NBSIR 80-2088

The Assessment of Safety Symbol Understandability by Different Testing Methods

NBSIR 80-2088

THE ASSESSMENT OF SAFETY SYMBOL UNDERSTANDABILITY BY DIFFERENT TESTING METHODS

Neil D. Lerner
Belinda L. Collins

Environmental Design Research Division
Center for Building Technology
National Engineering Laboratory
National Bureau of Standards
U.S. Department of Commerce
Washington, D.C. 20234

August 1980

U.S. DEPARTMENT OF COMMERCE, Philip M. Klutznick, *Secretary*

Luther H. Hodges, Jr., *Deputy Secretary*
Jordan J. Baruch, *Assistant Secretary for Productivity, Technology, and Innovation*

NATIONAL BUREAU OF STANDARDS, Ernest Ambler, *Director*

Abstract

This paper reports an experiment on the understandability of pictorial symbols proposed for fire-safety alerting. The purpose of the experiment was two-fold: 1) to determine the understandability of each symbol to a group of potential users; and 2) to assess the effects of variations in both presentation and response methods upon the measurement of understandability.

Twenty-five symbols, of which twenty-two had been proposed by the International Organization for Standardization (ISO) were evaluated for understandability with 91 U.S. subjects. These symbols were presented in three different ways: slides, placards, or booklets. Subjects indicated their understanding of each symbol's meaning by one of two response methods: by writing a brief definition or by selecting the correct definition from among four alternatives. In addition, for both methods, subjects rated their confidence in the correctness of their answers. In a second phase of the experiment, all subjects were given 15 different messages and asked to draw a symbol that would convey each message.

The results revealed no effect of the mode of symbol presentation. The definition and multiple choice response procedures led to generally similar conclusions. The confidence ratings provided useful additional information and helped to reconcile some of the discrepancies between the two response methods. Advantages and disadvantages of both methods are discussed.

The understandability of the 25 symbols ranged from near zero to virtually complete comprehension. The poor performance of some critical symbols such as "exit" was noted, and some potentially dangerous confusions in meaning were revealed. The drawings produced by subjects for given referents (or symbol meanings) were most likely to be different in image from ISO proposed symbols for those symbols which were poorly understood. These data underscore the importance of determining the understandability of safety symbols before symbol standards are adopted.

Table of Contents

Table of Contents (continued)

List of Figures

List of Tables

SI Conversion

The units and conversion factors given in this table are in agreement with the International System of Units or SI system (Systeme International d'Unites). Because the United States is a signatory to the 11th General Conference on Weights and Measures which defined and gave official status to the SI system, the following conversion factors are given.

Length

1 inch = 0.0254* meter
1 foot = 0.3048* meter

Area

1 square inch = $6.4516^{*} \times 10^{-4}$ $meter^2$
1 square foot = 0.0929 $meter^2$

Volume

1 cubic foot (ft^3) = 0.0283 $meter^3$

---

* Exactly

1.0 INTRODUCTION

1.1 SYMBOLIC SIGNS

The modern use of symbol signs began in 1909 with an international conference on highway sign systems in Europe (Eliot, 1960). This effort expanded rapidly to include most highway sign applications. Because symbols can be reduced to a small size, they are also extensively used for machinery and automotive applications, particularly for international trade. Now there is a growing international movement toward the use of safety and informational symbols within buildings. Pictorial symbols are often substituted for written signs because they require no knowledge of a particular written language and because they can convey some kinds of information more effectively than words.

Among the major advantages of symbols are that they can, in some cases, be perceived more rapidly (Janda & Volk, 1934), more accurately (Walker, Nicolay, & Stearns, 1965), and at a greater distance (Dewar & Ells, 1974) than words. Reaction time to symbols may be shorter (Ells & Dewar, 1979), even under conditions of stress (Smillie, 1978). Symbol meanings can often be rapidly learned and accurately remembered (Walker et al., 1965), with minimal confusion among alternatives (Green & Pew, 1978). Symbols may also be superior to words under conditions of interference either by distraction from another task (King, 1975) or by visual interference or degradation (Dewar, Ells, & Mundy, 1976). These advantages of symbols over words may not be true under all conditions for all symbols, however.

The perceived advantages to symbol use have led to more extensive production and use of pictorial signs. Unfortunately, this has resulted in a proliferation of confusing and contradictory symbols which often fail to communicate the desired information. The proliferation of different symbols for a common meaning and the use of confusing or contradictory symbols can cause serious problems for safety communication (Collins & Pierman, 1979). Yet, attempts at standardization may ultimately result in symbols which do not communicate any consistent meaning or which do not convey the correct meaning.

For example, the International Organization for Standardization (ISO) Technical Committee (TC) 21 recently proposed a set of symbols for fire safety and fire fighting information (1978). These symbols included such essential concepts as "exit", "no exit", "fire alarm call point", "do not block", and similar ideas. Although the symbols were intended for international standardization, they had never been evaluated for their understandability to any audience.

In an initial assessment of the understandability of 20 of the ISO symbols, Collins and Pierman (1979) reported that 9 symbols were understood correctly by less than 30 percent of their 143 subjects. Symbols which performed poorly included "exit", "no exit", "do not block", while symbols which were understood by more than 90 percent of the subjects included "telephone", "no smoking", "fire extinguisher".

These data indicated a wide range in the understandability of specific fire safety symbols. In response to these problems, the ISO committee proposed replacement symbols for the "exit" and "no exit" concepts.

The present paper reports an experiment on the understandability of the new egress symbols, together with other exit designations and the set of ISO-proposed fire-safety symbols. The purpose of the experiment was two-fold: 1) to determine the understandability of each proposed symbol to a group of potential users, and 2) to assess the influence of certain methodological variables in measuring symbol comprehension.

1.2 EVALUATION METHODS--BACKGROUND

Because a symbol must communicate the intended meaning to the target audience, understandability is one of the most important criteria for determining the effectiveness of a symbol. Understandability or meaningfulness has been assessed in many different ways. Brainard, Campbell, and Elkin (1961), Cahill (1975, 1976), King (1971), Walker, Nicolay, and Stearns (1965), Easterby and Zwaga (1976), and Easterby and Hakiel (1977) had subjects give short definitions for each member of a set of symbols. The definitions were then categorized as correct, incorrect, or occasionally, partially correct by one to three raters. The wrong answers given in the definition procedure provided insight into the kinds of confusions and misunderstandings associated with a particular symbol. In a somewhat different approach, other researchers have asked subjects to select a definition for each symbol from a long list of definitions. Brainard et al. (1961), Griffith and Actkinson (1977, 1978), Wiegand and Glumm (1979), and Freedman, Berkowitz, and Gallagher (1976) all used some variant of this selection procedure. Brainard et al. (1961) went one step further and compared the answers from the matching procedure with those for the definition procedure. As might be expected, there were fewer correct answers for the definitions, although there was a high correlation between the ordering of the two sets of responses in terms of symbol understandability.

Another technique is that of rated meaningfulness, in which a subject indicates numerically how meaningful a symbol is. When Dewar and Ells (1977) assessed the meaningfulness of a set of highway symbols, they found that rated meaningfulness was highly correlated with the accuracy of definitions of the same symbols. Green and Pew (1978) had subjects make magnitude estimations of the meaningfulness of automotive symbols. Again, symbols that were rated as poor also tended to be misidentified and confused.

Still another approach to determining understandability is that of rank ordering, in which a number of symbols for a given referent are ordered according to some criterion such as meaningfulness. Easterby and Zwaga (1976) and Easterby and Hakiel (1977) had subjects rank order at least six symbols for a given meaning in order to determine the "best" symbol for that meaning. This procedure can successfully identify several highly ranked symbols for later testing of understandability.

A number of investigators have used a behavioral measure to assess understandability. Heard (1974), for example, asked each subject to touch the appropriate automotive control symbol as meanings for the various symbols were read aloud. Several variations of each symbol were tested and compared by this procedure. Forbes, Gervais, and Allen (1963) and Dewar and Swanson (1972) observed driving behavior in response to symbolic instructions to determine the best symbol for a specific traffic control application. Freedman and Berkowitz (1977) measured the speed and accuracy of subject's way-finding in a "rally" using public information symbols.

In a different approach, Green (1979) and Brainard et al. (1961) gave subjects the symbol meaning and asked them to produce an appropriate drawing. Known as the "production method," this procedure provides insight into the common images that may exist for a particular idea.

In the investigations of symbol meaningfulness, researchers have used a variety of stimulus materials. Thus, Brainard et al. (1961), and Walker et al. (1965) used placard-type symbols. Griffith and Atkinson (1978), Dewar and Ells (1977), and Cahill (1975, 1976) used colored slides to present their stimuli. Still other researchers - Wiegand and Glumn (1979, Green and Pew (1978) Green (1979), Easterby and Zwaga (1976), Easterby and Hakiel (1977), and Freedman et al. (1976) - used booklets or reproductions which allowed subjects to work at their own pace. Finally, Forbes et al. (1963), Heard (1974), and Dewar and Swanson (1972) assessed symbols in an actual application. In no case, however, did any researcher compare the symbol's performance in the various stimulus presentation modes.

1.3 EXPERIMENTAL APPROACH

In the following experiment, three presentation modes and two response modes were assessed to compare their relative effectiveness. Because researchers may find it desirable to use slides, booklets, or placards depending upon the number and availability of subjects and the nature of the experiment, subject performance was compared for each of the stimulus modes. In addition, the problems inherent in the definition procedure - problems of administration, subjectivity, reliability, vagueness - made it desirable to compare it with the more objectively scored multiple choice procedure. In the response comparison, subjects either provided a short definition or selected an answer from among four multiple-choice answers. To provide an index of guessing and additional information on understandability, confidence ratings of the correctness of the answers were also obtained.

2.0 METHOD

2.1 EXPERIMENTAL DESIGN

The experiment consisted of a main portion on symbol meaningfulness and a supplementary one on symbol production. The meaningfulness portion evaluated how well a set of fire-safety symbols conveyed the intended messages. The production portion determined the kinds of images produced by subjects for a given message.

The procedure of the symbol meaningfulness portion defined a two-factor, independent groups experimental design. One factor was the mode of symbol presentation: placards, slide projections, or booklets. The other factor was the type of response required of the subject: providing a short definition of the meaning of each symbol, or selecting the correct meaning from among four multiple choice alternatives. Thus the understandability portion of the experiment was a 3 x 2 design (3 modes of stimulus presentation, 2 types of responses), with six independent groups of subjects.

2.2 SUBJECTS

A total of 94 paid participants, recruited from the Montgomery County, Maryland, area, were employed in the experiment. Subjects were screened to exclude those reporting any visual problems (other than corrective glasses). An attempt was made to recruit both male and female subjects over as wide an age range (from 18 years up) as possible. Three participants were discarded, two for failure to follow instructions, and another because of professional interests in symbols. The final sample of 91 subjects included 58 women (range 18-72 years, median age = 41) and 33 men (range 18-63 years, median age = 25).

2.3 SYMBOLS

Twenty-five fire safety signs were investigated. These included twenty symbols proposed by ISO TC 21 (1978), two proposed ISO replacement symbols for "exit" and "no exit", two U.S. "EXIT" signs (in red and in green), and a privately copyrighted "fire exit" symbol selected for use at the 1980 Winter Olympics. The symbols and their intended meanings are presented in Figure 1.

2.4 STIMULUS MATERIAL

For the understandability portion, each symbol was initially drawn on a 30 cm x 30 cm (1' x 1') placard. These placards presented the symbols to subjects in the placard condition. Color slides photographed from the placards were used to project the symbols (projected image size 90 cm x 90 cm, or 3' x 3') to subjects in the slide condition. Finally, 11 cm x 11 cm (4.25" x 4.25") photographs of the placards were photocopied in color for use in the booklet condition.

