7.21	.0212
M*S	5	128.00	3.83	.0295
Error	11	33.39		
Total	23			

TABLE 31

ANOVA for Incorrect Decisions (%) in The Mixed Level

Source of Variance	DF	MS	F	Alpha Hat
Method (M)	2	.1930	1.55	.2299
Subject (S)	5	2.0055	16.14	.0001
Trial (T)	1	.5487	4.42	.0451
Error	27	.1242		
Total	35			

TABLE 32

ANOVA for Incorrect Decisions (%) in The Easy Level

Source of Variance	DF	MS	F	Alpha Hat
Method (M)	1	.0032	.02	.8964
Subject (S)	5	.4345	2.36	.0868
Trial (T)	1	.1290	.70	.4145
Error	16	.1837		
Total	23			

TABLE 33

ANOVA for Incorrect Decisions (%) in The Difficult Level

Source of Variance	DF	MS	F	Alpha Hat
Method (M)	1	.8230	4.89	.0491
Subject (S)	5	3.6323	21.58	.0001
Trial (T)	1	1.0417	6.19	.0302
M*S	5	.6475	3.85	.0292
Error	11	.1684		
Total	23			

TABLE 34

Duncan's MRT for Incorrect Decisions (%) in The Difficult Level

Means with the same letter are not significantly different.

Alpha level = .05 DF = 11 MS = .1684

Method	Mean	N	Grouping
2	2.27	12	A
1	1.90	12	B

Method 1 = Inspect for one defect at a time.

Method 2 = Inspect for all defects at a time.

DISCUSSION

There were five types of defects defined (A, B, C, D, and E). A, B, and C were grouped as similar and easy defects because they were designated by a single letter or symbol. D and E were grouped as similar and difficult defects because they were designated by three continuous digits and were patterned with background objects.

There were three inspection methods compared: 1) Inspect for one defect at a time. 2) Inspect for all defects at a time. 3) Inspect for one group of similar defects at a time.

There were three difficulty levels of defects conducted: 1) Mixed level-- All of the five defects needed to be inspected. The three inspection methods were employed. 2) Easy level-- Only the three easy defects needed to be inspected. Method 1 and Method 2 were employed. 3) Difficult level-- Only the two difficult defects needed to be inspected. Method 1 and Method 2 were employed.

Defect Discriminability

The two ratings were very consistent ($r = 0.84$). D and E were rated as the least discriminable defects; A and B were in the middle; C was the most discriminable defect.

Inspection time was highly correlated with ratings (low rating = low discriminability). ADJ gave better prediction of inspection time than SEM. The less discriminable defect had a longer inspection time.

Inspection error rate was not consistent with time. Subjects tended to make more errors on the less discriminable defects. Here, unlike the results in ratings and inspection time, defect B was more difficult to detect than defect A. In other words, defect B was more difficult to inspect than the subjects had thought. The difficulty with B could be caused by similar "curved" background objects, such as 8 and R, while A, a "linear" letter, had no background objects with similar shapes.

Inspection Time

For the mixed level, inspecting for one defect at a time required a significantly longer time (about 50% longer) than the other two methods. But there was no significant difference of time between the latter two methods.

For the easy and difficult levels, inspecting for one defect at a time yielded a significantly longer time (88% longer in the easy level and 22% longer in the difficult level) than inspecting for all defects at a time.

Inspecting for less defects at a time needs more time for eye travel but saves some time for mental "gear shifting". The results showed that the eye travel time is more sensitive than the "gear shifting" time on inspection methods. In other words, the time spent on extra eye travel was not compensated from the time saved on less mental "gear shifting" (inspecting for one defect at a time yielded a significantly longer time than the other two methods). However, Method 2 and Method 3 in the mixed level showed a tradeoff of the two effects on time (no significant difference of inspection time between these two methods).

Inspection Errors

For the mixed level, there were no significant differences of the three error percentages among the three methods.

