NBSIR 81-2268

Evaluation of Exit Directional Symbols

Neil D. Lerner

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

May 1981

U.S. DEPARTMENT OF COMMERCE

NATIONAL BUREAU OF STANDARDS

NBSIR 81-2268

EVALUATION OF EXIT DIRECTIONAL SYMBOLS

Neil D. Lerner

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

May 1981

U.S. DEPARTMENT OF COMMERCE, Malcolm Baldrige, ***Secretary***
NATIONAL BUREAU OF STANDARDS, Ernest Ambler, ***Director***

ABSTRACT

This paper discusses visibility considerations for exit symbols and the relationship between understandability and visibility concerns. Two experiments evaluated directional indicators (arrows) in the context of building exit signage. The first experiment compared the visibility of 16 arrows under degraded visual conditions that were comparable to a smoke environment. This experiment had two objectives: (1) the development of a methodology for assessing exit pictogram visibility; and (2) a comparison of the visibilities of the specific arrows tested. A second experiment obtained subjective rankings of the arrow types on the basis of several criteria of concern for exit signage. These criteria included connotative meaning, uniqueness from other directional indicators, and appropriateness.

The visibility procedures were statistically sensitive and demonstrated differences in the visibility of different arrows. An analysis of the type of confusions that occurred in errors for each arrow suggests certain relations between graphic features and errors in detection. The second experment indicates substantial agreement between participants in ranking the arrows and revealed strong relationships between the several criteria. Together the results of the two experiments were used to evaluate the set of arrows for appropriateness for use with exit designators.

Methodological issues in evaluating symbol visibility are considered and other issues of concern in testing exit signage, such as special user groups, are also discussed.

TABLE OF CONTENTS

TABLE OF CONTENTS (cont.)

1. INTRODUCTION

Directional indicators, or "arrows," play an important role in building emergency signage. This is especially so in designating exits, where highly visible and readily comprehended directional information can be crucial to life safety. The manner of indicating direction is only one aspect of the larger issue of developing optimal graphic symbols to designate the location of exits. Therefore, broader issues of exit symbol visibility and understandability are discussed below. In this context, the special role and limitations of arrows in conveying the egress message become more apparent. The two experiments described in this report investigate the adequacy of various arrows. The results of this research are applicable both to directional indicators in particular and to building emergency signage in general.

1.1 PICTORIAL EXIT SYMBOLS

Pictorial signs are increasingly used as alternatives to written signs because they have a number of potential advantages as communication devices (Collins and Pierman, 1979). Although highway applications are probably the most commonly encountered pictorial signs, many public buildings now use symbols to provide informational, safety, and emergency messages to occupants. The terms "symbol", "pictogram", and "pictorial sign" refer here to any graphic image, whether highly representational, abstract, or arbitrary, that is used in place of written messages in signs. Unfortunately, although some pictograms are well understood by the public (e.g., "no smoking" symbols, Lerner and Collins, 1980), other messages remain difficult to communicate symbolically. Among these is the critical message of "exit". Outside of English speaking countries, the word "exit" is not widely used (American Institute of Graphic Arts (AIGA), 1979); thus, a sign based on this written message is not likely to be effective to foreigners visiting the U.S., nor are Americans abroad likely to encounter it. A meaningful exit pictogram would be of obvious value. However, an effective exit symbol not only must be well-understood, but must also meet demands of visibility as well. The exit message may be the most crucial life safety message to the general public during fire or other emergencies, and so the performance of an exit symbol may be most critical at a distance under conditions of emergency lighting or smoke. Thus, an ideal exit symbol would unambiguously convey to occupants the message of egress with optimal visibility (the term "visibility" is defined below).

Many symbolic versions of the "exit" message have been developed, including efforts by various international and national standards groups and agencies. These symbols show a great range of variability, both conceptually and graphically. For example, some exit symbols are simple, bold geometric shapes, based on highly abstracted (e.g., rectangle proportioned as a door) or arbitrary (e.g., bisected disc) images. At the other extreme there are highly detailed pictograms of a realistic figure fleeing flames, with an outside view of trees visible through a doorway. As these few examples may suggest, symbols may vary widely both in their meaningfulness to the public and in their visibility. Most examples have not been explicitly tested for these qualities, but several experimental evaluations (e.g., Lerner and Collins, 1980; Japanese proposal, 1980) have, in fact, demonstrated important differences in either of these factors.

The concern in this report is primarily with visibility. However, in any evaluation of exit symbol performance, understandability and visibility must be jointly considered. The term "visibility" is used in this report to refer generally to perception under poor visual conditions. In fact, depending upon the psychological response (detection, discrimination, recognition, interpretation, identification), a variety of technical terms have specific (though not always agreed upon and sometimes overlapping) operational meanings. These terms include "detectability," "discriminability," "readability," "legibility," "visibility," "conspicuity," and others. Since, visibility has been used widely in experimental literature similar to that described in this report (e.g. Smith and Weir, 1978; Jacobs, Johnston, and Cole, 1975), it is adopted throughout. However, it should be noted that the term has sometimes been restricted to the process of simple detection (presence or absence of a stimulus), and not to higher-order perceptual processes. The broader use of the term is followed here. "Understandability" refers to the comprehension of the intended sign message by potential users. Many different methods have been used to assess understandability (reviewed by Collins, 1981), such as having research participants provide definitions for a symbol, match symbols and messages, rate the meaningfulness of a symbol on some subjective scale, or engage in behavior as directed by a symbol sign. While the understandability of a symbol would certainly decrease as visibility is degraded, the understandability of exit symbols has typically been evaluated under optimal viewing conditions. Therefore, as used in this report, understandability may be interpreted as a measure of how well the intended symbol meaning is comprehended under good viewing conditions.

In the following sections, the issue of exit symbol visibility will be discussed, and then the problem of symbol understandability and its relation to visibility will be briefly considered. Finally, a special concern with exit directional indicators ("arrows") will be discussed. The experiments that follow focus on these directional indicators.

1.2 EXIT SYMBOL VISIBILITY

The visibility of an exit indicator during emergency conditions is an important concern. Smoke is a principal threat to life during fires, since problems of perception in smoke, which disrupt escape, usually precede the thermal and toxic effects of fires (Gross, Loftus, and Robertson, 1966). Compounding the visual disruption of way-finding by smoke are the disorienting effects of smoke, fear, lachrymation, and possibly even confusion due to central nervous system interference by some noxious smoke products (Phillips, 1978). While this disruption serves to emphasize the importance of egress signage during emergencies, smoke obscuration can be a special problem for signs: since smoke density increases with distance above the floor, obscuration may be greater at the level of the signs (such as above a door) than toward the floor (e.g., Demaree, 1979). Automatic sprinkler systems can drastically influence the pattern of smoke as well. Tests of simulated hospital room fires reveal that initiation of the sprinklers can invert the flow of gases from a burning room to a corridor (O'Neill, Hayes, and Zile, 1980). This rapidly results in a relatively uniform ceiling-to-floor obscuration in the adjacent corridor. Thus, while sprinklers may be an effective means of fire containment, they can also serve to accelerate visibility problems. The increased use of automatic sprinkler systems therefore should lead to even more concern with exit signage visibility.

Although there has been research on the visibility of signs in smoke, the focus of interest has generally been either on smoke parameters, such as composition and density, or on hardware parameters, such as illumination source and intensity or reflectivity (e.g., Demaree, 1979; Bono and Breed, 1965; Jin, 1970). Almost no research has focused specifically on the visibility of symbol signs for building environments. Some reports have compared the visibility of symbols versus word signs for highway applications (e.g., Jacobs, Johnston, and Cole, 1975; Dewar and Ells, 1974; Dewar and Swanson, 1972). Although symbols generally were more visible than written messages, these experiments used distances much too long and presentation times too brief to represent realistic conditions for buildings, however appropriate for highways.

One reported experiment has directly evaluated symbolic exit signs for visibility in buildings during emergencies (Japanese proposal, 1980). The Japanese proposal to the International Organization for Standardization (ISO) describes visibility experiments conducted by T. Jin. This research was part of a larger Japanese project evaluating exit symbols. Unfortunately, the translated report of the work does not provide great detail on research methods. Jin obtained "visual distances" for various symbols based on observations by 30 participants in a smoke-filled 14 m-long corridor. Thirteen pictographs were compared, most of which had been selected from a much larger set of images on the basis of judgments of graphic quality and general preference. Despite this preselection process, the 13 symbols still varied substantially in visibility, with the most visible having a visual distance about 35 percent greater than the least visible.

Jin's results are difficult to evaluate due to the sketchy discussion of methodology and uncertainty about what response was required of participants and how "visual distance" was determined. Hopefully, a more detailed translation of this work will be reported. Nonetheless, Jin's findings are valuable in providing a quantitative demonstration of how symbols can vary on the critical quality of visibility. In fact, given the preselection of symbols prior to the visibility experiment, Jin's data probably are quite conservative in revealing the range of visibility that could exist among currently used or proposed symbols.

1.3 VISIBILITY AND UNDERSTANDABILITY

Proposed exit symbols not only differ in their visibility but also in their understandability. Several previous research efforts at the National Bureau of Standards (NBS) concerned with the understandability of symbols for building and workplace safety (Collins and Pierman, 1979; Lerner and Collins, 1980; Collins and Lerner, 1980) included various exit designations among the symbols investigated. As a result, information has been collected for ten different "exit" and "emergency exit" symbols. The percentage of research participants that correctly identified the meaning of a given symbol ranged from under five percent to over 90 percent. The symbols most readily comprehended tended to be the most literal, showing walking or running figures, doors, and flames in various degrees of realism and detail. At the other extreme, highly symbolic, simple images were least understood.

Another finding emerging from the NBS studies was that symbols with intended meanings unrelated to egress may often be interpreted as indicating exit or refuge. These confusions raise the danger of an ambiguous symbol causing a person to seek safety in an inappropriate direction. With one exception the common feature of these symbols was their ambiguity, lacking any image of an identifiable object. An arrow was present in several, probably contributing to the egress interpretation (discussed further below).

These experiments not only indicate the need to explicitly evaluate the understandability of a symbol set; the data also suggest a relationship between understandability and visibility. The literal and relatively detailed images of the most readily understood images may be expected to be the least visible, while the simple, geometric images of the abstract symbols may be expected to be the most visible. While actual data on the visibility of most of the symbols is not available, these inferences may be grossly made by considering the features that contribute to perception of an image. These features include the size of major contrasting areas, size of critical detail, figure-ground relationships, continuity, simplicity, and contrast boundaries (Easterby, 1967; Follis and Hammer, 1979; Overington, 1976).

An interesting example of the trade-off between understandability and visibility comes from ISO's proposed emergency exit symbols. In the case of these two symbols there is in fact some data available on visibility. ISO initially proposed a symbol consisting of a white rectangle on a green background. In response to criticism that this symbol was unintelligible, ISO eventually replaced it with a new proposed image: a stylized figure running through a simplified image of a half-open door. This led to a substantial increase in understandability, as shown by Lerner and Collins (1980). Yet comparing visibility data from two similar experiments described in the 1980 Japanese proposal to ISO, this change decreased by about one-third the distance at which the symbol can be identified in a smoke-filled environment. Thus better understandability was achieved at a substantial cost in visibility.

The apparent inverse relationship between the understandability and visibility of exit symbols may not necessarily be true for other symbols. For instance, pictograms to designate the location of equipment ("fire extinguisher") or the use of required safety gear ("hard hat area") communicate effectively using simple, and probably highly perceptible, images. No single stereotypical image exists for the concept of emergency exit. Various images can contribute to the egress message -- figures, doors, arrows, fire, diagrams -- and these have been used in numerous combinations to produce exit pictograms. The design challenge is to simplify and minimize these images and their combination so as to maximize visibility while maintaining understandability. Consequently, the need to evaluate candidate exit symbols concurrently for both visibility and understandability is imperative. Testing and selecting a symbol on the basis of only one of these properties is likely to result in less-effective performance on the other property.

1.4 THE ARROW AS AN EXIT DESIGNATOR

Although no common stereotypical image exists for the concept of "exit", the arrow is probably the strongest element related to the exit idea. Lerner and Collins (1980) presented data to support this view. Participants in one portion of their experiment were asked to draw (produce) symbols to communicate various fire and safety messages, including "exit" and "fire exit."