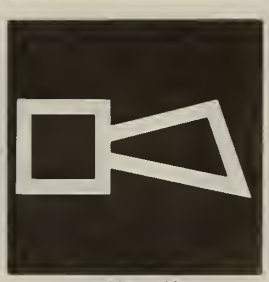
Figure 1. Twenty-five symbols intended for fire-safety alerting.

Do Not Block
(Black on White,
Red Circle & Slash)

Keep Fire Door
Shut
(White on Blue)

Emergency Phone
(White on Red)

General Phone
(White on Blue)



Fire Alarm Horn
(White on Red)

Fire Alarm
Call Point
(White on Red)

Fire Exit
(Copyright Yannone, 1979)
(Black on White,
Red Flame)

Emergency Exit,
Originally Proposed
(White on Green)

Emergency Exit,
Currently
Proposed
(White on Green)

Not an Exit
Currently
Proposed
(White on Black,
Red Circle & Slash)

Not an Exit,
Originally Proposed
(Black on Yellow)

U.S. Exit
Two Versions Tested:
Green on White
Red on White

Figure 1. Twenty-five Symbols Intended for Fire-Safety Alerting

In the production portion, referents were listed on otherwise blank sheets of paper (2 referents per sheet). Subjects were given black felt-tip marker pens to draw the symbols.

2.5 PROCEDURE

An experimental session consisted of the following sequence of events: a) participants read and signed consent forms outlining the general purpose and procedure of the experiment; b) instructions for the symbol meaningfulness portion of the experiment were given, and a practice example was worked and discussed; c) the symbol meaningfulness portion of the experiment was administered; d) instructions for the symbol production procedure were presented, and the production portion of the experiment was completed; e) subjects were debriefed. The entire session required 45 to 60 minutes.

Subjects were tested in groups of from 7 to 18 people, in conference rooms at the National Bureau of Standards. All of the subjects within a group received the same mode of stimulus presentation. However, a random half of each group was given the multiple choice response procedure, while the other half was given the definition procedure. The 25 symbols were presented in a different random order for each group. Two or three separate groups of subjects composed each stimulus presentation condition. The total number of subjects tested under each experimental condition is listed as part of Table 1.

One symbol at a time was presented to the subjects. For the placard and the slide conditions, subjects faced the front of the room, where the experimenter presented the symbols. Symbols were initially presented at a slow rate (of one every two minutes as specified in the instructions to the subjects). This permitted the subjects to become familiar with the procedure and allowed the experimenter to monitor their behavior. Because subjects required much less time to complete their response with experience, the rate of presentation was then increased so that a new symbol was presented whenever everyone had completed responding to the previous symbol. This usually took about 30 s per symbol. In the booklet condition, each participant had his or her own booklet. Each page contained a symbol and response form. Booklet subjects worked at their own pace.

Subjects in the definition condition were asked to "give a short definition of each fire-safety symbol that we present to you." Subjects in the multiple choice condition were asked to "circle the answer that is the best definition of the meaning of each fire safety symbol that we present to you." The incorrect alternatives were primarily derived from incorrect definitions given to a similar set of symbols used by Collins and Pierman (1979). All subjects were also asked to rate their confidence in the correctness of each answer on a scale from 1 (very uncertain) to 5 (very certain). Definition subjects rated definitions they provided; multiple choice subjects rated all the four alternatives for each symbol. A sample page from a booklet for the

Table 1 Mean Number of Correct Answers and Number of Participants for Each Experimental Group

	Slide	Placard	Booklet	
Multiple Choice	$\bar{X}$ = 17.1 n = 13	14.9 18	16.1 15	n = 46
Definition	$\bar{X}$ = 14.6 (11.9)* n = 14	15.1 (12.4) 16	15.8 (12.7) 15	n = 45
	n = 27	n = 34	n = 30	n = 91

* Numbers in parentheses refer to strict scoring

multiple choice group is shown in Figure 2 to illustrate the answer forms. Response forms for subjects in the slide and placard conditions were similar, except that the picture of the symbol was omitted. Complete instructions to the subjects are presented in Appendix A. The incorrect alternatives given for each symbol in the multiple choice procedure are listed in the second column of Appendix B.

After the symbol meaningfulness portion of the experiment had been completed, the supplementary symbol production procedure was introduced. Subjects were given booklets containing fifteen definitions and were asked to "draw a symbol that conveys the meaning of the definition." These definitions, which are listed in Appendix C, included 11 referents for fire safety messages (of which examples had been seen in the understandability portion of the experiment) and 4 referents for workplace safety ideas. Although no information about the correct answers was given in the first portion of the experiment, it is conceivable that viewing the symbols in the earlier portion could influence the subsequent production data. Therefore, the production portion of the experiment should be viewed as supplementary. After completion of this portion of the experiment, subjects were debriefed as to the purpose of the experiment and the intended meaning of each of the symbols they had seen.

2.6 SCORING OF RESPONSE PROTOCOLS

Answers for the subjects in the definition groups were scored as correct, partially correct, incorrect, or no answer. Three judges rated each answer. The judges initially discussed what constituted correct and partially correct answers for each symbol, and then independently scored all items. Where there was not initial unanimous agreement, they then attempted to resolve any discrepancy. Complete consensus was reached on all but 5 of the 1125 items (0.4%); these items were simply assigned the rating of the majority of the judges. Performance of the subjects in providing definitions will be reported for two scoring criteria - "strict" and "lenient." "Strict" scoring refers to "correct" answers only; "lenient" scoring refers to "correct" plus "partially correct" answers. An answer would be considered only partially correct if it: was too general to indicate that the specific function was understood (e.g., "door" for an exit, or "ladder" for "firefighting ladder"); was too narrow to indicate the intended range of the symbol meaning (e.g., "no striking of matches permitted" for "no open flame"); was too literal (e.g., "broken glass" for "break glass for access"); contained reference to the object symbolized although the full interpretation was not correct ("do not discard lighted match" for "no open flame"); or was otherwise ambiguous.

a) Fire shelter	Certainly wrong	1	2	3	4	5	Certainly right
b) Hard hat area	Certainly wrong	1	2	3	4	5	Certainly right
c) Fire fighter's equipment	Certainly wrong	1	2	3	4	5	Certainly right
d) Tunnel	Certainly wrong	1	2	3	4	5	Certainly right

Figure 2. Sample Page Showing Symbol, Multiple Choice Alternatives and Confidence Ratings

Multiple choice answers were scored as correct or incorrect. The data are presented as percentage correct, without any transformations which attempt to account for "guessing." Uncertainty or guessing is identified independently by the confidence ratings for response alternatives.

3.0 RESULTS

The three modes of symbol presentation (slide, placard, booklet) and the two types of response (multiple choice and definition) yielded 6 experimental conditions. Table 1 lists the mean number of correct answers (out of 25) for subjects in each group, as well as the number of subjects per group. The mean number correct given for the definition group is for lenient scoring ("correct" plus "partially correct"), with the number in parentheses representing the mean for strict scoring ("correct" only).

3.1 SYMBOL PRESENTATION/RESPONSE MODE

These data were analyzed by a two-factor analysis of variance. Separate analyses were conducted using strict and lenient scoring of the definition answers. Neither analysis revealed a significant (α=0.05) effect of the mode of symbol presentation, nor of the interaction of this factor with the type of response. For lenient scoring, there was no significant effect of the type of response, with the mean number correct under each condition being very similar. For strict scoring, the number correct for the definition groups was significantly lower than for the multiple choice groups. Table 2 summarizes these analyses.

Strict and lenient scoring of the definitions yielded generally similar information. The rank orders of the symbols under the two scoring criteria were in good agreement (rho = 0.96, $p < .001$). Differences exceeded 3 ranks for only 3 of the 25 symbols, and exceeded 4 only for a single symbol ("Emergency Exit").

Agreement among the slide, placard, and booklet groups for each of the 25 symbols is presented in Figure 3. This figure plots the percentage of correct answers for each symbol for each presentation mode. Each data point is an average across type of response (multiple choice and leniently scored definition). The figure confirms the results of the analysis of variance in showing generally close agreement between groups and no systematic tendency for one group to differ. The only dramatic discrepancy is for the initially proposed ISO emergency exit symbol, where the range between the best and worst scoring groups was 30 percent. The range was 10 percent or less for 17 of the 25 symbols, and less than 15 percent for 21 symbols.

Table 3 (left side) lists the percentage of correct answers for each symbol, for both the multiple choice and definition groups (data are combined across method of presentation). A breakdown of incorrect answers for both groups is contained in Appendix B. Figure 4 presents the number correct for both procedures as scatterplots. Each data point

Table 2 Summary of Analyses of Variance for Mode of Presentation and Type of Response

	Factor	Sum of Squares	df	Mean Square	F	
	Type of Response	16.65	1	16.65	2.62	
Lenient	Mode of Presentation	17.12	2	8.56	1.35	
Scoring	Interaction	31.48	2	15.74	2.48	
	Within groups	539.52	85	6.35		
	Type of Response	297.77	1	297.77	46.48	$p < .001$
Strict	Mode of Presentation	13.05	2	6.53	1.02	
Scoring	Interaction	28.96	2	14.48	2.26	
	Within groups	544.55	85	6.41		

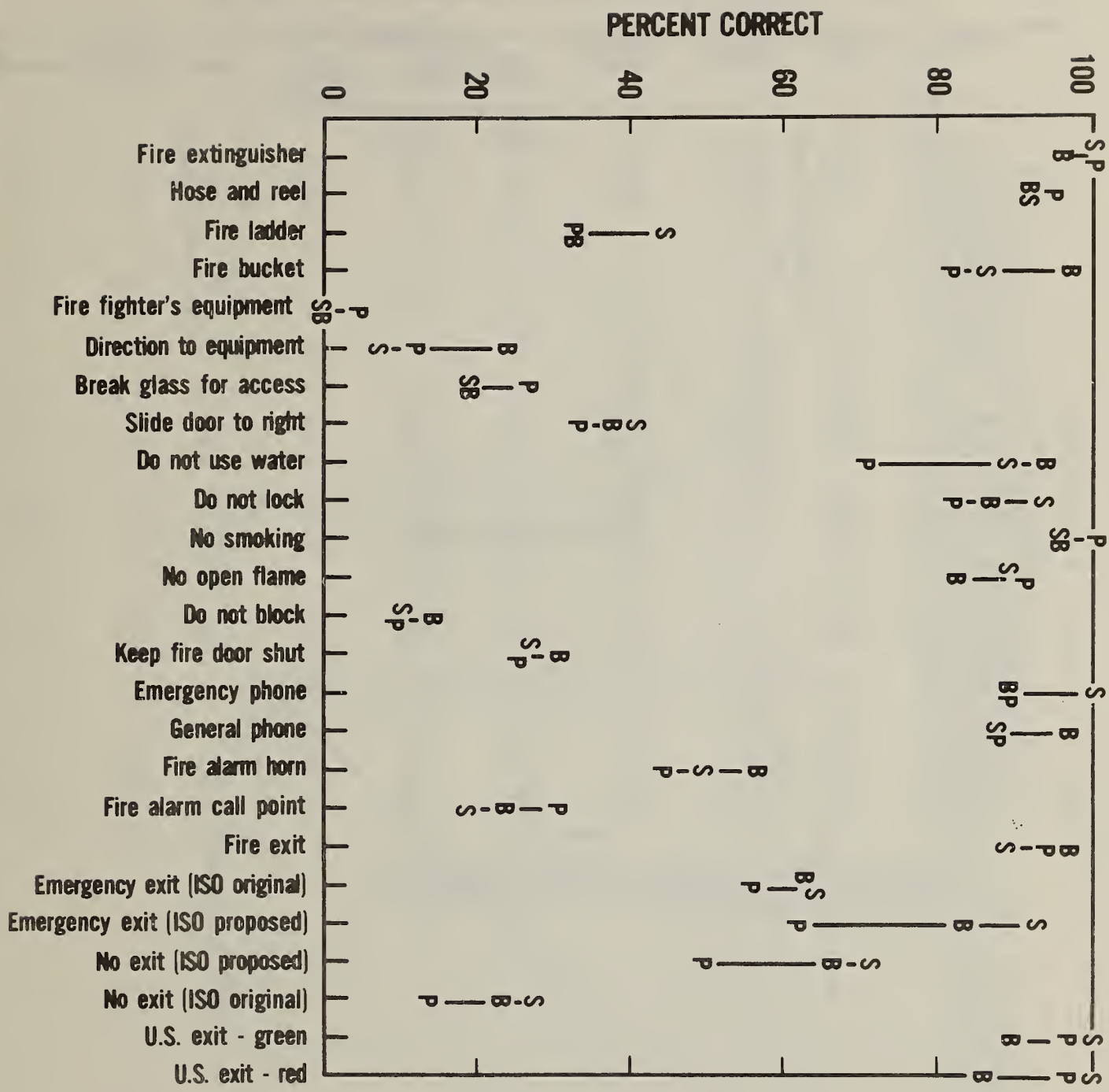


Figure 3. Percentage of Correct Answers for Each Mode of Presentation (slide, booklet, placard) for Each Symbol. S = slide; B = booklet; P = placard.