For the easy level, there were no significant differences of the three error percentages between Method 1 and Method 2.

For the difficult level, inspecting for one difficult defect at a time made significantly less incorrect decisions (17% less) than inspecting for both difficult defects at a time though there were no significant differences of type I and type II errors between the two methods.

Theoretically speaking, visual inspection is primarily a matter of comparing the inspected characteristic with a mental image that corresponds with an acceptable characteristic; inspection accuracy would be improved by reducing the number, frequency, and variety of mental images required at any given time (Harris and Chaney, 1969). This theory conflicts with the effect of defect rate on inspection accuracy. Decreasing the defect rate can reduce the number, frequency, or variety of mental images in inspection. According to the above theory, inspection accuracy should increase against the decrease in defect rate. But the previous experiments (Jenkins, 1958; Harris, 1968; and Fox and Haslegrave, 1969) showed that inspection accuracy dropped with reductions in defect rate and it accelerated as the defect rate approached zero. A reasonable explanation is that, during monotonous inspection work, the detection of defects has a stimulative effect that pulls the inspector out of the monotonous state. This can increase the inspection accuracy; this stimulation effect is more significant than the converse effect caused by increasing the mental load.

In multi-defect inspection situations, the defect rate decreases with the reductions of the number of defects inspected at a time. Decreasing the number reduces the mental "gear shifting" which has a positive effect on inspection accuracy. Decreasing the defect rate reduces the stimuli in the monotonous state which has a negative effect on inspection accuracy. The two effects come togather in the multi-defect inspections. The accuracy comparisons of the inspection methods in different inspection situations depend on which effect is significantly stronger than the other.

Ebeling and Orr (1978) found that the selection of the most accurate inspection method was influenced by the defective level. Inspecting for all defects at a time (simultaneous scanning method) was best at lower defective levels while inspecting for one defect at a time (sequential scanning method) was best at higher defective levels. For the low defective levels, the stimulation effect was so strong (Jenkin, 1958; Harris, 1968; and Fox and Haslegrave, 1969) as to overcome the "gear shifting" effect-- thus the simultaneous scanning was the best. For the high defective levels, the stimulation effect was not strong enough to balance off the "gear shifting" effect-- thus the sequential scanning was the best.

Table 35 gives the frequency of defect occurring (stimulus) for the present experiment. In the mixed level, the stimuli in Method 2 and Method 3 came about 45% more often than in Method 1

TABLE 35

Frequency (occur./min.) of Defect Occurring

Difficulty Level	M1	M2	M3
Mixed	125/66.3 = 1.9	125/44.8 = 2.8	125/45.9 = 2.7
Easy	75/29.2 = 2.6	75/15.6 = 4.8	
Difficult	50/37.1 = 1.3	50/30.3 = 1.6	

M1 = Method 1 = Inspect for one defect at a time.

M2 = Method 2 = Inspect for all defects at a time.

M3 = Method 3 = Inspect for one group of similar defects at a time.

((2.75 - 1.9)/1.9 = 45%). In the easy level, the stimuli in Method 2 came 85% more often than in Method 1 ((4.8 - 2.6)/2.6 = 85%). In the difficult level, the stimuli in Method 2 came only 23% more often than in Method 1 ((1.6 - 1.3)/1.3 = 23%). These might explain why the stimulation effect could balance off the "gear shifting" effect in the mixed and easy levels (no significant difference of inspection errors) but not in the difficult level (inspecting for one defect at a time made significantly less errors than inspecting for all defects at a time). Therefore the selection of the most accurate inspection method was influenced by the difficulty level of defects. Inspecting for one defect at a time was the most accurate method in the difficult level while no method dominated in the other two situations.

Consequently, the selection of the most accurate inspection method can be influenced by the defective level of items and the difficulty level of defects. More research is needed to evaluate the inspection methods under a wider range of defect difficulties.

Selection of Inspection Methods

Table 36 shows the comparisons of inspection methods in the three difficulty levels of defects.