"Exit" proved to be extremely difficult for participants to symbolize. Seven percent of the participants drew no image at all. Further, despite instructions not to use words (generally well-followed for other messages), another 42 percent of the participants included the word "exit" as all or part of the image. No common stereotype image was found, although the most frequently occurring image was an arrow. An arrow was incorporated into about 20 percent of all responses, and nearly 40 percent of those responses that did not use the word "exit." Thirteen percent of all responses used the arrow as the entire image.

For the "fire exit" message, participants again used the word "exit", this time in nearly one-third of the responses. Flame was a very common element of the images, occurring in about three-quarters of all responses. The arrow was again a frequent image, being incorporated into about one-third of the drawings.

An examination of existing and proposed exit symbols confirms the result of the Lerner and Collins experiment. The variety of different conceptualizations suggests that designers, too, do not share a stereotypical idea for exit. But again, the arrow emerges as a frequent component of the exit symbol. Figure 1 presents a number of examples of the ways in which arrows have been incorporated into exit symbols.

Another indication of the "exit" connotation of arrows comes from confusions in symbol meaning. In testing a set of fire-safety symbols, Lerner and Collins (1980) noted that ambiguous symbols that incorporated arrows were likely to be misinterpreted as suggesting egress. For example, a symbol showing an arrow entering a square area, on a triangular field, was proposed by ISO indicate "not an exit." More participants interpreted this symbol as egress-related than the correct "no egress" interpretation.

Despite its connotation of exit, an arrow by itself is not adequate to denote an exit unambiguously. Lerner and Collins (1980) included an arrow among the symbols they tested for understandability. An arrow was clearly interpreted to mean "exit" by only about one-third of the participants. Virtually all participants recognized the symbol as providing directional information, but most did not explicitly state this to be direction to an exit. This parallels the ubiquitous use of the arrow for all sorts of directional signage (Follis and Hammer, 1979), so that the meaning of an arrow by itself is ambiguous.

FIGURE 1. Use of arrows in exit symbols

Thus, the contribution of an arrow to building signage is not straightforward. As with many other messages, "exit" often requires some form of directional indicator to be used with it. In this case, however, the arrow also appears to contribute to the message itself. At the same time, the arrow can cause confusion from other messages, and is by itself insufficient to communicate the idea of an exit in an unambiguous manner. If an arrow is integrated into an exit symbol consisting of other elements, there is the further concern that the added feature may increase complexity and possibly reduce the size of the individual features, and so reduce the perceptibility of the symbol.

1.5 THE POSSIBILITY OF A SPECIFIC ARROW FOR USE WITH EXIT DESIGNATORS

The previous section indicated that despite its connotations, the arrow may not be desirable as a major element of an exit symbol. The American Institute of Graphic Arts (AIGA, 1979) has also discussed some problems with incorporating the arrow into the symbol itself. Nonetheless, it is clear that any exit sign, whether in symbol or word form, will frequently be used in conjunction with a directional indicator. Because existing exit symbols are often unclear in meaning, and because an arrow connotes "exit", a unique and effective directional indicator for use only in conjunction with exit signage may help reinforce the exit message. Making the use of a particular arrow specific to exits, and selecting an arrow that is distinct from other directional indicators, may reduce confusability with other signs and symbols.

Thus, the use of a unique directional indicator to be used only in conjunction with exit signs (but not in isolation) could provide the following advantages:

1. The arrow may reinforce the exit message both by contributing its own connotative meaning and by providing redundancy in the visual message.

2. The uniqueness of the arrow may reduce confusability with other messages, especially for poorly understood symbols and symbols containing arrows. Signs and symbols containing no arrow or an arrow other than the "exit arrow" could be interpreted as unrelated to exit even if their actual meaning remained ambiguous.

3. The arrow may increase the "learnability" of a standardized exit symbol. Many candidate exit pictograms are simply not well understood. Furthermore, as the drawings of Lerner and Collins' (1980) participants suggest, the "exit" word sign may be entrenched in the public mind as the way to indicate egress. These two factors impede the adoption of a standardized exit symbol, despite the potential advantages such a symbol may provide. If a unique arrow were adopted, it could be used with existing exit designators as well as symbols, and so become associated with the exit message. It could then be effective in helping to bridge understanding in those situations where observers are unfamiliar with the exit symbol.

4. The arrow may serve to increase the visibility of exit designators, especially under emergency conditions, such as smoke or emergency lighting. Some proposed symbols have poor legibility, whereas some arrows may have much better legibility. Problems of exit symbol interpretation

and confusion from other symbols would be more severe under poor visibility, and so the contribution of a unique arrow to the exit message might be amplified.

To realize these potential benefits of a unique exit-related arrow, the choice of the specific arrow may be critical. One goal of the experiments in this report is to provide initial data on visibility and other criteria important to exit directional indicators.

1.6 SUMMARY: THE EXIT SYMBOL PROBLEM

There is a recognized need for a standardized symbol to designate the location of exits. Many symbols have been used or proposed that fail to be well understood. Others may not be sufficiently visible under emergency conditions, such as smoke, during which symbol effectiveness may be critical. The most visible symbols, in fact, appear to be among the least understood although there are few empirical data to support this apparent relationship. Serious confusions with other symbols also exist, so that unrelated messages may be interpreted as related to egress. Furthermore, no strong single stereotypical image appears to exist among people for symbolizing the exit idea. The closest approximation to such a stereotype is the arrow, which by itself is too ambiguous and ubiquitous to function as the sole exit designator. Thus although the proliferation of exit symbols is itself a problem, emphasizing the need for standardization, no candidate symbol appears clearly superior and entirely adequate.

To mitigate this problem, it is suggested that a distinct arrow could be standardized uniquely for use with exit symbols or exit word signs. This could help reinforce the meaning, reduce confusability with other messages, increase visibility, and possibly aid in the learning or interpreting of newly encountered exit symbols. Adopting such a standard directional indicator for exits represents a partial solution. Development of an effective meaningful exit symbol is ultimately the optimal solution.

Although understandability and visibility have been emphasized here as the primary criteria for a standard exit symbol, other considerations exist as well. One such factor is the symbol's discriminability from other existing signage. Additionally, if a symbol is to find general acceptance it must be graphically compatible with other signage and be generally satisfactory to the design community. Specific attention may need to be paid to certain special user groups as well: these might include children or the elderly, for example. Despite these and other additional considerations, however, the primary step in selecting an effective exit symbol will be to identify an image that is well understood and highly visible for the general public.

One purpose of the research presented in this report is to develop an effective and efficient laboratory procedure for assessing the visibility of exit symbols. The discussion of exit symbol performance points up the need for explicitly testing proposed symbols for both understandability and visibility. In the case of understandability, a variety of evaluation methods have been employed and a sizable research literature exists (Collins, 1981). In contrast, there has been little research appropriate for comparing the visibility of graphic alternatives for building signs. Thus the visibility research reported here is addressed to the methodological problem of establishing an appropriate general procedure, rather than determining the visibility of any specific set of exit symbols. At the same time, this research presents an opportunity to explore further the suggestion of standardizing a directional arrow for exclusive use with exit designators. The following section provides an overview of the specific goals and approach of the experiments. Sections 3 and 4 will describe the method and results of each experiment in detail.

2. SCOPE AND RATIONALE OF EXPERIMENTS

Two experiments are described in this report. The first concerns the visibility of a set of directional indicators ("arrows") under conditions of poor visibility. The means of manipulating visibility to simulate emergency viewing conditions is described in section 3. The purpose of the experiment included both methodological issues regarding the general evaluation of symbol visibility and specific comparisions of directional indicators.

A specific set of arrows, described in the next section, was selected for evaluation. The visibility of the specific arrows is of interest both because of the widespread use of arrows, and also because of the susggestion that a unique arrow be specified for use with exit signage. The visibility experiment provided some initial comparison of various arrows and arrow features.

In addition to the interest in the performance of specific arrows, directional indicators are of value for addressing more general methodological questions in symbol visibility. Response measures to arrows are unambiguous and not strongly dependent on the symbol set. A participant merely indicates the direction in which an arrow is pointing, and the response is scored as correct or incorrect. In contrast, when pictograms are presented, answers may be more ambiguous, for example, a verbal description of what was seen. A variety of problems can arise, discussed in detail in section 5.5. To mention some of these briefly, performance may vary with the familiarity of the symbols, and this factor will itself vary during the course of experience in the experiment. Furthermore, recognition of symbols may be based on comparison of features within the set tested. For example, only one symbol in a set may have a strong diagonal component or a dark area in the lower left corner. Such factors may be irrelevent to how people perceive a symbol in a real environment, but may be a reliable means for experimental participants to distinguish symbols within a set. Data interpretation and scoring problems can also arise. Many of these problems have been encountered in other research (e.g., Zwaga, 1979), and using a directional response avoids them.

Among the questions concerning method are issues of the sensitivity of the procedure and the parameters selected: that is, are the data sensitive to variations in the visibility parameters selected, and can the visibility of the various stimuli be reliably discriminated. Another question concerns the stability of the performance of the research participants through the course of the experiment.

Data on arrow visibility also will be useful in providing reliable data for comparisons with future observations that may be made in smoke chambers. It is considerably more difficult and expensive to obtain psychophysical data from an actual smoke environment, and there are problems in maintaining a precise and uniform smoke field in a large area for durations typical of psychophysical experiments. The use of the more simple directional arrow procedure over more complex symbol recognition methods may be especially valuable in this case. Small-scale experiments with actual smoke situations could provide data against which laboratory procedures in the present experiment could eventually be compared and validated.

The second experiment was designed to supplement the first by providing data on the subjective responses of participants to the various directional indicators. At issue was whether research participants would show substantial agreement in evaluating how well various arrows served to connote the direction to an exit, how appropriate and attractive each arrow was, and how unique each arrow appeared. If there were to be substantial agreement between participants, then these data could provide further information on the suitability of different arrows as elements of exit designators.

The research described in this report was designed to develop a rationale, procedure, and working hypotheses for the systematic evaluation of exit indicators. It was intended as preliminary research, and therefore the findings should not be taken as recommendations for selecting any specific directional arrow. The set of arrows tested, and the sample of observers used, were determined by the requirements of an initial methodological investigation.

The arrows employed in the experiment were selected to provide a range of different graphic concepts. Several of these arrows are in common use and are sold widely through sign catalogues; others represent modifications of more common designs. Some were adapted from an experiment by Smith and Weir (1978). Some arrows differ from one another only in the value of some parameter (e.g., shaft width, head angle). This range of arrow types was intended to permit an evaluation of the sensitivity of the method to the form and parameters of the arrows, and to help suggest hypotheses about effective arrow elements or types. The collection of arrows was not intended to represent a set of candidates to be chosen from.

Similarly, the research participants were selected to meet the needs of a methodological pilot experiment. To assure a relatively homogeneous group with good acuity, all participants in the visibility experiment were females under thirty, with corrected-to-normal acuity and no reported visual defects.[1] While this may increase the sensitivity of the experiment, it obviously requires greater caution in generalizing the results to the general population.

---

[1] A number of age-related changes (some of which interact with sex) occur in visual perception (acuity, sensitivity to low illumination, difference thresholds for contrast) as well as in the physical and optical properties of the eye (Fozard, et al., 1977; Illuminating Engineering Society (IES), 1972; U.S. Public Health Service, 1964; Weale, 1963).

3. VISIBILITY EXPERIMENT

3.1 VISIBILITY METHOD

In the visibility experiment, viewing conditions were designed to simulate smoke effects. Smoke degrades visibility by two primary mechanisms: first, it reduces the light coming from the target; second, it scatters light from other regions into the viewing path, reducing contrast (Middleton, 1952). In the visibility experiment, light from the stimulus source was reduced by placing a neutral density filter in front of the stimulus slide projector. A second light source projected a veiling light superimposed upon the projected stimulus image. In addition, the ambient room lighting was dim (in the mesopic range) to typify emergency lighting and smoke conditions. The general procedure was adapted from research on highway signage described by Smith and Weir (1978), although a number of changes were made to better reflect a building emergency environment.

3.1.1 Research Participants

The participants were seven female volunteers, ranging in age from 18 to 28 years (mean = 21.3). They were recruited from the Gaithersburg, Maryland, area through local advertising and paid for participation. All reported normal or corrected-to-normal vision.