Table 3 Percentage of Correct Answers and Mean Confidence Ratings for Each Symbol

SYMBOL # MEANING	PERCENT CORRECT: MULTIPLE CHOICE	PERCENT CORRECT: DEFINITION (LENIENT)	PERCENT CORRECT: DEFINITION (STRICT)	CONFIDENCE RATINGS: MULTIPLE CHOICE: CORRECT ANSWERS ALL SUBJECTS	CONFIDENCE RATINGS: MULTIPLE CHOICE: CORRECT ANSWERS CORRECT SUBJECTS	CONFIDENCE RATINGS: MULTIPLE CHOICE: HIGHEST RATED ALTERNATIVE	CONFIDENCE RATINGS: DEFINITION: CORRECT	CONFIDENCE RATINGS: DEFINITION: PARTIAL	CONFIDENCE RATINGS: DEFINITION: INCORRECT
1. Fire Extinguisher	100.0	97.8	97.8	4.7	4.7	1.4	4.7	---	(2.0)
2. Hose & Reel	95.7	91.1	77.8	4.8	4.8	1.9	4.8	3.3	(2.0)
3. Fire Ladder	19.6	53.3	24.4	2.9	4.1	3.8	4.1	3.9	4.1
4. Fire Bucket	95.7	80.0	60.0	4.6	4.6	1.8	3.9	3.8	2.7
5. Fire Fighter's Equipment	0	2.2	0	1.5	---	4.0	---	(2.0)	1.8
6. Direction to Equipment	22.2	6.7	2.2	2.5	4.0	3.2	(4.0)	(4.0)	4.0
7. Break Glass For Access	28.3	15.6	11.1	2.8	4.2	3.0	2.8	(2.5)	3.0
8. Slide Doors To Right	56.5	15.6	4.4	3.3	4.1	2.8	(2.5)	2.8	2.6
9. Do Not Use Water	76.1	91.1	88.9	4.3	4.8	2.5	4.4	(1.0)	(3.0)
10. Do Not Lock	89.1	84.4	68.9	4.6	4.8	1.7	4.3	4.3	3.5
11. No Smoking	100.0	95.6	95.6	5.0	5.0	1.0	5.0	---	---
12. No Open Flame	87.0	88.9	75.6	4.1	4.5	1.8	4.2	4.5	(3.2)
13. Do Not Block	2.2	20.0	0	1.6	(5.0)	4.0	---	4.0	4.1
14. Keep Fire Door Shut	18.5	35.6	28.9	2.3	3.7	3.4	3.4	(3.3)	2.4
15. Emergency Phone	93.5	93.3	71.1	4.3	4.4	1.5	4.1	4.6	(3.0)
16. General Phone	87.0	95.6	82.2	4.6	4.8	1.6	4.3	4.0	(4.0)
17. Fire Alarm Horn	77.8	22.2	22.2	3.2	3.5	1.8	2.6	---	2.0
18. Fire Alarm Call Point	32.6	15.6	13.3	2.6	3.7	3.1	2.2	(2.0)	1.7
19. Fire Exit - (Yannone, 1979)	95.7	91.1	86.7	4.7	4.8	1.8	4.6	(3.5)	(3.7)
20. Emergency Exit ISO Original	57.8	62.2	22.0	2.5	3.2	2.2	2.9	2.4	2.4
21. Emergency Exit ISO Proposed	69.6	86.7	68.9	3.8	4.5	2.6	4.3	3.7	2
22. No Exit ISO Proposed	69.6	53.3	33.3	4.0	4.5	2.1	4.1	4.4	3.6
23. No Exit ISO Original	30.4	11.1	6.7	2.0	3.5	2.8	(3.3)	(4.0)	2.4
24. U.S. EXIT-Green	93.5	97.8	93.3	4.7	4.8	1.3	4.7	(5.0)	(4.5)
25. U.S. EXIT-Red	91.3	97.8	97.8	4.8	5.0	1.6	4.9	---	(4.0)
			Mean* =	3.6	4.3	2.4	4.0	3.8	2.9

* Calculation of these means excluded entries for symbols that were based on fewer than 5 observations (indicated by parentheses). A dash indicates no observations.

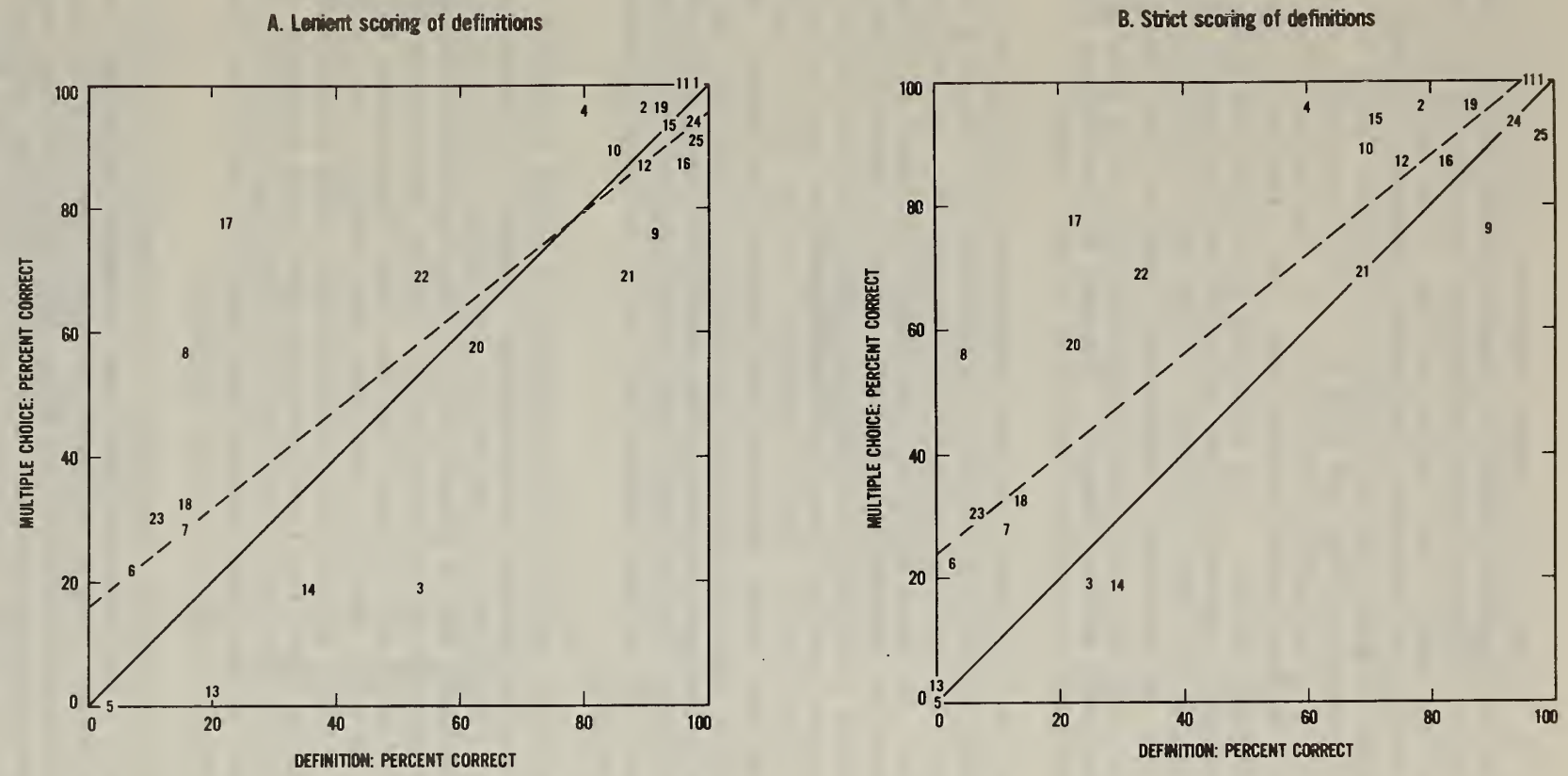


Figure 4. Comparison of Percentage Correct Obtained by Multiple Choice Procedure With Percentage Correct From Definition Procedure, for Both Lenient and Strict Scoring. The Numbers Correspond to Symbols in Table 3.

is represented by a number; the number indicates the particular symbol, as listed on the left side of Table 3. Panel A plots multiple choice performance versus leniently scored definition performance for each item; Panel B is similar, except that strict scoring of the definition answers is used. As the figure indicates, the general agreement between the response measures is good, the correlation coefficient being 0.86 ($p < .001$) for lenient scoring and 0.87 ($p < .001$) for strict scoring. However, there are a few marked deviations. The regression lines fit to the data points (broken lines) deviate from a 45° line (solid lines) at the lower percentage values. This indicates that while agreement between response procedures is generally good on relatively successful symbols, the multiple choice procedure typically yields higher estimates of the understandability of relatively unsuccessful symbols. (Of course, this is partly due to the fact that random guessing will yield about 25% correct for the multiple choice groups). Features of the deviant points will be further described in subsequent sections of this report.

Table 3 (right side) contains the mean confidence ratings given for each symbol, collapsed across the method of presentation. Because a few subjects ignored or incorrectly followed instructions for the confidence ratings, at least for some symbols, these means are based on slightly different numbers of observations for each symbol. Table 3 presents confidence ratings separately for multiple choice and definition groups. The multiple choice portion lists mean confidence ratings for a) the correct choice alternative, for all subjects, b) the correct choice alternative, for only those subjects that selected the correct answer, and c) the highest rated of the three incorrect alternatives, for all subjects. The definition portion of the table lists mean confidence ratings for a) subjects providing a correct definition, b) subjects providing a partially correct definition, and c) subjects providing an incorrect definition.

As Table 3 indicates, subjects who identified a symbol correctly, whether in the definition or multiple choice groups, tended to be more confident in their answer than those who answered incorrectly. Figure 5 shows a plot for the multiple choice groups, of the mean confidence rating against the percentage of correct answers. Again, the numbers on the scatter plot indicate which symbol the data point corresponds to, as indicated in Table 3. Panel A plots confidence ratings for all subjects; Panel B gives ratings for correct subjects only. Both panels show that ratings tend to be lower for items that are more often answered incorrectly. Relatively successful items (high percentage correct) tend to have confidence ratings that cluster from about 4.5 to 5.0 for correct subjects and about 4.0 to 5.0 for all subjects.