In the mixed level, Method 1 spent more time and made no less errors but there were no differences between the other two methods. Therefore either Method 2 or Method 3 should be selected.

In the easy level, Method 1 spent more time and made no less errors than Method 2. Therefore Method 2 should be selected.

In the difficult level, Method 1 spent more time but made

TABLE 36

Comparison of Inspection Methods

	Local Optimal*				
Difficulty Level	Time	Type I Errors	Type II Errors	Incorrect Decisions	Global Optimal**
Mixed	M2,M3	Any	Any	Any	M2,M3
Easy	M2	Any	Any	Any	M2
Difficult	M2	Any	Any	M1	depends

M1 = Method 1 = Inspect for one defect at a time.

M2 = Method 2 = Inspect for all defects at a time.

M3 = Method 3 = Inspect for one group of similar defects at a time.

* The columns under 'Local Optimal' indicate that if the selection criterion is only one of them, the methods listed there are the optimals.

** The column under 'Global Optimal' lists the optimal methods by using all of the four criteria.

less incorrect decisions than Method 2. A cost analysis of time and errors is needed to decide the best method. If accuracy is much more emphasized, Method 1 will be a wise selection.

Inspectors

Subject was found to be a very significant factor in inspection. Table 37 lists the mean times and error rates by subjects. The slowest subject spent 139% longer time than the quickest subject ((75.86 - 31.70)/31.70 = 139%). The least accurate subject made 258% more errors than the most accurate subject ((2.04 - 0.57)/0.57 = 258%). These figures show the importance of selecting inspectors-- a project beyond the scope of this study.

Future Research

This study showed that the number of defects inspected at a time, defect rate, and defect difficulty have notable effects in the multi-defect inspection. Possible research could be done by varying the ranges of the three factors and observing the possible changes of the optimal inspection methods. Also, more research is needed to explore the usefulness of Method 3, inspecting for one group of similar defects at a time.

TABLE 37

Mean Inspection Times and Error Rates of Subjects

Subject	Mean Time (min.)	Mean Error Rate (%)
1	31.70	1.90
2	33.54	2.04
3	52.28	1.21
4	68.98	0.57
5	75.86	1.33
6	51.61	0.81
Mean	52.33	1.31

CONCLUSIONS

Inspection time and defect discriminability were highly correlated. The less discriminable defect needed a longer inspection time. Inspection error rate was not consistent with time. D and E were the least discriminable defects; A and B were in the middle; C was the most discriminable defect.

In the mixed level, inspecting for one defect at a time yielded 50% longer time and no less errors than inspecting for all defects at a time or inspecting for one group of similar defects at a time. There were no significant differences between the latter two methods. In the easy level, inspecting for one defect at a time yielded 88% longer time and no less errors than inspecting for all defects at a time. In the difficult level, inspecting for one defect at a time required 22% longer time but gave 17% less errors than inspecting for all defects at a time. Consequently, inspecting for multiple defects at a time performed better than inspecting for one defect at a time in the mixed or easy level while a cost analysis on time and errors is needed to decide the best method in the difficult level.

The number of defects inspected at a time, defect rate, and defect difficulty were found to affect inspection accuracy in the multi-defect inspection. Possible research could be done by varying the ranges of the three factors and observing the possible changes of the optimal inspection methods. Also, more research is needed to explore the usefulness of Method 3, inspecting for one group of similar defects at a time.

REFERENCES

Bloomfield, J. R., Peripheral acuity with complex stimuli at two viewing distances. In Air to Ground Target Acquisition, AGARD Conference Proceedings No. 100. Brussels: AGARD, 1972, B6-1-B6-10.

Bloomfield, J. R., Marmurek, H. H., and Traub, B. G., Color and texture differences in embedded target visual search situations. In Proceedings of the 18th Annual Meeting of the Human Factors Society, 1974, 158-170.

Colquhoun, W. P., Vigilance and the inspection problem. Personnel Management, 1957, 39(341), 163-171.