3.1.2 Stimuli

The stimuli were 32 directional indicators ("arrows"). Each of the 16 basic arrow shapes occured once as a black figure on a white background and again as a white figure on a black background. The 16 shapes are shown (as black on white) in figure 2. The stimuli were initially drafted as 15.2 cm x 15.2 cm (6 in x 6 in) placards, with the largest dimension of each arrow being 12.7 cm (5 in). The only exception was stimulus F (11.4 cm, 4.5 in), which was the same size as the identical portion of stimulus E.

For the visibility experiment, the stimuli were projected as slides. Each placard was photographed against a medium grey background. The projected image was square, with the placard centered. The length of the placard image was ≈ 83 percent of the overall dimension of the slide image, so that the placard occupied approximately 70 percent of the area of the projected image.

3.1.3 Apparatus

The arrows were presented as rear-projected slides, with the participant seated 3.05 m (10 ft) from the screen. The image was at approximate eye level, 1.12 m (44 in) above the floor. The size of the projected image was 8.4 cm per side, which subtended a visual angle of 1.6°; the width of the placard alone subtended a visual angle of 1.3°; and the greatest dimension of the arrow itself subtended 1.1°. A 1.3° projection of the placard is equivalent to a 12 in placard viewed from about 44 feet.

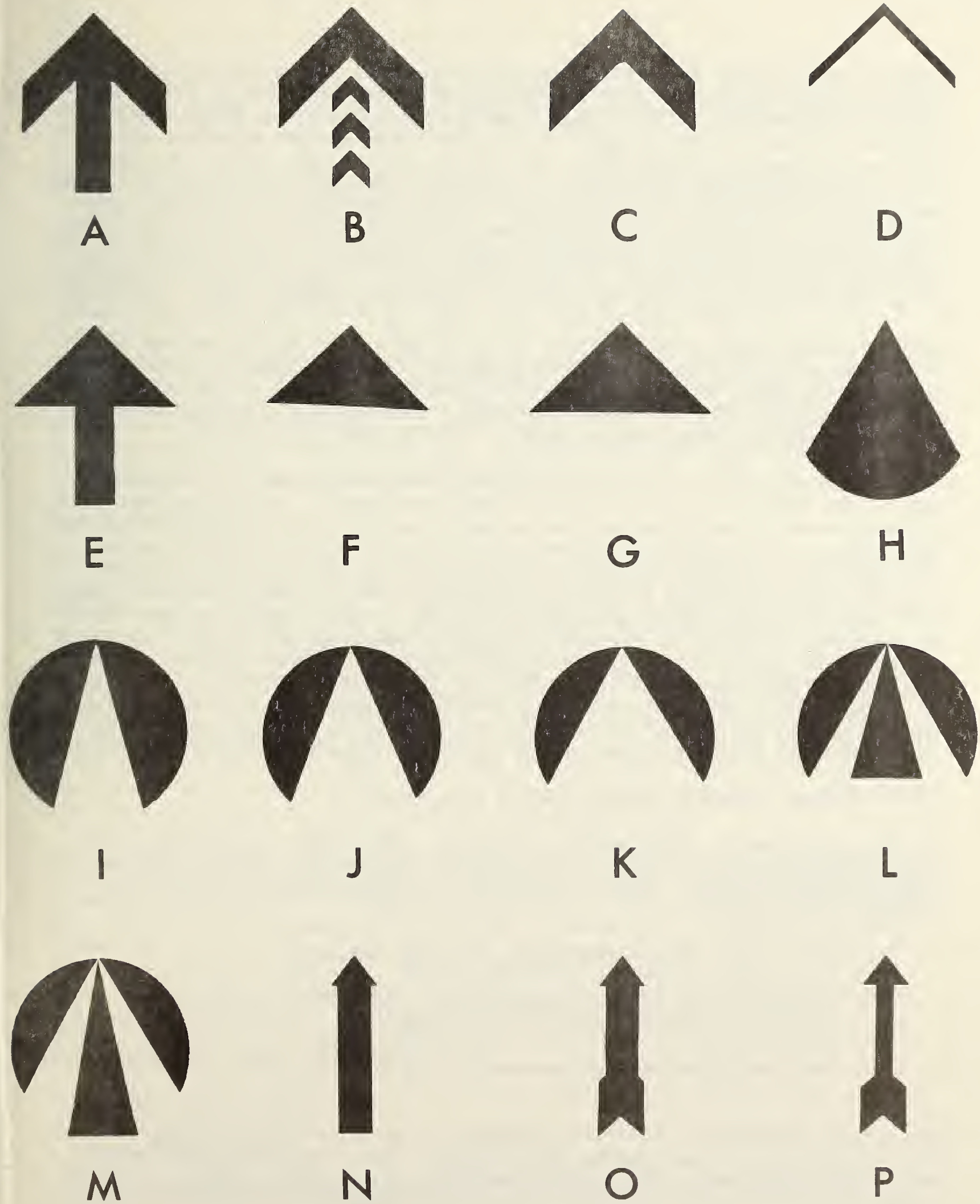


FIGURE 2. Sixteen arrow shapes used in the visibility experiment (white on black versions not shown).

Stimulus presentation was controlled by a two-channel projection tachistoscope system. One projector was used to present the arrow slides, while the other projector presented a veiling light. The two rear-projection images could be optically combined at the screen. The veiling light was projected through a 4.0 log unit neutral density filter and was on continuously during the experiment. This served two functions: a) it reduced contrast between figure (arrow) and ground when combined with the projected arrow slide; b) it provided a dim uniform visual field for illuminating the room. The only other source of room illumination was a shaded 4-watt night-light, located 1.2 m (4 ft) behind the observer, 1.4 m (55 in) above the floor. These light sources produced an ambient illumination level, measured at the seat of the participant's chair, of 0.21 lx.

Neutral density filters were affixed to the front of the shutter through which the arrows were projected. Thus, the observer viewed an image that was first projected through a neutral density filter and then combined at the screen with the veiling light from the second projector. Three levels of filtering of the stimulus projector image were included: 4.6, 5.0, and 5.1 log units.

Luminance measurements of the projected image were made using a 1° luminance meter. The night-light was on during these measurements. Due to the fact that the brightness of the screen was not perfectly uniform, the luminance values are approximate. The luminance from the veiling light projector only, projected through the 4.0 log unit filter, was 0.68 cd/m^2. Since the luminance of the veiling light was roughly an order of magnitude greater than the stimulus image, this value did not change substantially when the stimulus image was added to it. Figure-ground contrast for the arrow slides was measured with the veiling light projector off and no filter on the stimulus projector (night light on). The luminance of black areas was in the range of 50 to 140 cd/m^2, with the luminance of white areas in the range of 4600 to 5100 cd/m^2 (variability is due to minor nonunformity of the screen and slides). Thus the ratio of the luminances of white and black elements of the slides was in the range of 100-to-1 to 100-to-3.

3.1.4 Stimulus Sequences and Visibility Conditions

Sixteen sequences of slides were constructed, each sequence consisting of the 32 different arrows. The order of the arrows and the direction in which each arrow pointed--up, down, left, or right--were randomly determined, with the following constraints:

a) Each direction occurred eight times per sequence.

b) For each block of four sequences (1-4, 4-8, 9-12, 13-16), each direction occurred once for each arrow.

c) For each block of four sequences, each arrow occurred once in each each quarter of the list.

For actual presentation during the experiment, the series were shown in pairs, e.g., series one and two, series three and four, etc. Thus there were eight sets of 64 slides each. During the experiment a participant viewed a total of 18 such sets, so that each particular set was viewed two or three times. The sequence in which the sets were presented is shown in appendix A. As the appendix indicates, three sets (excluding practice) were viewed on days one and five, and four sets were viewed on days two, three, and four.

Three different visibility conditions were included in the experiment. The visibility condition was determined by the value of the neutral density filter through which the arrow was projected: 4.6, 5.0, or 5.1 log units. A given set of 64 slides was viewed under a given visibility condition. Since there were 18 sets of slides viewed, each visibility condition occurred six times per participant. The sequence of visibility conditions was different for each person, and was determined as follows. Since there are factorial 3 (3!) = 6 different orderings of three conditions, each participant received each of the six orderings of visibility conditions, thus completing the 18 sets. The six sequence of the six orderings was randomized for each participant, with the following constraints: a) each visibility condition occurred at least once within the three or four sets presented each day; and b) the same condition did not occur twice in a row within a day. Appendix A presents the sequence of visibility conditions for each participant.

3.1.5 Procedure

Each session began with a five-minute visual adaptation period. Following this, the participant viewed the stimuli while seated in the chair and recorded responses on an answer form. A stimulus slide was presented for 3 s, followed by a 5 s period in which the participant recorded the direction--up, down, left, or right--in which she believed the arrow pointed. A tone indicated the duration of the answer period; when the tone went off, the next slide was shown. An entire series of 64 slides was shown in sequence, all under the same visibility condition. This was followed by another set of 64 slides under a different visibility condition. Except for the first and last days of the experiment, a session consisted of four sets of 64 slides each. The slides were shown in the same sequence for each participant, but the visibility condition in effect for each list varied between participants, as discussed above.

In the first session, the participant read the instructions and was given a page containing an example of each arrow (appendix B). The instructions emphasized the importance of recording an answer for every slide, even when guessing was required. After familiarization with the arrow set, a practice list was presented. This differed from the final procedure in that only 32 slides were presented, and the slide image was not projected through a filter. Following this, the practice list was again presented, but this time projected through the 4.6 log unit filter. Subsequently actual data collection was begun. Six sets of 64 slides were viewed under each of the three visibility conditions during the five days of data collection.

3.2 VISIBILITY RESULTS

3.2.1 Practice Trial

None of the participants made any errors on the practice list when no filter was used to degrade the stimulus image. This indicates that all 16 directional indicators, with both positive and negative contrast, were interpreted correctly. Thus, errors observed under reduced visibility may be attributed to the visibility condition and are not confounded with problems of interpreting the arrow.

3.2.2 General Performance

The effectiveness of the filtering variable in manipulating visibility, and the stability of performance over the course of the experiment, were assessed by an analysis of variance. The mean number of errors per set of 64 slides was the dependent variable. The days-by-visibility repeated measures analysis (Keppel, 1973) is summarized in table 1 (this and all subsequent statistical tests were conducted at the $\alpha = 0.05$ confidence level). The visibility variable was highly significant, $F(2,12) = 37.04$, $p < 0.001$, and accounted for more than half of the variance among the scores, as estimated by ω^2 ($\omega^2 = 0.55$). The mean number of errors varied from only 1.24 per set (two percent) under the most visible condition, up to 21.69 per set (34 percent) under the least visible condition. Pairwise comparisons (Newman-Keuls test, Winer, 1971) among the three visibility conditions indicated that they all differed significantly from one another ($p < 0.01$ in each case). Performance was stable during the five days; neither the main effect of days nor the interaction of days with visibility were statistically significant factors.

Table 1. Summary of Days-by-Visibility Analysis

Source	Sum of Square	df	Mean Square	F
Subjects (S)	2637.08	6	439.51	
Visibility (V)	7334.98	2	3667.49	37.04 ($p<0.001$)
(V)x(S)	1188.21	12	99.02	
Days (D)	144.86	4	36.22	1.72 (not significant)
(D)x(S)	504.54	24	21.02	
Visibility x Days	103.91	8	12.99	1.03 (not significant)
(V)x(D)x(S)	605.81	48	12.62	
	12519.34	104		

3.2.3 Effect of Stimulus Variables

The percentage of incorrect answers for each arrow, under each visibility condition, is presented in table 2. Each percentage is based on 84 observations (6 occurrences of each visibility condition x 2 stimulus lists per occurrence x 7 participants). As the table indicates, only arrow N showed a substantial rate of errors (about 20-26 percent) under the most visible viewing condition (4.6 filter), with some difficulty for arrow P also indicated (about 3.6-4.8 percent errors). Under the more difficult viewing conditions, a wide range of differences between arrows was observed: from 2.4 percent to 56.0 percent errors with the 5.0 log unit filter, and from 8.3 percent to 65.5 percent errors with the 5.1 log unit filter. To assess these data, a visibility-by-arrow-by-direction of figure-ground contrast repeated-measures analysis of variance (Keppel, 1973) was performed, and the results are summarized in table 3. The analysis included only two of the three levels of the visibility factor. The 4.6 log unit filter condition was omitted since its inclusion would (a) badly violate statistical assumptions of homogeneity of variance and normal distribution (since there were almost no errors); and (b) the high error rate for only a single arrow (N) could in itself produce an arrow-by-visibility interaction, and thus could mask the possible existence of such an interaction between the other two visibility conditions.