The symbols with the lowest percentage correct generally had mean ratings of only 3.5 to 4.2 for those subjects that answered correctly; for all subjects, the poorest symbols had mean ratings of from 1.5 to less than 3.0. The correlation coefficient of group mean rating of the correct answer with the percentage correct is $r = 0.95$. For most of the 25 symbols, the correct multiple choice alternative was given the highest group

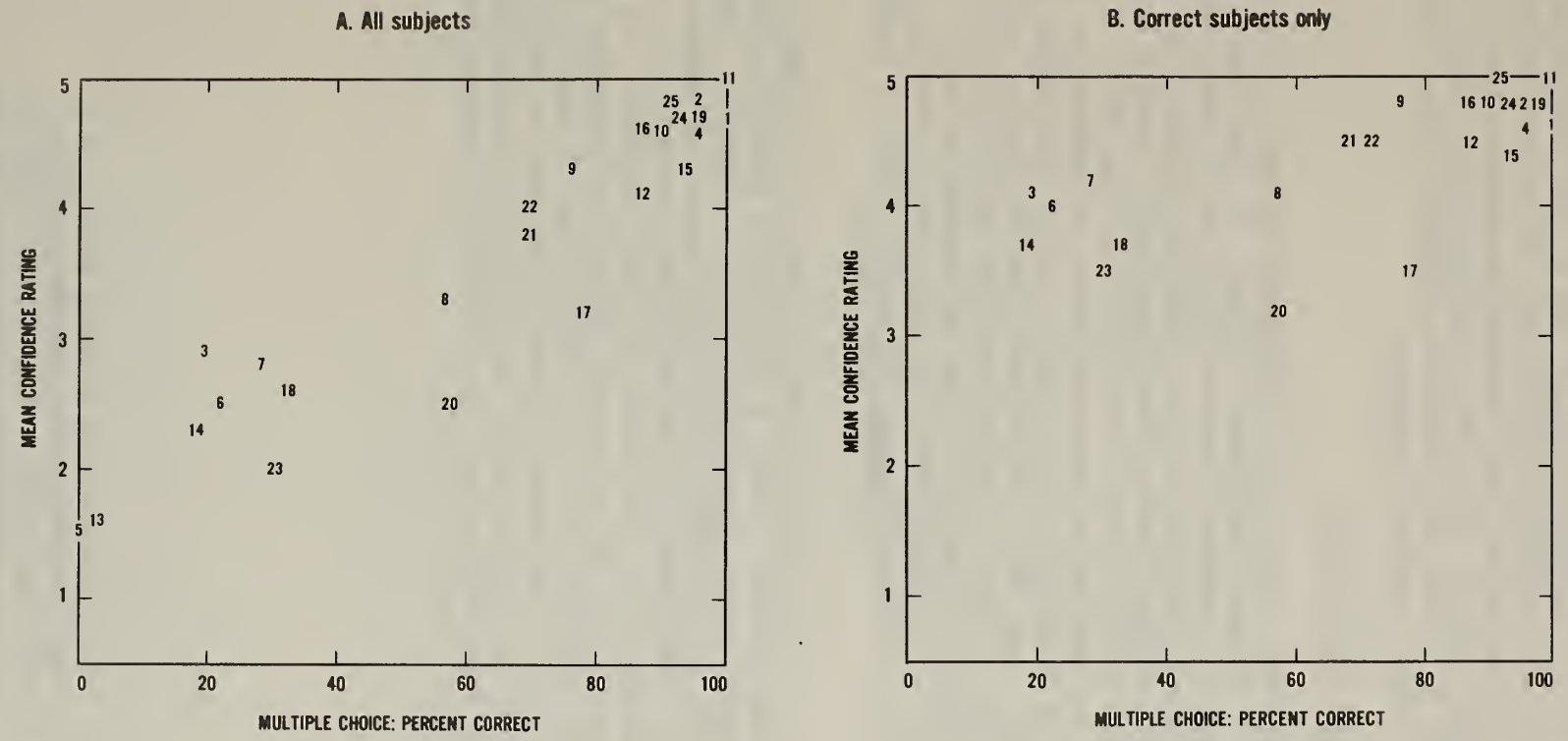


Figure 5. Multiple Choice Confidence Ratings as a Function of Percentage of Correct Answers for Multiple Choice. The Numbers Correspond to the Symbols in Table 3.

confidence rating, while the incorrect alternatives averaged much lower. However, in eight of the twenty-five cases, one of the incorrect alternatives was given a higher confidence rating than the correct alternative. These were the same alternatives that frequently appeared (usually more often than the correct answer) as incorrect answers in the definition groups (see Appendix B). Thus both response procedures identified similar confusions about a symbol's meaning.

3.2 PRECISION OF THE DATA POINTS

The precision of the data points can be estimated for both the percentage correct data and for the mean confidence ratings. In both cases, the relationship between the estimated standard error for a symbol and the percentage of correct answers for the symbol is described by an inverted-U function. Figure 6 shows these functions. Panel A shows the standard error of the multiple choice confidence ratings for the correct alternative. (The estimate of the standard error is obtained by dividing the observed standard deviation by the square root of the number of observations, in this case, 46). As indicated in the figure, there was better agreement among subjects' ratings when the symbols were very good (high percentage correct) or very poor (low percentage correct). The standard error of the mean ranges from less than 0.03 up to 0.22 rating scale units; the 95% confidence interval for a data point is roughly ± 2 standard errors.

The standard error for the percentage correct measure is estimated by using the binomial distribution. Panel B of Figure 6 plots the relationship using the normal curve approximation to the binomial. Based upon this relationship, the standard error of the number of items correct is estimated by $\sqrt{NPQ}$ (where N = number of observations; P = proportion correct answers; Q = proportion incorrect answers). To express the standard error in terms of percent rather than absolute number) correct, the estimated standard error is $\sqrt{NPQ}/N$. Panel B shows this function for an N of 46.[1/] The standard error is maximal at 50% correct, where it is about 7.4%; it reduces as the percentage correct deviates (in either direction) from 50%, so that for 98% (or 2%) correct, the estimated standard error is only about 2%.

1/ The approximation to the normal curve becomes poorer as P approaches 0 or 1 or as the sample size becomes smaller. Thus while Panel B shows the general shape of the function for larger groups, the extremes of the function may be somewhat inaccurate for an N of 46.

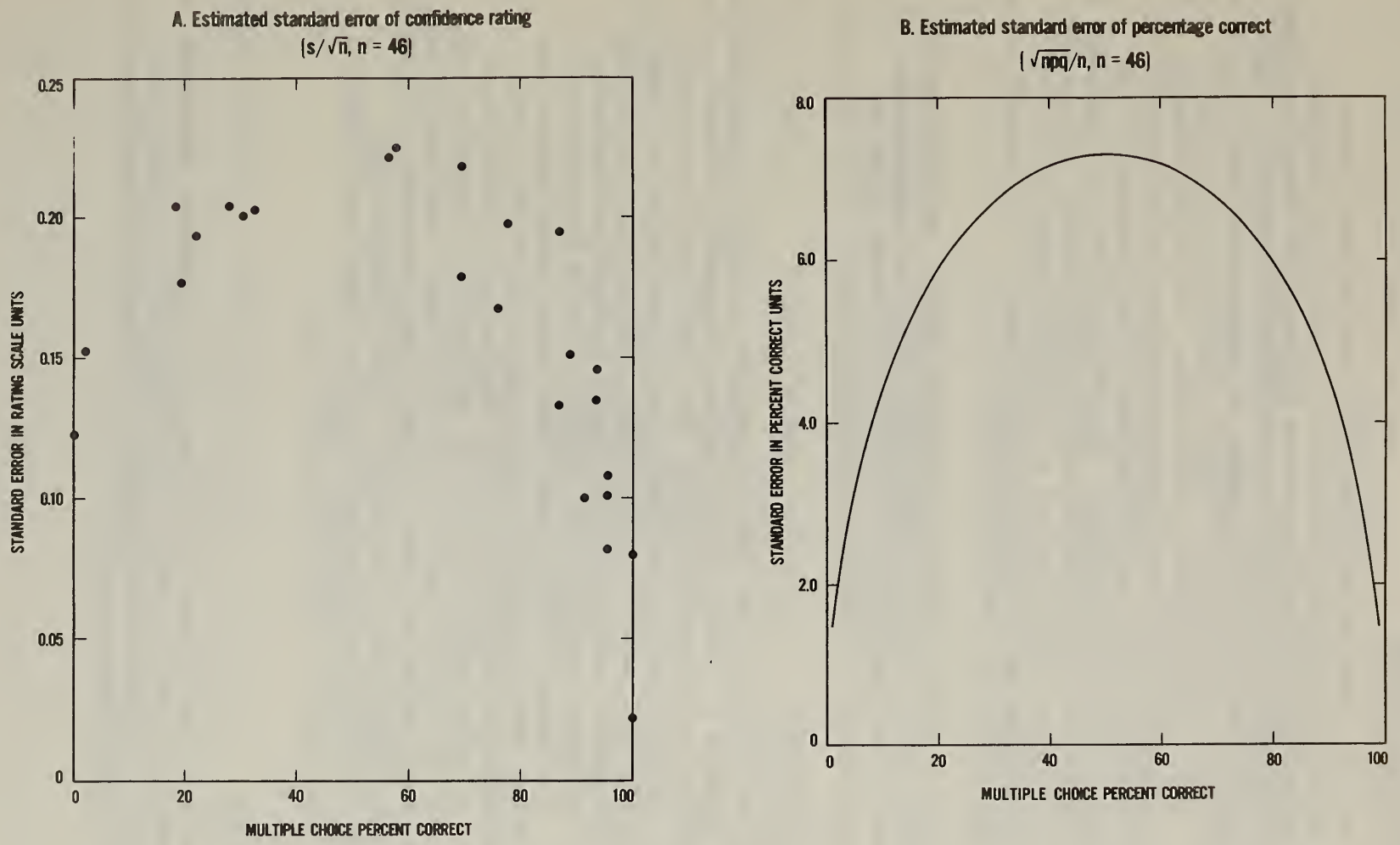


Figure 6. Estimated Standard Error as a Function of Percentage Correct Responses for Both Multiple Choice and Definition Procedures

3.3 EFFECTS OF AGE AND SEX

Performance was also examined as a function of age and sex. Since neither the method of stimulus presentation nor the type of response (using lenient scoring) had any discernible effect on the number of items correct, data were collapsed across the methodological variables. The correlation coefficient of age with number of items correct was $r = -0.30$, which, while significant ($p < .01$), accounted for only about 9 percent ($r^2 = .09$) of the variance in the scores. For an analysis of variance, age was broken into three categories (less than 25 years, 25 to 49, and 50 or older) and a two-factor (Age by Sex) analysis of the number of items correct was performed. The sex factor was not significant (overall means of 15.6 for males and 15.3 for females), nor was the age by sex interaction. While age was a significant factor ($F(2,85) = 4.92$, $p < 0.01$), the age categories could account for only about 8 percent of the variance in the data ($\omega^2 = 0.077$.

The mean number of correct items for the young, middle, and older age groups were 15.6, 16.4, and 14.3 respectively. Post hoc comparisons indicated that the only statistically reliable difference among age categories was between the middle and older age groups ($p < 0.01$).

3.4 SYMBOL PRODUCTION

A summary of the results of the symbol production procedure is given in Table 4 and Appendix C. Table 4 lists the most frequently produced image for each of the fifteen referents. For comparison, the proposed ISO image for each referent is included in the table. Appendix C summarizes the production results in more detail. For each of the fifteen referents, the appendix categorizes the images produced and indicates the number of subjects producing each image. To avoid idiosyncratic associations, the summary only includes those images given by more than one person (except where a commonly drawn image is described in finer detail). The summaries do not incorporate the factors of color, background shape, or added words, but focus on the pictorial content of the image. Although the table presents only symbol images, there were a few referents for which words were among the most common responses in spite of specific instructions not to use words. These referents (and responses) included Exit and Fire Exit ("Exit"), Hazard Area ("Danger"), and Blind Alley ("No Exit," "Dead End," "Stop").

4.0 DISCUSSION

This experiment had two primary objectives: 1) to determine the effect of certain methodological variables on the measurement of symbol meaningfulness, and 2) to provide information on the meaningfulness of a specific set of fire safety symbols.