Drury, C. G., The inspection of sheet materials--model and data. Human Factors, 1975, 17, 257-265.

Ebeling, K. A. and Orr, P. H., Simulation and evaluation of alternative grading methods in potato packaging operations. In Annual Proceedings of the North Dakota Academy of Science, 1978, 30(Part II), 44-50.

Fox, J. G. and Haslegrave, C. M., Industrial inspection efficiency and the probability of a defect occurring. Ergonomics, 1969, 12(5), 713-721.

Harris, D. H., Effect of equipment complexity on inspection performance. Journal of Applied Psychology, 1966, 50, 236-237.

Harris, D. H., Effect of defect rate on inspection accuracy. Journal of Applied Psychology, 1968, 52(5), 377-379.

Harris, D. H. and Chaney, F. B., Human factors in quality assurance. New York: Wiley, 1969.

Howarth, C. I. and Bloomfield, J. R., A rational equation for predicting search times in simple inspection tasks. Psychonomic Science, 1969, 17, 225-226.

Jenkins, H. M., The effect of signal-rate on performance in visual monitoring. American Journal of Psychology, 1958, 71, 647-661.

Konz, S. and Osman, K., Team efficiencies on a paced visual inspection task. Journal of Human Ergology, 1977, 6, 113-121.

Kristofferson, M. W., Types and frequency of errors in visual search. Perception and Psychophysics, 1972, 11, 325-328.

McFarling, L. H. and Heimstra, N. W., Pacing, product complexity, and task perception in simulated inspection. Human Factors, 1975, 17, 361-367.

Megaw, E. D., Factors affecting visual inspection accuracy. Applied Ergonomics, 1979, 10.1, 27-32.

Miller, J. W. and Ludvigh, E., Time required for detection of stationary and moving objects as a function of size in homogeneous and partially structured visual fields. In A. Morris and E. P. Horne (Ed.) Visual Search Techniques. Washington: National Academy of Sciences, 1960, 151-169.

Neisser, U., Novick, R., and Lazar, R., Searching for ten targets simultaneously. Perceptual and Motor Skills, 1963, 7, 955-961.

Williams, L. G., Studies of extrafoveal discrimination and detection. In Visual Search. Washington: National Academy of Sciences, 1973, 77-92.

Williges, R. C. and Streeter, H., Display characteristics in inspection tasks. Journal of Applied Psychology, 1971, 55(2), 123-125.

MULTI-DEFECT INSPECTION

by

JINN-YEN SU

B.S. (Industrial), Tunghai University
Taiwan, Republic of China, 1976

AN ABSTRACT OF A MASTER'S THESIS

submitted in partial fulfillment of the
requirements for the degree

MASTER OF SCIENCE

Department of Industrial Engineering

KANSAS STATE UNIVERSITY
Manhattan, Kansas

1981

ABSTRACT

This experiment tested three multi-defect inspection methods under three difficulty levels of defects. Subjects were six non-student females. A set of 360 matrices was used as inspected items. The single defect rate for each defect was 6.94%. An item could contain more than one but at most two types of defects. The inspection was self-paced. Inspection time and errors were used as performance criteria. The less discriminable defect needed a longer inspection time. The number of defects inspected at a time, defect rate, and defect difficulty were found to affect inspection accuracy. In the mixed level, inspecting for one defect at a time yielded 50% longer time and no less errors than inspecting for all defects at a time or inspecting for one group of similar defects at a time. There were no significant differences between the latter two methods. In the easy level, inspecting for one defect at a time yielded 88% longer time and no less errors than inspecting for all defects at a time. In the difficult level, inspecting for one defect at a time required 22% longer time but gave 17% less errors than inspecting for all defects at a time. Consequently, inspecting for multiple defects at a time performed better than inspecting for one defect at a time in the mixed or easy level while a cost analysis on time and errors is needed to decide the best method in the difficult level.