As table 3 indicates, the main effects of visibility and arrow were highly significant [$F(1,6) = 89.10$; $F(15,90) = 26.91$; p's < 0.001], accounting for approximately nine and 35 percent of the variance ($\omega^2 = 0.091$ and 0.350, respectively). The main effect of the direction of contrast was not significant, but this factor did interact with the arrow factor [$F(15,90) = 4.43$; $p < 0.001$]; that is, the effect of the direction of contrast depended on the particular arrow. The magnitude of this interaction was small, however ($\omega^2 = 0.023$).

3.2.4 Type of Error

Incorrect responses were broken into two categories: those in which the participant's response differed from the correct answer by 90°, and those in which the response differed from the correct answer by 180°. If incorrect answers represented random guessing, the expected number of 90° errors would be double the number of 180° errors. However the data differed substantially from this expectation. Table 2 presents the ratio of 90° errors to 180° errors for each arrow, under each visibility condition (numbers in parentheses indicate that the ratio is based on fewer than ten errors).

For comparing the type of errors, the data for each arrow were collapsed over the visibility conditions and direction of contrast. For three arrows-- N, O, and P -- the majority of errors are at 180°. These are the arrows with predominant shafts, so that the orientation of the arrow (horizontal or vertical) may be detectable even when the head is difficult to distinguish. More surprising is the occurrence of 90° errors in greater than a two-to-one frequency for the remaining group of arrows. The actual ratio of 90°-to-180° errors for these 13 arrows taken together was better than three-to-one. Only three arrows (E, H, I) showed ratios of about two-to-one; for the remaining ten arrows the ratios ranged from 2.6 to 8.0 (see table 2). All seven participants showed greater than two-to-one errors at 90°, averaged over the 13 arrows (range of 2.5 to 5.0). Thus there was a substantial tendency for participants to make 90° errors (excluding the data from arrows N, O, P).

To be certain that no special guessing strategy or response bias could account for the pattern of errors, responses to three different arrows (A,F,K) were analyzed in detail. The analysis included both directions of contrast (black on white, white on black), but was confined to the 5.1 log unit filter condition.

Table 2. Error Rate and Ratio of 90° to 180° Errors for Each Arrow Under Each Experimental Condition

		4.6 Log Filter		5.0 Log Filter		5.1 Log Filter	
Contrast	Arrow	% Errors	90°/180°	% Errors	90°/180°	% Errors	90°/180°
Black on White	A	0	-	8.33	(2.50)	16.67	3.67
	B	1.19	(1.00)*	13.10	2.67	21.43	3.50
	C	0	-	2.38	(∞)	8.33	(6.00)
	D	2.38	(0.50)	45.78	5.33	57.14	2.83
	E	0	-	28.57	5.00	39.29	1.54
	F	0	-	21.43	17.00	40.48	4.67
	G	1.19	(1.00)	22.62	∞	41.67	4.83
	H	0	-	14.29	2.00	32.14	3.50
	I	0	-	17.86	0.88	36.90	1.58
	J	0	-	4.76	(1.00)	23.81	1.22
	K	0	-	11.90	2.33	17.86	4.00
	L	0	-	13.10	1.20	22.62	2.17
	M	0	-	10.71	(3.50)	15.48	1.60
	N	20.24	0	45.24	0.15	65.48	0.83
	O	2.38	(0)	39.29	0.32	57.83	0.81
	P	4.76	(0.33)	27.38	0.64	52.38	1.10
White on Black	A	0	-	11.90	4.00	23.81	3.00
	B	0	-	8.33	(6.00)	21.43	5.00
	C	0	-	4.76	(3.00)	19.05	7.00
	D	1.19	-**	44.05	4.29	65.48	2.44
	E	0	-	5.95	(1.50)	26.19	1.44
	F	1.19	(1.00)	17.86	2.75	38.10	5.40
	G	0	-	10.71	(∞)	20.24	4.67
	H	0	-	9.52	(1.67)	30.12	1.50
	I	0	-	15.48	0.86	35.71	5.00
	J	0	-	23.81	9.00	36.90	2.88
	K	0	-	5.95	(∞)	28.57	3.80
	L	0	-	5.95	(∞)	13.10	10.00
	M	0	-	2.38	(∞)	10.71	(3.50)
	N	26.19	0.05	55.95	0.81	60.71	0.65
	O	1.19	(0)	33.33	0.42	52.38	0.38
	P	3.57	(0)	30.95	0.62	47.62	0.90

* parentheses indicate ratio is based on fewer than ten errors
** error of omission - answer was skipped

Table 3. Summary of Visibility-by-Arrow-by-Direction of Contrast Analysis

Source	Sum of Squares	df	Mean Square	F
Subjects (S)	829.69	6	138.28	
Visibility (V)	341.26	1	341.26	89.10 ($p<0.001$)
V x S	22.98	6	3.83	
Contrast (C)	5.36	1	5.36	2.52 (not significant)
C x S	12.75	6	2.12	
Arrow (A)	1340.23	15	89.35	26.91 ($p<0.001$)
A x S	298.60	90	3.32	
Visibility x Contrast	0.64	1	0.64	(<1.0) (not significant)
V x C x S	10.03	6	1.67	
Visibility x Arrow	29.78	15	1.98	1.24 (not significant)
V x A x S	144.23	90	1.60	
Contrast x Arrow	108.11	15	7.21	4.43 ($p<0.001$)
C x A x S	146.53	90	1.63	
Visibility x Contrast x Arrow	25.25	15	1.68	(<1.0) (not significant)
V x C x A x S	234.33	90	2.60	
	3549.77	447		

Table 4. Percentage of Responses in Each Direction for Each Direction of the Stimulus Arrow

		Response				
		Λ	V	>	<	(N)
Stimulus	Λ	73.3	5.5	11.0	10.3	(146)
	V	1.6	75.2	11.2	12.0	(125)
	>	12.0	11.1	69.2	7.7	(117)
	<	8.6	12.1	7.8	71.6	(116)

A summary of these data are presented in table 4 in the form of a response matrix: for each direction the actual stimulus arrow pointed, the percentage of each direction of response is given. The data are grouped for all three arrows, and represent a total of 504 observations.

The percentage of correct responses was very similar (72.3 percent $\pm$ 3 percent) regardless of the direction the arrow actually pointed. Ninety degree errors predominated (four-to-one over all), and the number of errors at +90° and -90° was quite similar in every case. A chi-square test found no significant change in the distribution of responses among the categories "correct", "90° error," and "180° error" as a function of the direction of the stimulus arrow ($\chi^2 = 6.1$, df = 6). There was however some suggestion of more 90° errors for the vertical arrows, although the ratio of 90° to 180° errors exceeded two-to-one in every case. There was no apparent response bias, each response occuring with a similar frequency as an error. Thus, for the group of seven participants (although not necessarily for individuals) there was no substantial response bias and no apparent pattern of responding that could account for the preponderance of 90° errors.

4. RANKING EXPERIMENT

4.1 METHOD

The ranking experiment provided further information on the suitability of a set of arrows (taken from the visibility experiment) for use as directional indicators for exits. Each arrow was ranked on several subjective dimensions relevant to building exit signage, which are described below. The results of this experiment, taken together with the visibility findings, can help identify effective arrows, or arrow features, for exit applications.

Thirteen arrow shapes were tested. Only the black figures on white backgrounds were used, with arrows F, J, and L omitted (to avoid highly similar stimuli). Each arrow was presented on a separate 21 cm x 28 cm (8 1/2 in x 11 in) sheet. The image was a same-size photocopy of the original 15.2 cm x 15.2 cm arrow placard described under the visibility experiment method. Each participant received a packet of 13 sheets, one sheet for each arrow tested.

Participants were 52 female students in upper level undergraduate psychology courses at Hood College, in Frederick, Maryland.

Each arrow was rank-ordered by participants according to one of four criteria: 1) how well the arrow conveyed the idea of "exit"; 2) how well the arrow served to indicate direction; 3) how unique the arrow appeared, relative to ordinarily encountered directional arrows; and 4) how attractive and appropriate the arrow would look as a sign in a building. (Exact instructions for ranking are attached in appendix C.)

Each participant rank-ordered the arrows according to only one criterion (randomly assigned). A total of 12 to 14 participants was included for each criterion. Each participant was given a packet of 13 randomly ordered sheets, with a different arrow on each sheet. The participants then reordered the sheets according to the required criterion.

4.2 RESULTS

A rank of one was assigned to the arrow a participant judged to possess most the quality being rated, and a rank of 13 was assigned to the arrow judged to least possess that quality. Table 5 presents the mean rank given each arrow by the participants in each of the four groups. These group means are based on a highly significant degree of agreement between the participants within each group. Table 6 (first numerical column) indicates the degree of agreement between participants expressed by the coefficient of concordance (W). For all four stimulus criteria, W is significant (χ^2 approximation, Hays, 1963, pg. 658) beyond the $\alpha = 0.001$ level.

Given significant agreement among participants, pairwise comparisons of group rankings of individual arrows for a given dimension were made. A multiple comparison procedure (Hollander and Wolfe, 1973, pg. 151) was used to protect against spurious differences; however, given the large number of pairwise comparisons (78 for each dimension), such a test is necessarily relatively

Table 5. Ranking of Arrows for Four Different Criteria

	EXIT		DIRECTION		APPEARANCE		UNIQUENESS	
	Mean of Individual Rankings	Rank of Group Mean	Mean of Individual Rankings	Rank of Group Means	Mean of Individual Rankings	Rank of Group Means	Mean of Individual Rankings	Rank of Group Means
Arrow								
A	2.71	1	3.15	4	4.08	3	10.42	10.5
B	3.57	3.5	5.77	6	3.69	1.5	5.25	5
C	7.36	7	7.08	7	6.38	6	6.25	6
D	8.36	8	7.92	8	7.00	7	6.42	7
E	3.50	2	3.08	3	4.46	4	10.92	12
G	8.43	9	10.23	10	9.31	10	6.50	8
H	11.21	12	11.31	12.5	11.38	13	4.00	3
I	10.93	11	10.15	9	9.62	11	4.08	4
K	11.50	13	10.77	11	10.38	12	3.83	2
M	10.71	10	11.31	12.5	8.38	9	1.50	1
N	5.43	6	4.46	5	7.31	8	11.08	13
O	3.71	5	2.92	2	5.31	5	10.33	9
P	3.57	3.5	2.85	1	3.69	1.5	10.42	10.5

Table 6. Rater Agreement for Arrow Rankings

Criterion	Concordance (W)	χ^2	
Exit	0.76	127.85	$p<0.001$
Direction	0.79	123.86	$p<0.001$
Uniqueness	0.71	101.66	$p<0.001$
Appearance	0.47	72.70	$p<0.001$

insensitive. The results of the pairwise comparisons are presented in table 7. As the table indicates, many of the comparisons were significant, despite the conservative test. Furthermore, there was good agreement among the four criteria as to which pairs of arrows differed significantly from one another.

The degree of agreement between the group rankings of each stimulus criterion is shown as a rank-order correlation in table 8. The rank for each arrow, from first through thirteenth, was ordered by the group mean rank of each arrow (table 5). The rank correlation for each pair of criteria is presented in table 8. All of these correlations are significant beyond the $\alpha = 0.05$ level. The "Exit", "Direction", and "Appearance" dimensions show a strong positive correlation with one another (rho's from + 0.84 to + 0.93); all three show a negative correlation with "Uniqueness" (rho's from − 0.57 to −0.82). Thus an arrow ranked high on the "Exit" dimension would also tend to rank high on the "Direction" and "Appearance" dimensions and low on the "Uniqueness" dimension. Figure 3 shows these relationships as scatterplots, with each arrow indicated by the letter corresponding with figure 2.