Table 4 ISO Images and Most Frequently Produced Images for Each Referent

Referent	ISO Image	Production Image
Extinguisher	Canister with hose	Same
Fire Alarm Call Point	Concentric circles	Telephone receiver
No Smoking	Circle, slash, lighted cigarete	Same
No Open Flame	Circle, slash, lighted match	Flame rather than match
Do Not Use Water To Extinguish Flame	Circle, slash, bucket pouring water on flames	Same
Do Not Lock	Circle, slash, padlock	Same
Do Not Block	Circle, slash, box and can	Corridor with objects and slash
Keep Fire Door Closed	Door with arrows to close	Same
Blind Alley	Triangle, closed box, arrow (initially proposed) Circle, slash, open door, figure (currently proposed)	Corridor with arm obscured
Exit	Rectangle (initially proposed) Open door, figure (currently proposed)	Arrow
Fire Exit	Same as "Exit"	Flame, figure, door or arrow
Hazard	Triangle, exclamation point	Explosion-like object
Hearing Protection	Head with ear muffs	Same
Head Protection	Head with hard hat	Same
Eye Protection	Head with eyeglasses	Eyeglasses alone

4.1 COMPARISON OF METHODS

4.1.1 Mode of Symbol Presentation

Two methodological variables were investigated in the experiment: mode of symbol presentation and type of response. The means by which the symbols were presented--placards, slides, or booklets--had no discernible effect on how accurately subjects identified the symbol meaning. No effect on the number of correct answers was revealed in an analysis of variance, and Figure 3 showed generally close agreement between the three presentation methods. Although one symbol (currently proposed ISO "emergency exit") showed a range of 30 percent between the three conditions of symbol presentation, a single such observation among 25 triads is statistically probable due to the variability inherent among groups of this size (about 30). The close agreement among the responses to most of the symbols provides little basis for assuming any effect of mode of presentation. Thus for investigating the meaningfulness of symbols (though not necessarily for other attributes), it appears to make little practical difference how the stimuli are presented.

Each mode of presentation has certain advantages. Slides are useful for presentation to large groups. Placards present the symbols in a form which may most closely approximate real signs as seen in buildings. Booklets are efficient in that subjects are self-paced, independent, and do not require the presence of an experimenter (in fact, subjects could be reached by mail or through intermediary organizations). From a methodological viewpoint, booklets have the further advantage of allowing a different random order of symbols for each subject, thus minimizing the possibility of sequential effects in the group data. Because the present experiment suggested that presentation method does not differentially affect symbol meaningfulness, it appears that the most convenient method may be selected for a particular application.

4.1.2 Method of Response

The other methodological factor investigated in this experiment was the type of response required of the subject: providing a short definition or selecting the correct choice alternative. Additional information was obtained for both procedures by requiring the subjects to rate their confidence in the correctness of the answers.

Many different procedures have been employed to evaluate the meaningfulness of symbols; these methods have included giving definitions, matching symbols with referents, providing various sorts of ratings or rankings, drawing symbols (production), and providing behavioral measures. The relative merits of each evaluation method depend in part upon the stage of symbol development, which can range from initial conceptualization and design through large scale population testing and standardization. One point at which thorough large-scale testing is critical is after a set of symbols has been selected. One, or at best a few, symbolic renditions of each referent are presented to large groups of potential users, who are often sampled with regard to demographic variables. Typically, the

method used in these experiments (e.g., Easterby & Zwaga, 1976; Freedman & Berkowitz, 1977; Brainard et al., 1961) has been to require subjects to write down a definition for each symbol; ISO TC 145 has recommended this procedure. In the present experiment the multiple choice procedure was investigated as a possible alternative method to the definition procedure. While this experiment employed only 91 subjects, the methodological comparisons really were directed at appropriateness for programs of larger scale and expense.

Those favoring definition procedures cite the advantage of the "open ended" quality of the response. Techniques such as multiple choice potentially constrain answers and facilitate guessing (Easterby & Hakiel, 1977). However, in the evaluation of refined symbols ready for adoption, these concerns may not be primary. In addition, the definition procedure poses problems such as the difficulty of administering and scoring large numbers of definition tests as well as the subjectivity and reduced reliability (rater agreement) of scoring (Freedman, 1979). Among the frequently encountered interpretive problems are vague, partially correct, or partially incorrect answers, possible differences in terminology between scorers and test participants, and illegibility. It is often desirable to have answers scored by more than one rater to avoid idiosyncratic scoring and to provide an index of scorer reliability. This requires additional time, effort, and expense. When several raters are used, the problems become increasingly difficult, and for international testing, subtle semantic distinctions could confound attempts at consistency.

In fact, scoring problems did emerge as a significant concern with the definition procedure in the present experiment. Although the rank-order correlation was good (rho = 0.96), the mean difference in the percentage correct for lenient and strict scoring of individual symbols was about 11 percent, and the discrepancy ranged up to 40 percent. While agreement among the individual scorers was good, it was due in part to adopting quite strict criteria for "correct" scoring and quite lenient criteria for "partially correct" scoring. Thus the agreement was sometimes achieved at the cost of increasing the size of the ambiguous "partially correct" category; this category accounted for 20 percent or more of the cases for about one-fourth of the symbols. This indicates a need for caution in interpreting the results of experiments where only a single "percent correct" is given.

Because there are criticisms of the multiple choice procedure as well as of the definition method, this experiment employed the multiple choice method with two important additional features. These features were included to mitigate the major concerns with the multiple choice method (constraints on the variety of answers and ease of guessing). First, by using as alternative choices those incorrect definitions frequently provided by subjects in the Collins and Pierman (1979) experiment, severe constraints on answers could be avoided. That is, since the alternatives cover the general types of answers given by most of the subjects using a definition procedure, it is likely that one of the alternatives will be similar to that which the subject would provide in a definition proce-

dure. Secondly, ratings of the subject's confidence in the correctness of each alternative allow guessing to be identified, and have the further advantage of providing more than simple binary information (correct or incorrect) about a given answer. Thus it was hoped that an easily and objectively scored multiple choice procedure with confidence ratings could provide substantial information with minimal response constraint or bias. One purpose of the present experiment was to permit explicit comparison of the results from such a multiple choice procedure with the results obtained from a definition procedure for measuring symbol meaningfulness.

The correlation coefficients relating the two methods were fairly high, $r = 0.86$ and 0.87 for lenient and strict scoring, respectively. As Figure 4 indicates, the two methods agreed most closely for "good" symbols (high percentage correct). The multiple choice procedure yielded generally higher estimates of meaningfulness for poorly understood stimuli.

Despite the general agreement between the methods, several substantial deviations occurred, which are discussed more fully below. However, it is important to recognize that some of the differences between definition and multiple choice results could reflect in part random variation in the data, rather than true differences in the methods. Figure 6 presents the estimated standard errors for the data.

4.1.3 CONFIDENCE RATINGS

The confidence ratings provide information in addition to percentage correct that is valuable in evaluating symbol effectiveness (particularly for multiple choice data) and in comparing definition and multiple choice data. For example, in some cases a pair of symbols may be quite similar in terms of the percentage of correct answers, yet may differ considerably in how certainly the subjects believe the answers to be correct. For the multiple choice data, for instance, "Slide Door to the Right" and "Emergency Exit (initally proposed)" were both correctly chosen about 56-58 percent of the time. Nonetheless, confidence ratings given for "Emergency Exit" were quite low; even those subjects who correctly selected it gave a mean rating of only 3.2. The rating for "Slide Door" was higher than this even for all subjects together (3.3), and much higher (4.1) for those answering correctly. Similarly "Fire Alarm Horn" and "Do Not Use Water to Extinguish" were each correctly selected about 76-78 percent of the time. Still, subjects rated "Do Not Use Water" as a very reasonable answer (4.3 for all subjects) while "Fire Alarm Horn" was seen as much less likely (3.2 overall, 3.5 for those correct). Thus even where the percentage of correct answers is reasonably high, the confidence ratings may reveal guessing or uncertainty.

The confidence ratings can also provide information about serious confusions and reveal why some symbols do not perform as well as might be expected. For example, for multiple choice subjects, "Fire Ladder" was correctly answered much less frequently than "Not an Exit (initially proposed)" (19.6 vs 30.4%), yet the confidence ratings for "Fire Ladder"

were considerably higher (2.9 vs 2.0 for all subjects). Alternative answers to "Fire Ladder" received even higher ratings (3.8 for "Use ladder for fire escape"). Thus the ratings suggest that while the "Not an Exit" symbol did poorly because it was relatively meaningless to subjects, "Fire Ladder" did poorly because of confusion with other possible interpretations.

Two interesting observations come from the confidence ratings of the definition group. First, inspection of Table 3 indicates that even for participants that provided the correct answer, confidence in the answer was relatively low for symbols that had a lower percentage of correct answers. Thus "understandability" was reflected not only in the percentage correct, but also in the confidence ratings of those that in fact answered correctly. Secondly, all symbols that showed good understandability (> 80% correct) also had uniformly high confidence ratings from correct subjects (> 4.25). Thus in no case does it appear that a symbol was correctly identified by definition subjects largely on the basis of guessing. In general, however, the confidence ratings provide less essential information for the definition procedure than they do when used with the multiple choice procedure.

The confidence ratings may also help explain some of the discrepancies between multiple choice and definition groups on certain symbols. Certain data points in Figure 5 stand out by having relatively low confidence ratings, typical of poorly understood symbols (< 35% correct), even though there are a substantial number of correct answers (50-80%). These tend to be most of the symbols in which multiple choice subjects did substantially better than definition subjects (see Figure 4). Thus because multiple choice subjects have alternatives to guess at, the use of a multiple choice procedure may overestimate the interpretability of a symbol. Use of confidence ratings provides a check on this overestimation.

Two previous experiments compared various sorts of rating data with a measure of the understandability of symbols. Dewar and Ells (1977) first had subjects provide definitions for twenty highway symbols, and then rate each symbol using the semantic differential test. This test involves rating stimuli on various bipolar adjective scales (e.g., "good" - "bad," "weak" - "strong," etc.) to derive values for each of four factors: evaluative, potency, activity, and understandability. Dewar and Ells found all four factors correlated significantly with percentage recognition, with correlation coefficients ranging from -0.69 to -0.78. Green and Pew (1978) measured "association norms" for automotive control symbols by having subjects touch the appropriate symbol (from a set of 19) when the control's function was described in a driving scenario. After this, subjects were given each symbol with its correct definition and asked to rate the symbol's "communicativeness". The rating correlated with the "association norms" (% correct), $r = 0.73$. In both these experiments, the usefulness of ratings was discussed as a possible substitute for understandability measures, rather than as supplementary data. However, it is questionable whether the correlations with understandability measures were sufficiently high to permit treating

these measures as equivalent. Even for the confidence ratings of the present experiment, which correlated considerably more highly with understandability (r = 0.95), some serious deviations occurred. All of these rating procedures have somewhat different approaches, and are useful in different ways. It would not appear appropriate to view them as analogous to a measure of understandability, however.

4.2 SYMBOL PERFORMANCE

Since the mode of symbol presentation (placard, slide, booklet) had no statistically discernible effect, it will be ignored in discussing the performance of individual symbols. The four primary measures of symbol performance in this experiment--percentage correct for strictly scored definitions, percentage correct for leniently scored definitions, percentage correct for multiple choice, and mean confidence ratings--showed generally good agreement. Therefore, the overall relative performance of individual symbols will be discussed in detail. Disagreements among measures will be noted.

Regardless of the measure used, the twenty-five symbols differed widely in understandability. The percentage correct measures ranged essentially from 0 to 100 percent. Group mean confidence ratings for the correct alternative similarly ranged from 1.5 to 5.0 on a 5-point scale. The precision of the data should be kept in mind in this evaluation of symbol performance, however. While the total of 91 subjects can provide a good initial assessment, definitive statements require much more extensive sampling. It should also be noted that, although individuals of both sexes and a range of ages were employed, the volunteers in this experiment do not constitute a random sample. At the least, the less educated and less affluent were underrepresented. Full-scale evaluations of symbol performance require more extensive and representative sampling.

There is no obvious criterion for determining "acceptable" performance for a symbol, especially given the influence of scoring criteria. It would be unwarranted to assume that the percentage of correct responses found in a laboratory experiment corresponds to the actual absolute percent of the population that would understand the symbol under actual conditions of application. Nonetheless, there has been a certain degree of consensus among researchers on criteria for symbol acceptability. In three large scale evaluations (Brainard et al., 1961; Freedman & Berkowitz, 1977; Heard, 1974), the various researchers set levels of "acceptable" performance which ranged from 75 to 85 percent. As a result, in the present experiment, a criterion of 80 percent was chosen as an initial level to define "acceptable" meaningfulness. In addition, a level of 50 percent was chosen below which an "unacceptable" category was defined. While these labels should not be taken too literally or these criteria viewed as final, they do provide a rough initial index of acceptability. Table 5 presents the category of acceptability for each of the 25 symbols for the multiple choice and leniently scored definition groups. As the table indicates, the two methods were in good general agreement, classifying 20 of the 25 symbols in the same manner (for strict scoring, the decisions would have agreed on 15 symbols).

Table 5 Categories of Acceptability for Each Symbol for Both Multiple Choice and Definition (Lenient) Groups

Definition (Lenient) 50-80% Percentage Correct	Multiple Choice Percentage Correct: Less than 50%	50-80%	Greater than 80%
Less than 50%	Fire Fighter's Equipment Fire Alarm Call Point Do Not Block Direction to Equipment Keep Fire Door Shut Break Glass For Access No Exit - ISO Original	Fire Alarm Horn Slide Door to Right	
	Fire Ladder	Emergency Exit ISO Original No Exit- ISO Proposed	Do Not Use Water to Extinguish Fire
Greater than 80%		Exit-ISO Proposed	Fire Extinguisher Hose and Reel Fire Bucket General Telephone Emergency Telephone No Smoking No Open Flame Do Not Lock Fire Exit (Yannone) U.S. Exit (Red) U.S. Exit (Green)

Eleven symbols were classified as "acceptable" ($\geq$ 80%) by both methods; seven were "unacceptable" (< 50%) by both methods; the remaining seven symbols were "questionable" (50-80%) for at least one of the methods. As reference to Table 3 indicates, those instances where the two response methods led to different categorizations were usually where a) there was a large effect of the definition scoring criterion, or b) multiple choice subjects were able to guess (as indicated by low confidence ratings).

In addition to the overall categories of performance presented in Table 5, specific symbol performance can also be analyzed in more detail.

Perhaps the most important fire safety symbols are those relating to safe egress. Five exit-related symbols and two "no exit" symbols were included in this experiment. The two exit symbols which tested rather poorly included both ISO "emergency exit" symbols. The poor performance of the initially proposed ISO "emergency exit" symbol appears due to the symbol having little intrinsic meaning of any sort. This is suggested by the vagueness of the definitions given, the low confidence ratings given to all the multiple choice alternatives, and the low confidence ratings given even by correct definition and multiple choice subjects. Much of the information in this symbol was probably conveyed by the green background color, a cue which may be lost in heavy smoke or dim illumination. The currently-proposed ISO "emergency exit" symbol was more meaningful but suffered primarily by connoting information about running; this might be improved through graphic modification. On the other hand, the "fire exit" symbol (copyrighted by Yannone) tested quite well, and was comparable in meaningfulness to the red and green "EXIT" word signs. It should be noted however that the "fire exit" symbol included a directional arrow as an integral component of the symbol design. Although the other exit symbols may be used in conjunction with a directional arrow, they were not tested this way. Since the responses to the arrow alone and the results of the production method (see below) both indicated that an arrow suggests the idea of egress, the superior performance of the "fire exit" symbol may be partially due to the presence of the arrow.

The currently proposed ISO "not an exit" symbol conveyed meanings regarding running or pedestrians, or was taken to indicate a general "do not enter." However, the confidence ratings indicated that subjects at least found "not an exit" to be a reasonable interpretation. In contrast, the initially proposed ISO "not an exit" symbol received a low percentage of correct answers and low confidence ratings. Potentially dangerous misinterpretations regarding egress were common for this symbol: "exit," "shelter area," "elevator," "go this way" accounted for about 40 percent of all subjects ("overhead hazard" accounted for another 23%). Thus the initial ISO symbol not only failed to suggest its intended meaning to most subjects, but also suggested dangerous contradictory meanings. Similar contradictory meanings were also reported by Collins and Pierman (1979).

The findings for the arrow indicating the direction to fire fighting equipment must be qualified. Because the arrow was designed to be used in conjunction with a symbol indicating equipment, its low meaningfulness here is partly due to its being tested out of context and alone. However, it was tested in this fashion to determine whether it denotes strong meanings of its own, especially because it may be used with relatively meaningless symbols, such as the one for fire fighter's equipment. The arrow did in fact convey the idea of "exit," "one way," or "go (or don't go) this way" to 40 of the definition subjects and 35 of the multiple choice subjects (82% overall). Thus while the arrow may be more meaningful than tested when it is actually used with another sign, the data suggest that it is important to integrate the arrow clearly with another symbol such as equipment so that it does not convey the idea of egress.

Several other symbols in addition to "not an exit" and "direction to equipment" also suggested meanings that could lead to dangerous confusions. These include the symbols for "firefighting ladder" (interpreted as related to fire escape), "firefighting equipment" (interpreted as shelter or tunnel), and "slide door" (interpreted as indicating the direction of the exit). Appendix B details these incorrect answers. In all of these cases, misinterpretation could send an occupant seeking egress in a hazardous direction.

Twenty-two of the symbols investigated in the present experiment had previously been tested by Collins and Pierman (1979) using a definition procedure. The results of the two studies are in generally good agreement. The correlation coefficient for the percentage correct reported by Collins and Pierman with either the leniently scored or the strictly scored percentage correct in the present experiment was $r = 0.94$ in both cases. Despite the overall agreement, the percentage correct for a few symbols differed noticeably. The difference in one case ("slide door") was attributable to differences in scoring criteria. In the other cases, the difference was attributable primarily to the performance of a group of elderly people in the Collins and Pierman experiment. Although there was some effect of age on performance in the present experiment, the effect was small. However, the older participants in the present experiment were volunteers who came to the NBS site to participate. The Collins and Pierman experiment included a group of elderly residents tested at their retirement home; this group was somewhat older (mean age of 76 years) and presumably less active and alert than participants in the present experiment. In fact, answers from the retirement home residents frequently suggested confusion and disorientation. While residents of such homes are an important group to consider, their inclusion in the group data accounted for poorer performance on a number of symbols. Nevertheless, even with this difference, the replicability of the experimental findings remained quite good.

4.3 PRODUCTION DATA

Inspection of the production data images indicates that the image content does not necessarily follow that proposed by either ISO TC 21 ("Equipment for Fire Protection and Fire Fighting", 1978) or TC 80 ("Safety Colors and Signs", 1979). In Table 4, the referent, the ISO image, and the most frequently occurring production image are presented. Since symbols for the fire safety referents (but not the workplace safety referents) had been presented in the earlier portion of the experiment, this may have influenced the production data, which should therefore be viewed as supplementary. Presumably any bias would be in the direction of increasing the liklihood of obtaining the ISO image. As can be seen in Table 4, in seven cases the ISO image and the production image are essentially the same. In two of these seven cases, however, another alternative is a close second (see Appendix C). Thus, for "do not lock," although subjects gave the "circle, slash, padlock" image most frequently, the image of "circle, slash, keyhole and key" was given almost as frequently. Similarly, for "keep fire door closed," the image of "circle, slash, open door" was drawn almost as often as "door with arrows pointing closed." In the other five cases, "extinguisher," "no smoking," "do not use water," "hearing protection," and "head protection," the production image was most frequently the same as the ISO image.

Where the production image content differed from the ISO image content, it tended to be for symbols which had performed poorly in the understandability test. "Fire alarm call point," "do not block," "not an exit," and "exit" tested below 25 percent understandability for definition data with a strict criterion for the ISO symbols. These referents produced images which were different from those proposed by ISO. (The "fire exit" symbol which tested well was not proposed by ISO.) See Table 4 for a listing of the most frequently drawn images for these referents. Although "General hazard" had not been tested for understandability, the image drawn for it tended to resemble the "break glass for access" symbol (another poorly understood symbol). "No open flame" had tested reasonably well for understandability although its production image resembles the U.N. hazard warning image rather than the ISO TC 21 image.

Brainard et al. (1961) also found that concepts which were more accurately understood tended to be drawn more similarly. These authors suggested that their data indicate some common "stereotypes" for selected meanings.

From the subjects' drawings, it can be inferred that the image content proposed by ISO for several of the fire-safety symbols should be reconsidered. In particular, "fire alarm call point," "not an exit," "exit," and "hazard" were generally portrayed differently from the ISO suggestion. In addition, for the egress-related concepts no consistent graphic image emerged. Furthermore, few of the proposed egress symbols test well. Yet, these are some of the most critical of the life safety symbols. In fact, subjects were not able to produce any consistent image

for "exit" other than the arrow, and the majority insisted on using the word "EXIT." Thus the egress-related symbols were frequently misidentified while no common images surfaced for egress-related referents.

It should be noted that subjects may have felt inhibited or constrained by their inability to draw images for a specific referent. Green (1979) commented that lack of drawing skill can be a problem with production data. Another concern is that the specific wording of the message provided to the subjects may have influenced the images produced. Subsequent investigations might benefit from multiple phrasings, which would help indicate the desired degree of generality and discourage literal responses (e.g., figure with cane for "blind alley"). Nevertheless, the production data indicate where there may be problems with existing imagery, particularly when considered in conjunction with the understandability data. They also suggest alternative images for consideration.

5.0 CONCLUSION

A number of conclusions can be drawn from the results obtained in this experiment. The data indicate that the presentation method does not significantly affect the understandability of the symbols. As a result, the use of slides, placards or booklets may be governed by the exigencies of the particular experiment, based upon the practical advantages of each method.

While the method of response (multiple choice or definition) and the method of scoring (strict or lenient) does have an effect, the agreement between methods remains generally good. Rank-orders of the symbols are similar and the same symbols tend to be classified as acceptable or unacceptable by each method.

In conjunction with the use of multiple choice answers, the data indicate that the use of confidence ratings provides a good indication of guessing and of the overall acceptability of the symbol. Instances where there is a discrepancy between the percentage correct and the confidence rating suggest that the symbol is not fully understood. Furthermore, the selection of a high confidence rating for the wrong multiple choice alternative can indicate that dangerous or confusing misconceptions may exist for that image. Ideally, however, the most successful use of the multiple choice method is contingent upon the prior use of a definition procedure for generating "wrong" answers. These answers can then be used as plausible responses in the multiple choice procedure. For this reason, the multiple choice procedure is not recommended for initial or small-scale studies.

The production data, along with the understandability data, provide valuable insight into the effectiveness of a particular image for a given referent. A selected image is most likely to be optimal if it is both accurately recognized on the understandability task and frequently drawn on the production task. Similarly if images, such as those proposed for

egress, do poorly in both tasks, this suggests that the image concept is more difficult to portray graphically, and that graphic redesign and user education are needed. Finally, of course, the production data suggest alternative ways of portraying a particular concept that should be assessed further.

The overall data on the understandability of fire safety symbols indicate that some symbols are understood extremely well, while others perform very poorly. Symbols which do well include "fire extinguisher," "hose and reel," "no smoking," "general telephone," and "fire exit," while symbols which perform less well include "fire alarm call point," "do not block," "break glass for access," "not an exit," "fire fighter's equipment," and "direction to equipment." With the exception of "fire exit," which was not tested, the same pattern of responses for these symbols was found by Collins and Pierman (1979) for a different subject sample. In fact, the major differences found between the previous experiment and the present experiment could be attributed to the much poorer performance of the very elderly subjects used by Collins and Pierman. This difference reinforces the need to evaluate symbols with subjects of diverse backgrounds and ages.