Table 7. Pairwise Comparison of Arrow Rankings

Arrow		A	B	C	D	E	G	H	I	K	M	N	O	P
A	Exit			*	**		**	**	**	**	**			
	Direction				*		**	**	**	**	**			
	Appearance						**	**	**	**				
	Uniqueness		*					**	**	**	**			
B	Exit				*		*	**	**	**	**			
	Direction							**		*	**			
	Appearance						**	**	**	**				
	Uniqueness					**						**	*	*
C	Exit													
	Direction													
	Appearance							*						
	Uniqueness													
D	Exit					*							*	*
	Direction					*							*	**
	Appearance													
	Uniqueness										*			
E	Exit						**	**	**	**	**			
	Direction						**	**	**	**	**			
	Appearance						*	**	**	**				
	Uniqueness							**	**	**	**			
G	Exit												*	*
	Direction											**	**	**
	Appearance													**
	Uniqueness										*			
H	Exit											**	**	**
	Direction											**	**	**
	Appearance												**	**
	Uniqueness											**	**	**
I	Exit											**	**	**
	Direction											**	**	**
	Appearance													**
	Uniqueness											**	**	**
K	Exit											**	**	**
	Direction											**	**	**
	Appearance												**	**
	Uniqueness											**	**	**
M	Exit											**	**	**
	Direction											**	**	**
	Appearance													
	Uniqueness											**	**	**
N	Exit													
	Direction													
	Appearance													
	Uniqueness													
O	Exit													
	Direction													
	Appearance													
	Uniqueness													
P	Exit													
	Direction													
	Appearance													
	Uniqueness													

* $p<0.10$

** $p<0.05$

Table 8. Rank Correlations (rho) Between Arrow Rating Criteria

	Exit	Direction	Uniqueness	Appearance
Exit	--	+0.87***	-0.76**	+0.93***
Direction	--	--	-0.82***	+0.84***
Uniqueness	--	--	--	-0.57*
Appearance	--	--	--	--

* $p<0.05$

** $p<0.01$

*** $p<0.001$

Table 9. Composite Ranking of Arrows

	Visibility	Uniqueness	Exit Direction Connotation	Means of Ranks	Rank of Means
A	3	10.5	2.5	5.33	4
B	4.5	5	5	4.83	3
C	1	6	7	4.67	1.5
D	12	7	8	9.00	12
E	8	12	2.5	7.50	8
G	7	8	9	8.00	10.5
H	6	3	13	7.33	7
I	9	4	10	7.67	9
K	4.5	2	12	6.17	5
M	2	1	11	4.67	1.5
N	13	13	6	10.67	13
O	11	9	4	8.00	10.5
P	10	10.5	1	7.17	6

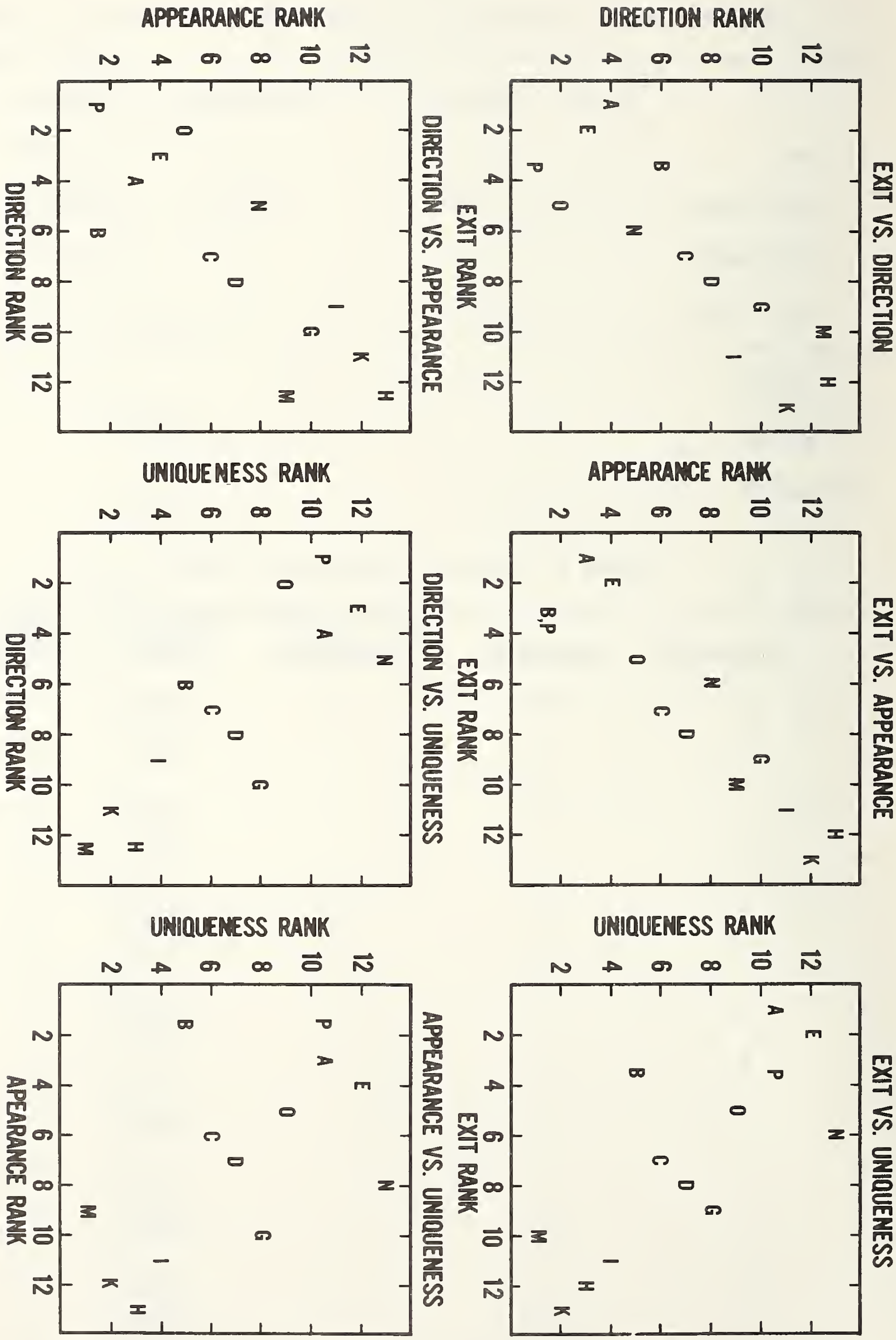


Figure 3. Pair-wise comparisons of arrow rankings for four different criteria

5. DISCUSSION

5.1 EVALUATION OF VISIBILITY METHODOLOGY

The method used to manipulate visibility appeared to be quite successful. Performance was related in a highly significant ($p<0.001$) manner to the visibility manipulation, with the visibility condition variable accounting for over half ($\omega^2 = 0.55$) of the variability in the data (Visibility-by-Days analysis). The method discriminated well between arrows ($p<0.001$), with the percentage of correct responses per arrow varying widely (44 to 92 percent in the most difficult viewing condition). This sensitivity was achieved despite the relatively small number of participants and the large set of stimuli among which the observations were divided. Furthermore, performance remained stable, with no significant effect of days and no interaction of days with visibility condition. Thus, the method appears to be a relatively sensitive and efficient means of evaluating the relative visibility of arrows under different viewing conditions, and the values of the visual parameters employed were effective.

This procedure differs from most experiments on symbol visibility in aspects of both stimulus presentation and the participant's response. These methodological issues are further discussed in section 5.3.

5.2 VISIBILITY OF THE ARROWS TESTED

The arrows differed widely in visibility, with D, N, O, and P consistently poor, and with C and M most visible overall. A, B, K, and L were also among the better-seen arrows. The arrows differed not only in how frequently they were incorrectly responded to, but also in the types of errors that were made. Three of the 16 arrows (N, O, P) showed predominantly 180° confusions; in contrast, ten arrows showed 90° confusions in greater than a two-to-one ratio to 180° errors.

The predominance of 90° errors for ten of the arrows is an interesting finding. Close analysis did not suggest that this result was related to any overt response bias or obvious guessing strategies. One possible explanation is that such errors may be made when a participant detects an arrow feature that is 90° out of phase with the actual arrow direction. In fact, arrows with a dominant perpendicular element and points (F, G) did show a strong bias toward 90° errors. However, other stimuli also showing this error bias have no perpendicular elements at all (e.g., B, C, J, K). Another possibility is that the errors may be related to search strategies used by the observer. As an example, an observer may try to match the image on the screen to a "search image" oriented in a particular way; that is, to search for an arrow pointing either up, down, or to one side. If failing to match the search image were to differentially favor either detection of stimuli 180° out of phase or evoke answers 90° out of phase, it could establish the preponderance of 90° answers. Quite possibly, graphic features and perceptual strategies both contribute in some way to the distribution of error types. While no satisfactory account of the error data is obvious, the finding of predominantly 90° errors may be worthy of further attention. The non-random distribution of the errors indicates that even when a participant answered incorrectly, some information about the stimulus must have been obtained. Uncovering the source of the error bias may help reveal how people search for signage in poor visibility or what graphic features reduce symbol effectiveness.

In comparing the various arrows, the effect of the arrow shaft as a graphic feature is interesting. For three different arrow types, a comparison of the arrow with and without a shaft is possible (A and B vs. C; E vs. F; L and M vs. K); in addition, three arrows (N, O, P) have a predominant shaft, with only shaft width varying between two of them (O and P). The shaft appears to increase the number of 180° errors relative to 90° errors, but whether it serves to increase or reduce the overall number of errors depends upon the particular arrow. When the head is made more predominant relative to the shaft (P vs. O, E vs. N), performance improves. Removing the shaft altogether actually increased the percentage correct for C (vs. A and B), but not for F (vs. E) or M (or L) (vs. K). One interpretation of these effects is that the shaft may serve to improve performance by reducing dominant perpendicular cues (strong vertical edge in E) and by reinforcing the directional message (as in L and M), but can also interfere by contributing conflicting cues, 180° out of phase with the arrow head. The positive effect of the shaft may have been somewhat obscured by a possible experimental artifact. If most of the information is contained in the head, then since each arrow was centered on the background, for some of the arrow heads without shafts (C, F, G), the most informative portion of the figure was centered in the display. It is not possible to assess the effect, if any, of this factor. The relatively long display time (3s) is ample to allow search of the entire display area, and so should help reduce this possibility. However, it is interesting to note that most visibility experiments do center the images and use very brief display times (see section 5.3.1), so that a bias may exist in much of these data. An optimal procedure may be to vary the location of the stimulus within the display field.

In any case, the data show that for some arrows the shaft can contribute substantially to improved performance, particularly for arrow M. This may have been an especially effective case because the sides of the shaft itself are not perpendicular to the intended direction and the tapering shaft results in multiple points at the vertex. Although the various comparisons among the arrows with different shaft types suggest a number of hypotheses, perhaps the main conclusion is that the arrow head should be the dominant graphic element. The shaft may be useful in reinforcing the message, but should be a secondary feature. Interestingly, the uniqueness rankings provided by participants, as well as a review of sign catalogues, suggests that arrows with predominant shafts (e.g., N, O, P) have actually been the most frequently used. Nonetheless, arrows more similar to A are now used and recommended for building signage (Follis and Hammer, 1979), and more recent guides and standards for informational signs have favored arrows more similar to A (e.g., ISO, 1979; Veteran's Administration, 1980).

The direction of figure-ground contrast had relatively little effect on visibility. There was no overall advantage to either black on white or white on black. Although the interaction of the direction of contrast with the particular arrow type was significant, the difference in percentage correct between the two contrast conditions was small for most arrows (see table 2). No obvious features appear to distinguish the arrows that seem to show a direction of contrast effect, and of course, some of the apparent differences could be fortuitous. In fact, the literature on figure-ground relationship shows generally small effects of the direction of contrast, and any effect may be highly specific to graphic, illumination, or methodological details (Konz and Mohan, 1972). In practice, the direction of contrast employed may also need to take into account factors in addition to visibility, such as compatability with other signage and contrast with building elements.