In summary, the extreme range of understandability of the symbols investigated in this experiment indicates the need to test symbols before widespread adoption. Particularly troubling is the low meaningfulness and the incidence of potentially dangerous confusions of some symbols proposed for international standardization. As a result, it may be beneficial to incorporate some testing procedures as integral parts of the symbol development process, and not solely as a post hoc evaluation. Criteria other than understandability--such as visual range, detection in smoke, and noticeability--also require consideration for many applications. Nonetheless, understandability remains of primary concern in achieving effective and widely-accepted symbols.

REFERENCES

Brainard, R. W., Campbell, R. J., and Elkins, E. H., Design and Interpretability of Road Signs, Journal of Applied Psychology, Vol. 45, 1961, pp. 130-136.

Cahill, M. C., Interpretability of Graphic Symbols as a Function of Context and Experience Factors, Journal of Applied Psychology, 1975, Vol. 60, No. 3, pp. 376-580.

Cahill, M. C,, Design Features of Graphic Symbols Varying in Interpretability, Perceptual and Motor Skills, 1976, 42, pp. 647-653.

Collins, B. L., and Pierman, B. C., Evaluation of Safety Symbols, National Bureau of Standards (U.S.), NBSIR 79-1760, 30 pages, (June 1979).

Dewar, R. E., and Ells, J. G., Comparison of Three Methods for Evaluating Traffic Signs, Transportation Research Record, 1974, 503, pp. 38-47.

Dewar, R. E., and Ells, J. G., The Semantic Differential as an Index of Traffic Sign Perception and Comprehension, Human Factors, 1977, 19(2), pp. 183-189.

Dewar, R. E., Ells, J. G., and Mundy, G., Reaction Time as an Index of Traffic Sign Perception, Human Factors, 1976, 18(4), pp. 381-392.

Dewar, R. E., and Swanson, H. A., Recognition of Traffic Control Signs, National Academy of Sciences, NRC, Highway Research Board, 1972, 414, pp. 16-23.

Easterby, R. S., and Hakiel, S. R., Safety Labelling and Consumer Products: Field Studies of Sign Recognition, AP Report 76, Applied Psychology Department, University of Aston, Birmingham, England, December 1977.

Easterby, R. S., and Zwaga, H. G., Evaluation of Public Information Symbols, ISO Tests: 1975 Series, AP Report 60, Applied Psychology Department, University of Aston, Birmingham, England, March 1976.

Eliot, W. C., Symbology on the Highways of the World, Traffic Engineering, 1960, 31, pp. 18-26.

Ells, J. G., and Dewar, R. E., Rapid Comprehension of Verbal and Symbolic Traffic Sign Messages, Human Factors, 1979, 21(2), pp. 161-168.

Forbes, T. W., Gervais, E., and Allen, T., Effectiveness of Symbols for Lane Control Signals, Highway Research Board Bulletin, 244, 1963, pp. 16-29.

Freedman, M., Symbol Signs - The Testing of Passenger/Pedestrian Oriented Symbols for Use in Transportation Related Facilities, Personnal Communication, 1979.

Freedman, M., and Berkowitz, M. J., Preliminary Report on Laboratory and Pilot Field Testing, Testing Criteria and Techniques of Evaluation for Passenger/Pedestrian Oriented Symbols for Use in Transportation Related Facilities, DOT-OS-60071, FIRL No C448, January 1977.

Freedman, M., Berkowitz, M. J., and Gallagher, V. P., Symbol Signs, Testing Criteria and Techniques of Evaluation for Passenger/Pedestrian Oriented Symbols for Use in Transportation Related Facilities, Interim Report, U.S. DOT, September 2, 1976.

Green, P., Development of Pictographic Symbols for Vehicle Controls and Displays. Society of Automotive Engineers, Technical Paper (Warrendale, Pennsylvania) - #790 583, February-March, 1979.

Green, P., and Pew, R. W., Evaluating Pictographic Symbols: An Automotive Application, Human Factors, 1978, 20(1), pp. 103-114.

Griffith, D., and Actkinson, T. R., International Road Signs: Interpretability and Training Techniques, Proceedings of the Human Factors Society - 21st Annual Meeting, 1977, pp. 392-395.

Griffith, D., and Actkinson, T. R., International Road Signs: Interpretability and Training Techniques, U.S. Army Research Institute for the Behavioral and Social Sciences, Report 1202, September 1978.

Heard, E. A., Symbol Study - 1972, Society of Automotive Engineers, Automotive Engineering Congress, Detroit, Michigan, February 25-March 1, 1974, 740 304.

ISO/TC 21/SC 1, Equipment for Fire Protection and Firefighting - Safety Signs, Draft Proposal 6309, Geneva: International Organization for Standardization, 1978

ISO/TC 80, Safety Colours and Safety Signs, Draft International Standard - ISO/DIS 3864.3, Geneva: International Organization for Standardization, 1979.

Janda, H. F., and Volk, W. N., Effectiveness of Various Highway Signs, National Research Council - Highway Research Board Proceedings, 1934, 14, 442-47.

King, L. E., A Laboratory Comparison of Symbol and Word Roadway Signs, Traffic Engineering and Control, 1971, 12, pp. 518-520.

King, L. E., Recognition of Symbol and Word Traffic Signs, Journal of Safety Research, 1975, 7, pp. 80-84.

Smillie, R. J., Continuing Research in Job Performance Aids: The Interaction of Speed Stress and Media, Proceedings of the Human Factors Society - 22nd Annual Meeting, 1978, pp. 502-506.

Walker, R. E., Nicolay, R. C., and Stearns, C. R., Comparative Accurary of Recognizing American and International Road Signs, Journal of Applied Psychology, 1965, Vol. 49, #5, pp. 322-325.

Wiegand, D., and Glumm, M. M., An Evaluation of Pictographic Symbols for Controls and Displays in Road Vehicles, Technical Memorandum 1-79, U.S. Army, Human Engineering Laboratory, February 1979.

Yannone, A. G., "Fire Exit" Symbol, Copyright 1979. Available from International Safety Signs, Inc., 33 West Elm Street, Brockton, MA 02501.

Appendix A

Complete Communications to Subjects (Agreement Form, Instructions, and Debriefing).

NBS-783
(2-75)

U.S. DEPARTMENT OF COMMERCE
NATIONAL BUREAU OF STANDARDS

3. Cost Center No.
7434110

RESEARCH PARTICIPANT AGREEMENT

1. Principal Investigator: Belinda Lowenhaupt Collins

2. Division/Section: 743

4. Location: [X] Gaithersburg [ ] Other (specify)

5. Experiment Name Code: Evaluation of Fire-safety Symbols

6. Description of Experiment

This study will supply information on the recognizability and understandability of a set of 24 Fire Safety symbols for a sample of office workers.

7. Risks to Participant

This study presents no hazards to the participants other than those hazards common to everyday office existence, such as trips and falls. No injuries are anticipated.

8. Responsibilities of Participant

The participant will: 1) attend one, one-hour experimental session as designated on schedule; 2) follow instructions given by experimenter; and 3) fill out response form accurately. Participants are free to terminate their participation at any time.

9. Responsibilities of Investigator(s)

The investigator(s) will:

1) ensure the safety of all participants at all times;
2) schedule the experimental sessions at times initially agreed to by the participants.
3) provide participants with knowledge of results upon completion of the experiment.
4) maintain the confidentiality of records in accordance with the requirements of the Privacy Act of 1974.

10. IT IS UNDERSTOOD THAT EITHER THE PRINCIPAL INVESTIGATOR, THE PARTICIPANT, OR THE PARTICIPANT'S PARENT OR GUARDIAN MAY TERMINATE THE PARTICIPANT'S INVOLVEMENT IN THE RESEARCH AT ANY TIME WITHOUT INCURRING LEGAL LIABILITY FOR SUCH TERMINATION.

11. I hereby certify that my participation is voluntary and that I have read and accept the terms of this agreement.

Participant, or Parent or Guardian *(Signature)*	Date
12. Principal Investigator *(Signature)*	Date
13. Early Termination by *(Signature)*	Date

USCOMM-NBS-DC

Instructions [Definition]

Please give a short definition of the meaning of each fire-safety symbol that we present to you. Each definition should be no longer than ten words. Even if you do not know what the symbol means, please guess at a definition. After you have given the definition for the symbol, please rate how confident you are that this answer is correct. Circle the number from 1 to 5 that best describes how certain you are that your answer is correct. One means very uncertain and 5 means very certain.

There are 25 symbols in all. You will have two minutes per symbol.

Instructions [Multiple Choice]

Please circle the answer that is the best definition of the meaning of each fire-safety symbol that we present to you. There are four choices for each symbol. After you have circled the best answer, we would like you to rate each of the four choices. Using the rating scale, please circle the number from 1 to 5 which best indicates how certain you are that each choice is correct. One means very uncertain and 5 means very certain.

There are 25 symbols in all. You will have two minutes for each symbol.

Instructions [Production]

On each of the following pages, there is a definition. We want to know how you think this definition should be symbolized. In the space below the definition, please draw a symbol that conveys the meaning of the definition. Please do not use words. We are not interested in a great work of art-a simple sketch will do.

There are 15 definitions in all. Please do not spend too much time on any one item.

[Debriefing Statement]

The purpose of this experiment is to determine the effectiveness of different methods of presenting symbols to you. It is also intended to see whether these symbols are understandable to you. Because these symbols are currently being proposed for adoption as an international standard for fire-safety information in buildings, it is important to determine whether they convey the intended meaning to people in the United States. The "answers" are given on the next page. We appreciate your cooperation with us in determining the understandability of these symbols.

1. (Fire emergency) exit
2. Fire extinguisher
3. Fire fighter's equipment
4. Do not lock
5. Do not block - keep passageway clear
6. Fire alarm call point
7. Do not use water
8. No smoking
9. No open flame
10. (General) telephone
11. (Emergency) telephone
12. Break glass to obtain access
13. Blind alley
14. Slide door to right to open
15. Fire ladder
16. Standpipe (hose and reel)
17. Direction to fire-fighting equipment
18. Fire bucket location
19. Exit (green)
20. Exit (red)
21. Fire alarm horn
22. Fire door - keep shut
23. Russian exit
24. Russian no exit
25 Fire exit

Appendix B

Frequency and Kind of Incorrect Responses for Each Symbol for Both Response Types

Referent	Multiple Choice	N	Definition	N
#1. Fire Extinguisher	Gas Pump	0	No Gas	1
	No Water	0		
	Flammable Material	0		
#2. Hose and Reel	Sprinkler	1	Radiator	2
	No Water	1		
	Radiator	0		
#3. Fire Ladder	Escape Ladder	33	Escape Ladder	19
	Stairway	2	Stay off Ladder	2
	Not Fire Escape	1		
#4. Fire Bucket	Flammable Liquid	1	Trash Can	6
	Garbage Dump	1	No Water	1
	No Water	0	Wash Bucket	1
#5. Fire Fighter's Equipment	Tunnel	25.5	Tunnel	7
	Shelter	17.5	Danger	4
	Hard Hat	3	Exit	3
			Halfmoon	4
			Handle	1
			No Passage Thru Here	1
			Opening	1
			Fire Hazard	1
			Sprinkler System	1
			Fire Bucket Location	1
			Crawl Space	1
			Bridge, Underpass, Sewer	1
			Skylight Available	1
			Bald Head Rising in Red Sky	1
			Half Way Down	1
			Leave Lights On - Unsafe When Poorly Lit	1

Appendix B - Continued

Referent	Multiple Choice	N	Definition	N
#6. Direction to Equipment	One Way	21	Emergency Exit	13
	Exit	12	Go Left	11
	Don't Go This Way	2	Go Right	2
			Follow Arrow	10
			This Way Only	3
			No Left Turn	1
			Direction-Dangerous	1
			Emergency Route	1
#7. Break Glass for Access	Explosion	17	Explosion	11
	Broken Glass	15	Blasting	9
	Fireworks	1	Broken Glass	9
			Fire Hazard	2
			Plate Glass, Breakable	1
			Starburst	1
			Firework Area	1
			Shatterproof	1
			Sign of A Clash, Accident	1
			Caution, Ending Fast	1
#8. Slide Door to Right to Open	Exit	16	Exit	15
	Narrow Passage	4	Exit to Right	8
	Safety Shelter	0	Elevator	3
			Door	3
			Something on Right	4
			Something Can Move Danger	1
			Map for Fire Drill	1
			To A Specified Area	1
			White Square Ahead	1
			Four Corners Station	1
#9. Do Not Use Water To Put Out Fire	Do Not Put Out Fire	6	Do Not Put Out Fire	1
	No Burning	5	No Fires Permitted	1
	Extinguisher	0	Extinguish Fire Immediately	1
			Bucket of Water - Pour On Fire	1

Appendix B - Continued

Referent	Multiple Choice	N	Definition	N
#10. Do Not Lock	Keep Locked	4	Do Not Unlock	2
	Do Not Enter	1	Do Not Enter-Locked	2
	Stored Materials	0	No Exit-Locked	1
			Padlock	1
			Locked Door or Cabinet	1
#11. No Smoking	Smoking Permitted	0	Do Not Discard Lighted	
	Ashes Permitted	0	Cigarette	1
	Fireworks Permitted	0	Cigarette Burning	1
#12. No Open Flame	Safety Match	4	No Blasting	1
	Grass Fire	1	Don't Fuel Fire	1
	Fire Extinguisher	1	Don't Throw Match	1
			Exit With Burning Match	1
			Match Burning	1
#13. Do Not Block	No Dumping	35	Do Not Place Trash Here	19
	Do Not Store	9	No Dumping	6
	Garbage Area	1	No Trash Cans, Collection	5
			Don't Throw Flammables Into Trash	1
			Trash Burning-Not Permitted	1
			Don't Put Lid on Garbage Can	1
			Window Bars and Boxes	1
			No Unloading	1
			No Boxes in Cans	1
#14. Fire Door	Door Swings	25.5	Door Opens/ Closes This Way	15
	Close Cover	7	Close Window	2
	Revolving Door	5	Exit Emergency	3
			Fold Down	2
			Open, Unlocked Door	2
			This Side Face Down	1
			Closed Cover	1

Appendix B - Continued

Referent	Multiple Choice	N	Definition	N
#15. Emergency Telephone	Pay Phone Only	2	No Phone	2
	Dangerous Phone	1	Telephone Repair	1
	No Phone	0		
#16. General Telephone	Phone Out of Order	3	Please Hang Up	
	Hang Up Receiver	2	Receiver	1
	Do Not Use Phone	1		
#17. Alarm Horn	Camera	5	Dead End	2
	Flash Light	3	Escape, Exit	2
	No Radios	2	Pathway	2
			Shelter, Aid	2
			Projector Overhead	1
#18. Fire Alarm Call Point	Target Area	24	Life Preserver	4
	Life Preserver	6	Target Area	3
	Fire Lane	1	Doughnuts Served	3
			Dangerous Area	3
			Do Not Enter	3
			Hot Circle	2
			Evacuation, Escape Route	2
			Slide Down Pole	1
			Dead End	1
			Fire Station Sign	1
			Do Not Touch	1
			Stop, Proceed w/ Caution	1
			Receptacle	1
			Break In Case of Emergency	1
			Fire Hazards Located Here	1
			Circles on Squares	1
			Red Circle w/ White Center	1
			Aid Station	1
			Wide Road	1
#19. Fire Exit- (Yannone)	Stand Back From Fire	2	Dangerous Fire	2
	Do Not Leave	0	Run From Fire	2
	Do Not Exit	0		

Appendix B - Continued

Referent	Multiple Choice	N	Definition	N
#20. Exit - ISO Original	Container	12	Entrance	2
	Oxygen	7	Keep Door Closed	2
	No Exit	1	Open	2
			Window	1
			Parking Area	1
			Elevator	1
			What's Behind the "White Door"	1
			Right-Hand Turn	1
			OK To Build Fires in Open Places	1
			Refrigerator Ahead	1
			Four Corners	1
#21. Exit - ISO Proposed	Running Permitted	12	Running Permitted	2
	Radiation	1	Shut Door	1
	Watch Out For Pedestrians	1	Door Opens In	1
			Entrance Permitted	1
			Pass with Caution	1
#22. No Exit - ISO Proposed	No Running	21	Do Not Enter	10
	No Pedestrians	5	Do Not Open Fire Door	3
	Poorly Lit Area	2	No Running	2
			Go Through Door and Down Side	1
			Don't Tip Over Bed	1
			Don't Use Elevator	1
			Be Careful On Stair	1
			Do Not Leave House	1

Appendix B - Continued

Referent	Multiple Choice	N	Definition	N
#23. No Exit - ISO Original	Overhead Hazard	21	Shelter Area	9
	Exit	9	Go Up to Roof	8
	Elevator	2	Go Straight Ahead	3
			Enter With Caution	5
			This Way To Elevator	2
			Yield At All Stops	2
			Stay Within Certain Area	1
			Pyramid	1
			Arrow Pointing Inside Building	1
			This End Up	1
			Skylight Above	1
			Closed In-One Way	1
			A Study in Black, Yellow, and White	1
			Fire Alarm Box	1
#24. Exit - U.S. Green	Blocked Exit	2	Daily Use	1
	Break Window	1		
	No Exit	0		
#25. Exit - U.S. Red	No Exit	3	No Fire Exit Here	1
	Stop	1		
	Smoking	0		

Appendix C - Description of Results of Production Method

Referent	Pictorial Content (frequency of occurence)
1. Fire Extinguisher	Fire extinguisher alone (79) Extinguisher spraying fire (5) Extinguisher next to fire (2)
2. Fire Alarm Call Point	Telephone receiver (58) Phone alone (25) Phone and flame (22) Phone and various other objects (11) Horn, bell, gong (11) 4 of these shown with phone Lever, button (10) 1 shown with phone Concentric figures of various sorts (4) Fire helmet (2) Fire truck (2)
3. No Smoking	Cigarette with slash or X (88) Pipe, cigar and/or cigarette and slash (3)
4. No Open Flame	Flame with slash or X (62) Lighted match with slash or X (23) Flame alone (3) Cigarette and slash (2) Flame and "No" (2)
5. Do Not Use Water To Put Out Fire	Fire, bucket, slash or X (76) Fire, hose, slash (4) Bucket and slash (3) Fire and bucket (2) Dripping faucet and slash (2)
6. Do Not Lock	Lock, slash or X (55) Key, keyhole, slash or X (20) Door with padlocks, slash or X (7) Key, lock, slash or X (5)
7. Do Not Block - Keep Passageway Clear	Door or corridor with objects, slash or X (36) Boxes and/or objects, slash or X (22) Door or corridor, slash or X (9) Open doorway or corridor (6) Figure, door, slash (5)

Appendix C - Continued

Referent	Pictorial Content (frequency of occurrence)
8. Keep Fire Door Shut	Door shown closing with arrows (32) Open door, slash or X (25) Closed door (23) Figure or hand closing door (4) Closed door, slash (2)
9. Blind Alley	Various attempts to show corridors with one path or arm obscured (21) Rectangular corridor with the end darkened or slashed (17) Door and slash (11) No figure (7) With figure (4) Converging lines (8) Blind person and door (6) Arrow going into square (4) Arrow, slash (4)
10. Exit	Arrow (18) Arrow alone (11) Arrow and word "Exit" (5) Arrow and flame (1) Row of arrows (1) (38 answers used word "Exit", 33 alone, 5 with arrow).
11. Fire Exit	Flame (66) Flame and arrow (20) Flame and door (14) Flame and word "Exit" (13) Flame, figure, arrow (10) Flame, figure, door (9)
12. Hazard Area	Explosion, shattered object, asterisk (intended image often not clear in sketches) (20) Figure amidst objects (9) Dangerous objects, e.g., guns, dynamite (6) Falling objects (3) Skull and bones (5) Circle and X (5) Circle and dot (3)

Appendix C - Continued

Referent	Pictorial Content (frequency of occurance)
12. Hazard Area (continued)	Triangle (3) Zig-zag (3) Hard hat (2) Exclamation point (2) "No Entry" message - door or figure and slash or X (3)
13. Hearing Protection Must Be Worn	Ear Muffs (76) Alone (27) On head (45) On figure (4) Ear(s) (16) Alone (10) With stopple (6)
14. Head Protection Must Be Worn	Hard hat or helmet (90) Hat not on head (37) With falling or bouncing objects (2) With arrows pointing to hat (1) Hat on head (46) Alone (37) With objects (4) With arrows (5) Hat on figure (7) Alone (3) With objects (1) With arrows (3)
15. Eye Protection Must Be Worn	Eyeglasses alone (46) Goggles alone (12) Glasses or goggles on face (25) Glasses or goggles with projectiles (7)

NBS-114A (REV. 9-78)

U.S. DEPT. OF COMM. BIBLIOGRAPHIC DATA SHEET	1. PUBLICATION OR REPORT NO. NBSIR 80-2088	2. Gov't. Accession No.	3. Recipient's Accession No.

4. TITLE AND SUBTITLE

The Assessment of Safety Symbol Understandability by Different Testing Methods

5. Publication Date

August 1980

6. Performing Organization Code

7. AUTHOR(S)

Neil D. Lerner and Belinda L. Collins

8. Performing Organ. Report No.

9. PERFORMING ORGANIZATION NAME AND ADDRESS

NATIONAL BUREAU OF STANDARDS
DEPARTMENT OF COMMERCE
WASHINGTON, DC 20234

10. Project/Task/Work Unit No.

11. Contract/Grant No.

12. SPONSORING ORGANIZATION NAME AND COMPLETE ADDRESS *(Street, City, State, ZIP)*

National Bureau of Standards

13. Type of Report & Period Covered

Final

14. Sponsoring Agency Code

15. SUPPLEMENTARY NOTES

☐ Document describes a computer program; SF-185, FIPS Software Summary, is attached.

16. ABSTRACT *(A 200-word or less factual summary of most significant information. If document includes a significant bibliography or literature survey, mention it here.)*

This paper reports an experiment on the understandability of pictorial symbols proposed for fire-safety alerting. The experiment was designed to determine the understandability of specific symbols and to assess the effects of variations in both presentation and response methods.

The symbols were presented as slides, booklets, or placards. Subjects indicated their understanding of each symbol's meaning either by writing down a brief definition or by selecting the correct answer from among four alternatives. For both methods, subjects rated their confidence in the correctness of the answers. In the second phase of the experiment, subjects were given fifteen different messages, and asked to draw a symbol for each idea.

Mode of symbol presentation had no effect on understandability, while the use of definition and multiple choice procedures led to generally similar conclusions. The confidence ratings provided additional information about discrepancies between the two response methods.

The understandability of the 25 symbols ranged from near zero to virtually total comprehension. These data underscore the need to determine the understandability of safety symbols prior to standardizing a symbol set.

17. KEY WORDS *(six to twelve entries; alphabetical order; capitalize only the first letter of the first key word unless a proper name; separated by semicolons)*

communication; evaluation method; fire-safety; hazard warnings; meaningfulness; method; pictogram; response; symbol; understandability

18. AVAILABILITY ☒ Unlimited

☐ For Official Distribution. Do Not Release to NTIS

☐ Order From Sup. of Doc., U.S. Government Printing Office, Washington, DC 20402, SD Stock No. SN003-003-

☒ Order From National Technical Information Service (NTIS), Springfield, VA. 22161

19. SECURITY CLASS (THIS REPORT)	21. NO. OF PRINTED PAGES
UNCLASSIFIED	60
20. SECURITY CLASS (THIS PAGE)	**22. Price**
UNCLASSIFIED	$7.00

USCOMM-DC