5.3 COMPARISON WITH OTHER SYMBOL VISIBILITY RESEARCH

5.3.1 Visibility Manipulation

The visibility of a symbol has been manipulated in a variety of ways in laboratory research. These have included manipulations of contrast (Smith and Weir, 1978), size (or correspondingly, distance) (Mackett-Stout and Dewar, in press), blur (Smith and Weir, 1978), and viewing time (Markowitz, Dietrich, Lees, and Farman, 1968). The viewing conditions and the means of manipulating visibility are not necessarily arbitrary or interchangable, since the relative performance of different symbols can vary with these factors. Several experiments in which symbols or arrows were compared under multiple viewing conditions indicate this. For example, Mackett-Stout and Dewar (in press) compared 32 public information symbols for both "glance legibility" (stimulus duration varied) and "distance visibility" (viewing distance varied). The authors reported a correlation between these two measures of 0.25, which only "approached" statistical significance. Smith and Weir (1978) compared the visibility of eight directional arrows in two experiments: one in which blur was varied (to simulate differences in acuity) and one in which contrast was varied (related to glare). The rank-order correlation for the arrows between the two experiments was 0.14. Jacobs, Johnston, and Cole (1975) varied blur in comparing a set of 16 symbols for traffic signs (they also included word signs). The rank-order correlation between conditions of no blur and the most severe blur was 0.58. Jin (Japanese proposal, 1980) compared the visibility of 13 exit symbols under both "normal" illumination and in smoke. The rank-order correlation for the set of symbols in the two conditions was 0.04. To some extent the low correlations observed in all these experiments may be attributed to possible experimental artifacts or to large variability in the measures relative to the differences between symbols. Nonetheless, it is obvious that a reasonable simulation of actual conditions of interest is desirable, since generalization from one condition to another may not be warranted. With the exception of Jin's work, the experiments on symbol and arrow visibility did not use procedures most appropriate for evaluation of building emergency signs. However, many of the experiments were explicitly directed at the highway driving situation.

Two parameters of viewing conditions should be addressed in particular. One is the temporal conditions of viewing. In the present experiment, participants had ample viewing time. However, duration is an easily manipulated experimental variable and has often been used to degrade performance in visibility experiments (e.g., King, 1971, 1975; Dewar, 1976). Sometimes the duration may be only milliseconds (e.g. 15 to 30 ms, Markowitz et al, 1968); in other cases duration may be longer, but still brief enough to stress the viewer under difficult viewing conditions (e.g. 0.5 s, Smith and Weir, 1978). Such brief times may be related to "glance legibility" for the highway driving situation. Brief durations may also be useful in manipulating task demand and stress, and can be used to study the visual information processing chain in sign perception. However, such brief presentation intervals are not necessarily appropriate for simulating viewing conditions in a building environment, and the validity of extending these observations to building applications is not demonstrated. The poor correlations between measurements made under very different viewing conditions (including "glance" vs. "distance" viewing) demands caution in generalizing results obtained with brief viewing times. The three-second stimulus duration of the present experiment was not demanding, with participants typically recording their responses before the viewing time was even complete.

The other parameter of special note is viewing distance (or correspondingly, image size). Several studies have determined visibility by obtaining the distance at which symbols are recognized under "good" or "normal" viewing conditions; a number of these experiments specifically addressed symbols for building use. While the authors typically emphasized the differences between symbols, perhaps the overlooked finding is that visibility under good viewing conditions may not be a critical concern. Virtually all the symbols tested may be adequate under these conditions. The results of the experiments on building symbols were transformed, using the relationship tan visual angle = size/distance, to determine at what distance a 12 in (30.5 cm) sign would be recognized. Jin's (Japanese proposal, 1980) results for exit signs indicate distances of about 48 to 92 m (about 150 to 300 ft). Mackett-Stout and Dewar's (in press) data for public information signs indicate distances of about 60 to 150 m (about 200 to 500 ft). Zwaga's (1979) recognition data for public information symbols indicate a distance of at least 105 m (about 345 ft). These distances are all large relative to most building spaces, especially given code requirements for distances to exits. For example, Sharry (1978, p. 111) has summarized the maximum permissible distances to exits (for new construction) prescribed in the Life Safety Code. For public and industrial settings, these range from 75 to 150 ft (22.9 to 45.7 m) for various unsprinklered occupancies, and from 75 to 200 ft (22.9 to 61.0 m) for sprinklered occupancies (greater distances are permitted in certain low risk storage and business occupancies). Thus symbols generally appear adequately visible under ideal viewing conditions. But given the often poor correspondence between relative visibility measures obtained under different viewing conditions, it may be unwarranted to generalize the relative performance under normal building conditions to poor visibility conditions. The present experiment therefore attempted to use a realistic visual angle for the stimulus, and visual conditions of true concern. The subtended visual angle of 1.3° is equivalent to viewing a 12 in sign at 13.4 m (44 ft), while stimulus presentation simulated emergency lighting and smoke.

5.3.2 Participant's Response

Another methodological factor in visibility research is the task which the research participant must perform. There have been three general categories of response: (1) for directional indicators, the participant indicates the appropriate direction; (2) for other symbols, the participant indicates which of a specified set of stimuli was presented ("matching" or "recognition"); or (3) for unfamiliar or unspecified symbols, the participant describes or interprets what was presented. The particular measure based on the response is most often proportion correct, although reaction time of the response, or distance at which the response occurs, have also been used. Although there has been little formal comparison of these methods, the participant's task can be an important factor in symbol visibility research.

This experiment used arrows as stimuli not only because of an interest in specific arrows, but also because of the methodological advantages inherent in using a directional response (section 2.0). There are no "partially correct" or ambiguous answers, and performance for any given arrow is not strongly influenced by which particular arrows comprise the test set. The procedure appears useful not only for explicitly comparing directional indicators, but also for assessing a number of graphic and display variables of concern for any signage under emergency viewing conditions. These would include such factors as the type and level of symbol illumination, color, contrast, size, graphic features, symbol surround, location, clutter from competing visual elements, and so on.

In directly comparing various non-directional symbols, such as exit designators, for visibility, the directional response used with arrows is not possible. The selection of an appropriate response measure has been a significant problem in the assessment of symbol visibility. Typically a small set of symbols is used, and the response is to indicate which member of this known set has been presented. This is often done by matching from among a list of of alternative choices.

A primary problem with this procedure is that it may be described as a discrimination experiment, in which the participant must decide which of several known images is presented. Thus the findings may reflect how well a set of symbols can be discriminated from one another, rather than how well any individual symbol can be identified when encountered in the environment. To determine which stimulus of a specific set was presented, the participant may rely on minor irrelevant cues which effectively distinguish one symbol in the set from another. These may be unrelated to the features which are actually used to recognize a symbol encountered in a real environment. For example, Zwaga (1979) examined the legibility of public information symbols proposed by ISO. He compared the results obtained by two methods. In one procedure, participants were made familiar with the set of symbols, and on each trial indicated which symbol of this set was presented. In the other procedure, participants were unfamiliar with what they were to see, and the task was to describe what was presented. Very different results were obtained with the two methods. The same symbol was not necessarily best under both procedures, and the procedure in which participants were familiar with the symbols consistently yielded much higher visibility estimates. Perhaps most critically, Zwaga reported that for "familiar" participants, legibility was related to the uniqueness of subsets of symbol features. Thus the flaw in this general procedure is that "visibility" is strongly determined by exhaustive knowledge of all the symbols that might occur and the relationship between the specific symbols included.

Another possibility is to use a relatively large set of familiar stimuli. This may help reduce reliance on a sub-set of discrimative cues. It will not eliminate the problem however, and for specific stimuli, some feature might remain a reliable indicator. Furthermore, this method faces the practical problem of establishing, demonstrating, and maintaining the familiarity of the large set of stimulus alternatives.

As an alternative to having participants be familiar with the set of symbols, the symbols can be presented without prior exposure. There are a number of problems with this approach. One is simply the question of how the observer indicates what he saw and how to score what he reports. This method also confounds the visibility of a symbol with its meaningfulness, which may be a particular problem for proposed exit symbols since they differ widely in understandability. Another problem is that the symbols become more familiar as they are seen during the course of the experiment. Either each symbol may be seen by each participant only once, which may be an inefficient or impractical procedure; or familiarity is confounded with experience in the experiment, and the problem of discrimination on the basis of possibly minor cues again occurs.

Each of the possible procedures discussed must also face the issue of how well it approximates the real-life situation to which the findings may be applied. It would seem unlikely that a person would have good knowledge of the entire set of symbols that might be encountered. (This issue distinguishes research on symbols from the many experiments that have investigated legibility using alphabetic characters. Letters are in fact a set for which appropriate participants may be assumed to have exhaustive knowledge.) Using a totally unfamiliar set of symbols may not be appropriate either, except to simulate the first encounter a person may have with a particular symbol. This not only introduces the confounding of visibility with understandability, but also is unlike the more typical case in which some familiarity with the symbol under good visibility has preceded the poor-visibility encounter.

One procedure that appears to be practical and realistic does not seem to have been used to assess symbol visibility. This would be to familiarize participants with a selected group of target stimuli and then present those stimuli in the context of a large number of unfamiliar distractor stimuli. For example, to adapt this procedure for exit symbols, the participant would first become familiarized with a number of exit symbols. These would then be presented along with a large number of unspecified symbols, which may also include some additional (unfamiliar) exit symbols. The participant would respond in a yes-no manner, to indicate whether or not each stimulus presented was an exit symbol.

This procedure would provide an objective, scorable response. It is realistic in that the observer is familiar with possible exit symbols, but not with the range of possible symbols that might also occur. This would also make feature-by-feature comparison unadaptive, since the graphic details of the non-exit symbols are unknown. This means that some minor graphic detail, which might in fact make one stimulus especially discriminable from some sub-set of symbols, can not be identified and used to discriminate it from all potential symbols that could occur in an experiment. This procedure would also have the additional advantage of identifying symbols that might be mistakenly interpreted as exit designators.

5.3.3 Other Research Comparing Arrows

There have been a number of studies which have compared the visibility of various directional indicators. All of these have been addressed to highway applications rather than building signage.

The most similar research to the present one is that of Smith and Weir (1978). They compared a set of arrows in two experiments. In one, blur was used to degrade visibility. In the other, a veiling illumination was used to reduce contrast, as in the present experiment. However, certain experimental parameters were very different, due to the highway driving application. While luminance levels were not specified, they were undoubtedly much higher, as the veiling light was used to simulate glare from sunlight. Furthermore, the stimuli were very small (visual angle circumscribed 8.6 min of arc, about 1/9 the size of the present experiment) and were presented for only 0.5 s. Only black arrows on a white background were used.

The Smith and Weir results showed only moderate agreement between the blur and glare conditions. The correlation for the scores of the arrows under the two conditions was $r = 0.49$, with a rank-order correlation of $\rho = 0.14$. There was some overlap with the arrows used in the present experiment: these are arrows H and J and versions similar to A and E. A chevron of intermediate width to C and D was also included. Because only H and J were identical, it is difficult to know to what extent any differences in results were due to graphic differences. Furthermore, the results for many of Smith and Weir's arrows were not statistically discriminable, so that performance can only be meaningfully described for groupings of arrows. Despite this, one clear discrepancy emerges with the results of the present experiment. The arrow comparable to arrow H performed most poorly in both of Smith and Weir's experiments. This arrow was intermediate in the present experiment. The many procedural differences make this disagreement difficult to account for.

Markowitz et al (1968) examined seven arrows using a glance legibility procedure. The size and luminance of the stimuli were not reported, but presentation times were extremely brief: 15 to 30 msec. The arrows were all similar in having relatively long shafts. One finding to emerge from the experiment was much better performance for arrows in vertical, rather than horizontal, orientations. No such bias was found in the present experiment (sec. 3.2.4), and the possibility of some experimental artifact must be considered. One interesting finding was that the best performing arrow was one that had a tapering shaft. While very different graphically from arrows L and M of the present experiment (more like arrow E with a much longer, tapered shaft), the finding of good visibility in both studies for the arrows with tapered shafts suggests a possible area for further investigation.

Carr (1969) summarized British research on highway arrow visibility. These studies also used very brief (5 to 40 msec) presentation times. The main finding emphasized by Carr is that the visibility of one arrow relative to another is strongly dependent on the viewing conditions: "sunlight", "fog", or "night". Very different rank-orderings were obtained for 10 arrows under the three conditions. Carr also warned of possible differences when illuminated vs. non-illuminated signs are used. The arrow set used in these experiments is not readily compared to the present experiment.

Because there have been important differences in procedure and little overlap in the stimulus sets, comparisons between arrow visibility experiments are difficult. General recommendations or rules bearing on arrow visibility are tenuous. Perhaps the most general conclusion to emerge is again the importance of testing each visibility condition of concern in as appropriate a manner as possible.

Although the arrow visibility issue is complex, even definitive results would present only a partial answer to the selection of an exit directional indicator. The arrow should also contribute to the egress message as much as possible, and furthermore be compatible as an interior design element. The arrow ranking experiment addressed these issues. The following section discusses the ranking experiment. Subsequently, a composite evaluation of the arrows is presented, based on both the visibility and ranking results.

5.4 ARROW RANKINGS

Participants ranked the arrows in a meaningful way, as indicated by the substantial coefficients of concordance. Agreement was poorest on the "appearance" criterion, but for all four criteria it remained highly significant. Even with the test insensitivity engendered by the large number of pairwise comparisons, and the relatively small number of participants (12 to 14) ranking on a given dimension, many comparisons of pairs of arrows showed statistically significant differences. Thus the ranking procedure appears to be a simple and efficient means of collecting statistically sensitive data on subjective evaluations of these arrows.

The rankings showed a substantial amount of agreement between dimensions. Arrows most strongly suggesting exit also tended most strongly to indicate direction, and furthermore these tended to be considered the most attractive and appropriate. In contrast, the negative correlations with uniqueness rankings indicated that more familiar arrows were typically viewed as more attractive and better functionally.

Arrows rated most highly on the "exit", "direction", and "appearance" dimensions tended to be A, B, E, O, and P. Those rated as most unique were B, H, I, K, and M. Arrow B is the only one included in both sets, although not rated first on any dimension. The scatter plots involving the "Uniqueness" criterion (figure 3) confirm that B is the only arrow for which the data point is substantially below the cluster of negatively correlated points.

5.5 COMPARISON OF ARROWS FOR POTENTIAL USE WITH EXIT DESIGNATORS

In an attempt to provide an overall evaluation of the arrows tested as possible exit designators, the visibility and ranking data were combined. Each arrow was considered on three dimensions: visibility, uniqueness, and connotation of "direction to an exit" ("exit" and "direction" rankings combined). Table 9 presents this summarization, with each of 13 arrows ranked on the three qualities. The "visibility" rank is based on the mean number of errors for each arrow, collapsed across visibility conditions and direction of contrast. A rank of one indicates the fewest errors; a rank of 13 indicates the most errors. The "uniqueness" rank simply presents the ordering of the arrows, based upon mean rank for that criterion in the ranking experiment. For the "exit direction connotation," the mean ranks across participants for "exit" and "direction" rankings were combined; these combined totals were then ranked, so that a rank of one indicates the arrow that most suggests "exit" and "direction".

In the fourth column of table 9 the ranks are combined and averaged for the three qualities; these means are then themselves ranked in the far right column. This final column, then, provides a composite ranking for each arrow on a combination of desirable dimensions for an exit-related arrow: its visibility, its distinctiveness from other arrows, and its ability to suggest the idea of direction to exit. While the most appropriate weighting to give each factor may be an open question, this summation at least reflects each major factor.

As the final rankings indicate, the best-rated arrows are C, M, B and A. Thus, if one were restricted to the set of thirteen arrows tested, these four might be considered as the most reasonable candidates for a standard arrow to be used in conjunction with exit signage. Of these, A is less desirable because similar versions are included in recent standards for informational signage, and its use in applications other than egress will be counter productive. Arrow C has also been included as an acceptable directional arrow in a proposed ISO standard (ISO, 1977). Since arrow M received such a low rating in terms of its connotation, B may be the best alternative. B also rated highly in appearance, suggesting its esthetic acceptability. Thus arrow B emerged as a possible candidate should a standard exit arrow be desired. However, as emphasized earlier, the arrows in the set tested were chosen to provide a range of graphic types for methodological reasons. Many other candidate arrows might have been included. Arrow B may only be considered as the most appropriate among the set tested.

The success of arrow B does suggest one strategy in developing an exit-related arrow that is visible and compatible with other signage, yet has a unique quality. That is to modify the arrow shaft, which appears to contribute little to visibility if the arrow head is effective.

5.6 FURTHER CONSIDERATIONS IN TESTING OF PROPOSED EXIT SYMBOLS

To provide a sufficiently sensitive test of methodology with a relatively small group of participants, the present experiments employed a homogeneous set of observers, with normal or corrected-to-normal vision. If a set of exit designators is to be empirically evaluated for its understandability and visibility for the general public, broader and more extensive sampling is of course required. In addition to the need to be broadly representative, however, is the need to give specific attention to special user groups. These include the aged and the visually impaired, children, and the illiterate.

Symbol signs may be especially effective for these groups, since problems in seeing, reading, or interpreting written messages limit the usefulness of word signs. These special groups are also those who may most need the benefit of egress directions. Death rates from fire in public settings (as well as for all settings combined) are disproportionately high for children and those over sixty-five (National Safety Council, 1979). Yet several researchers (e.g., Easterby and Hakiel, 1977; Hulbert, Beers, and Fowler, 1979; Lerner and Collins, 1980) have reported poorer comprehension of symbols for the elderly (under good viewing conditions), and a number of visual defects are common in older vision which may alter sign perception (Fozard, Wolf, Bell, McFarland, and Podolsky, 1977; Koncelik, 1979). Children may also interpret symbols in a manner quite different from adults. If these special groups are ignored, or undifferentiated as a subset of the "general public," important life safety considerations in exit symbology may be overlooked.

Another special consideration in sign effectiveness concerns the role of alcohol and drug effects. A disproportionately high number of fire victims show significant levels of alcohol or drugs (Berl and Halpin, 1979). Although the need for unambiguous alerting may be critical for an intoxicated person, or even someone on medication, perception of neither word nor symbol signs appears to have been investigated for these conditions. Such research might contribute to life safety during building emergencies, and could be of obvious benefit as well for traffic symbols.

In summary, a complete evaluation of symbol effectiveness must not only consider who may encounter the sign, but also for emergency alerting, who the potential victims are likely to be.

Another aspect concerns the international usage of symbols. If an exit designator is to communicate to foreign visitors to this country, or is to be used internationally, the need to take cultural factors into account is apparent. Cross-cultural comparisons of symbols have been undertaken for some symbol sets (e.g., Easterby and Zwaga, 1976). However, the obvious difficulties of any full scale effort, especially if not under the auspices of some international group, have probably severely limited cross-cultural testing.

As a final consideration, there are a number of pragmatic concerns that must be taken into account in evaluating a symbol. These may include reproducibility, compliance with existing standards, compatibility with other signage, potential for vandalism, and so forth. However, understandability is paramount for any symbol, and visibility may be critical for exit designators. Thus, while other factors require attention, understandability and visibility should be the focus of research efforts. Related to these are questions of salience ("attention getting" quality), memorability, and behavioral effectiveness.

The symbol image is but a single consideration in exit signage, with other concerns including usage, location, size, illumination, and so on. Problems of signage enter the broader issue of rapid and safe egress from public buildings, transportation, and workplaces. The prevalence of the exit message and its critical importance to life safety stresses the importance of optimizing all components of the egress system. This paper has considered some of the problems in identifying and interpreting the exit message in symbol signs and has discussed some approaches to those problems. The research reported has concerned an important element of egress signage: the directional indicator to identify the appropriate direction to an exit. Specific issues in the use of directional arrows, as well as more general issues in evaluating exit-related symbols, were explored. Careful consideration of these issues in subsequent design, research, and standardization may contribute significantly to increased public safety.

6. SUMMARY

The increasing usage of pictorial signage and recent efforts to standardize such signs have underscored the need for an effective symbol to represent the "exit" idea. Exit symbols were discussed in this report in terms of two primary criteria: understandability to users and visibility under degraded visual conditions. Despite the absence of much formal evaluation (especially for visibility), the adequacy of many suggested exit symbols appears questionable. Furthermore, there seems to be an apparent trade-off among existing symbols between meaningfulness and visibility.

Among the problems contributing to the development of a good exit symbol is the absence of a common stereotype image (graphic concept). In previous research (Lerner and Collins, 1980), participants showed relatively little agreement in proposing exit pictograms, except for the frequent use of a directional arrow.

The arrow was discussed as a graphic element in exit designation. While it appears to connote "exit" strongly, the arrow itself is too ambiguous and too widely used to serve as the primary element of an exit symbol. On the other hand, a number of advantages were noted for standardizing some particular arrow uniquely for use with exit designators. These advantages include increased understandability of the exit message, reduced confusability with irrelevant signage, better learnability for exit symbols, and improved visibility.

Two experiments were described in this report. The first concerned the visibility of a set of 32 arrows under conditions of reduced luminance and contrast. The experiment was directed at investigating a methodology for assessing symbol visibility, and also at comparing the visibilities of the specific arrows tested. The second experiment obtained subjective rankings of thirteen arrows according to how well they suggested "exit", how well they indicated direction, how unique they were, and how appropriate and attractive they appeared.

The visibility experiment procedure proved to be an effective method, and wide differences in visibility among the arrows were detected. The arrows also differed in the type of confusion they engendered when an error in perception occurred. The subjective arrow rankings showed substantial agreement among participants, and strong correlations between the subjective dimensions. Together, the visibility and ranking data were used to compare the set of arrows for appropriateness for use with exit designators, and suggestions were made for developing other candidate arrows.

Some methodological problems in assessing the visibility of pictograms were discussed, and a procedure was recommended.

Finally, the need to give special consideration to some specific groups of potential symbol users--including the elderly, children, and others--was emphasized. It may not be sufficient simply to sample broadly in evaluating a symbol, if the problems for groups of special concern become lost in the collective data.

REFERENCES

1. AIGA. The development of passenger/pedestrian oriented symbols for use in transportation-related facilities, U.S. Department of Transportation, DOT-RSPA-DPB-40-79-2, 1979.

2. Berl, W. G. and Halpin, B. M., "Human fatalities from unwanted fires," Fire Journal, 1979, 73, pp. 105-115, 123.

3. Bono, J. A., and Breed, B. K., "Study of smoke ratings developed in standard fire tests in relation to visual observations," Underwriters' Laboratories Bulletin of Research, 1965, 56, pp. 1-59.

4. Carr, R., "The legibility of signs," Print, 1969, pp. 28-31.

5. Collins, B. L. "The development and evaluation of effective symbols," National Bureau of Standards, NBSIR 1981, in preparation.

6. Collins, B. L. and Lerner, N. D., "Preliminary assessment of workplace safety symbols: pilot study results," unpublished letter report, 1980, (to be published as part of larger report, in preparation).

7. Collins, B. L. and Pierman, B. P., "Evaluation of safety symbols," National Bureau of Standards, NBSIR-79-1760, 1979.

8. Demaree, J. E., "A preliminary examination of interior aircraft emergency lighting under simulated postcrash fire and smoke conditions," Federal Aviation Administration, NA-79-46-LR, 1979.

9. Dewar, R. E., "The slash obscures the symbol on prohibitive traffic signs," Human Factors, 1976, 18, pp. 253-258.

10. Dewar, R. E. and Ells, J. G., "Comparison of three methods for evaluating traffic signs," Transportation Research Record, 1974, 503, pp. 38-47.

11. Dewar, R. E. and Swanson, H. A., "Recognition of traffic control signs," Highway Research Record, 1972, 414, pp. 16-23.

12. Easterby, R. S., "Perceptual organization in static displays for man/machine systems," Ergonomics, 1967, 10, pp. 193-205.

13. 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, 1977.

14. Easterby, R. S. and Zwaga, H. T. G., "Evaluation of public information symbols, ISO tests: 1975 series," AP Report 60, Applied Psychology Department, University of Aston, Birminghmam, England, 1976.

15. Follis, J. and Hammer, D., "Architectural signing and graphics," New York: Whitney Library of Design, 1979.

16. Fozard, J. I., Wolf, E., Bell, B., McFarland, R. A., and Podolsky, S., "Visual perception and communication," In J. E. Birren and Schaie, K. W. (Eds) Handbook of the Psychology of Aging, New York: Van Nostrand Reinhold Company, 1977.

17. Gross, D., Loftus, J. J., and Robertson, A. F., "Method for measuring smoke from burning materials," in ASTM, Symposium on Fire Test Methods - Restraint and Smoke, Philadelphia, PA, American Society for Testing and Materials, 1966.

18. Hollander, M. and Wolfe, D. A., "Nonparametric Statistical Methods," New York, John Wiley and Sons, 1973.

19. Hulbert, S., Beers, J., and Fowler, P., "Motorists' understanding of traffic control devices," AAA Foundation for Traffic Safety, Falls Church, VA, 1979.

20. IES. IES Lighting Handbook, New York, Illuminating Engineering Society, 1972.

21. ISO/TC80, "Safety colours and safety signs," draft International Standard ISO/DIS3864.3, Geneva, International Organization for Standardization, 1979.

22. ISO/TC145, "Graphic symbols - use of arrows," ISO/4196, Geneva, International Organization for Standardization, 1977.

23. Jacobs, R. J., Johnston, A. W., and Cole, B. L., "The visibility of alphabetic and symbolic traffic signs," Australian Road Research, 1975, 5, pp. 68-86.

24. Japanese Proposal of Pictograph Regarding ISO/TC21/SC1 N25, DP6309 Safety Sign No. 4 "Emergency Exit," May 1980.

25. Jin, T., "Visibility through fire smoke," Report of Fire Research Institute of Japan, No. 30, 1970.

26. Keppel, G., Design and Analysis: A Researcher's Handbook, Englewood Cliffs, NJ, Prentice Hall, 1973.

27. King, L. E., "A laboratory comparison of symbol and word roadway signs," Traffic Engineering and Control, 1971, 12, pp. 518-520.

28. King, L. E., "Recognition of symbol and word traffic signs," Journal of Safety Research, 1975, 7, pp. 80-84.

29. Koncelik, J. A., "Human factors and environmental design for the aging: aspects of physiological change and sensory loss as design criteria," in T. O. Byerts, S. C. Howell, and L. A. Pastalan (Eds.), Environmental Context of Aging, New York, Garland STPM Press, 1979.

30. Konz, S. and Mohan, R., "The effect of illumination level, stroke width, and figure ground on legibility of Namel numbers," Proceedings of the Human Factors Society, 1972, pp. 431-435.

31. Lerner, N. D. and Collins, B. L., "The assessment of safety symbol understandability by different testing methods," National Bureau of Standards, NBSIR-80-2088, 1980.

32. Mackett-Stout, J. and Dewar, R., "Evaluation of symbolic public information signs," Human Factors, in press.

33. Markowitz, J., Dietrich, C. W., Lees, W. J., and Farman, M., "An investigation of the design and performance of traffic control devices," Bolt Beranek and Newman, Report No. 1726, 1968.

34. Middleton, W. E., Vision Through the Atmosphere, Toronto, University of Toronto Press, 1952.

35. National Safety Council, Accident Facts, Chicago, IL, National Safety Council, 1979.

36. O'Neill, J. G., Hayes, W. D., and Zile, R. H., "Full-scale fire tests with automatic sprinklers in a patient's room: Phase II," National Bureau of Standards, NBSIR-80-2097, 1980.

37. Overington, I., Vision and Acquisition, New York, Crane, Russak, and Company, 1976.

38. Phillips, A. W., "The effects of smoke on human behavior," Fire Journal, 1978, 72, pp. 69-77, 122-123.

39. Sharry, J. A.(Ed.), Life Safety Code Handbook, Boston, National Fire Protection Association, 1978.

40. Smith, G. and Weir, R., "Laboratory visibility studies of directional symbols used for traffic control signals," Ergonomics, 1978, 21, pp. 247-252.

41. U.S. Public Health Service, Binocular Visual Acuity of Adults, U.S. Department of Health, Education, and Welfare, 1964.

42. Veterans Administration, Signage System Standards Manual, Washington, D.C., Veterans Administration, 1980.

43. Weale, R. A., The Aging Eye, London, H. K. Lewis, 1963.

44. Winer, B. J., Statistical Principles in Experimental Design, New York, McGraw-Hill, 1971.

45. Zwaga, H. J., "Legibility of public information symbols," Proceedings of the Human Factors Society, 1979, pp. 448 (Abstract).

APPENDIX A

Sequence of Stimulus Sets and Visibility Condition (Filter)[a]

Day	Lists	Participants 1	2	3	4	5	6	7
Day 1	P	0	0	0	0	0	0	0
	P	4.6	4.6	4.6	4.6	4.6	4.6	4.6
	1–2	4.6	4.6	5.1	5.0	4.6	5.1	5.0
	3–4	5.1	5.0	4.6	4.6	5.1	5.0	5.1
	5–6	5.0	5.1	5.0	5.1	5.0	4.6	4.6
Day 2	7–8	5.1	5.0	4.6	5.1	4.6	5.0	4.6
	9–10	5.0	5.1	5.0	5.0	5.0	5.1	5.0
	11–12	4.6	4.6	5.1	4.6	5.1	4.6	5.1
	13–14	5.0	5.1	5.0	5.1	5.0	5.0	4.6
Day 3	15–16	5.1	5.0	5.1	4.6	5.1	4.6	5.1
	5–6	4.6	4.6	4.6	5.0	4.6	5.1	5.0
	3–4	5.0	5.1	5.1	4.6	5.1	4.6	5.1
	1–2	4.6	4.6	5.0	5.1	5.0	5.0	4.6
Day 4	13–14	5.1	5.0	4.6	5.0	4.6	5.1	5.0
	11–12	4.6	4.6	5.0	4.6	5.1	4.6	5.1
	9–10	5.0	5.1	4.6	5.0	4.6	5.1	5.0
	7–8	5.1	5.0	5.1	5.1	5.0	5.0	4.6
Day 5	15–16	5.1	5.0	5.0	5.0	5.0	5.1	5.0
	1–2	4.6	4.6	4.6	5.1	4.6	4.6	4.6
	3–4	5.0	5.1	5.1	4.6	5.1	5.0	5.1

[a] Each entry in the main body of the table is the optical density (D) of the filter through which the arrow image was projected.

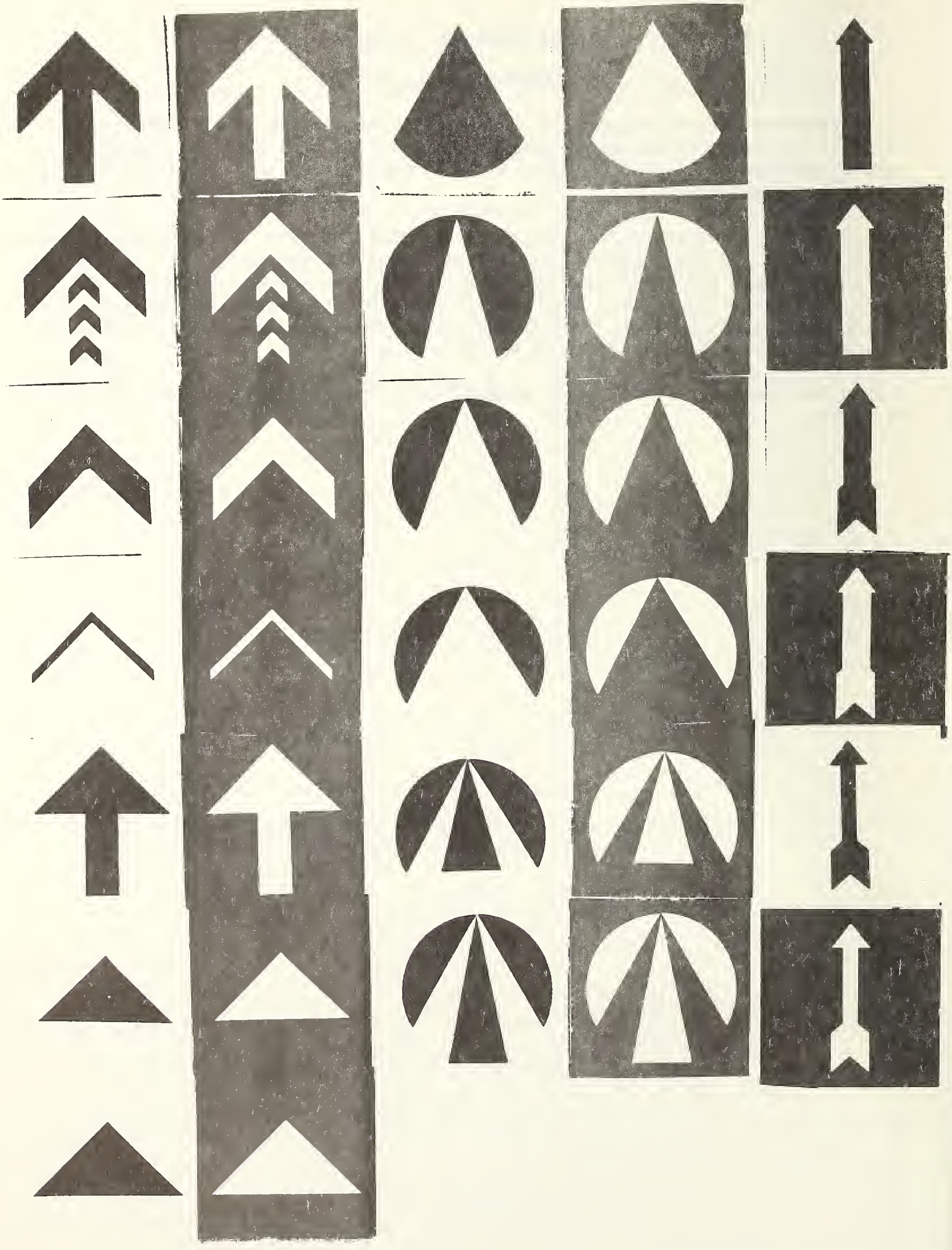
Appendix B. Instructions for visibility experiment

INSTRUCTIONS (A)

The purpose of this experiment is to determine how visible various directional "arrows" are under difficult viewing conditions. We will show you slides of arrows, and for each slide you will be asked to indicate the direction in which the arrow is pointing.

Before we start, we would like you to become familiar with the sorts of arrows you will see. On the attached page are copies of all the arrows in the set. Notice that each one occurs as both a black figure on a white background and again as a white figure on a black background. All of the arrows are pointing in the same direction (up). During the actual tests, the arrows will occur in a random sequence and may point in any direction.

Take a few minutes to familiarize yourself with these directional arrows before we go on.



INSTRUCTIONS (B)

Be seated in the chair directly in front of the screen. You will see arrows appear on the screen, one at a time. For each arrow, indicate on the answer sheet the direction the arrow is pointing: up, down, left, or right. Indicate the direction by drawing an arrowhead in the appropriate direction: Λ, if what you see points up; V, if what you see points down; <, if what you see points left; >, if what you see points right.

Each arrow will appear on the screen for 3 seconds, followed by 5 seconds with nothing on the screen. Then the next arrow will appear for 3 seconds, and so on. Use the 5 seconds between arrows to write down your answer. A tone will indicate this answer period--when the tone goes off, the next slide will be presented.

We will do a practice list to become familiar with the procedure. Be sure to write down an answer for each slide. If you are not sure, then guess. There must be an answer written down for every slide you see.

Although the arrows in the practice list will appear quite plain to you, they will be much more difficult to see once we begin the actual experiment. This is because we are interested in the limits of visibility--everything you see, except for the practice list, will be faint. At some times you may feel that you don't see anything. It is very important that you concentrate and make your best guess. This is because in a task such as this, even when people feel they are only "guessing," they actually turn out to be correct much more often than they think. Be sure to make your best guess for every slide you see.

Appendix C. Instructions for Arrow Ranking Experiment

(Directional Criterion)

We would like you to rank various arrows according to how well each arrow serves as an indicator of which direction to go. The arrow that you feel best serves to indicate direction should be placed on the top of the stack, and the one that you feel least serves to indicate direction should be placed at the bottom, with the others ordered in between.

(Appearance Criterion)

We would like you to rank various arrows according to how well you like each arrow's appearance. Consider how attractive and appropriate an arrow would look when displayed as a sign in a public building. Order the sheets from best appearance to worst. The most attractive and appropriate arrow should be placed at the top of the stack, and the least attractive and appropriate should be placed at the bottom, with the others ordered in between.

(Exit Criterion)

We would like you to rank various arrows according to how well each arrow conveys the idea of "exit." Order the sheets from the one that you feel most suggests exit to the one you feel least suggests exit. The arrow that best denotes exit should be placed on the top of the stack, and the one that least denotes exit should be placed at the bottom, with the others ordered in between.

(Uniqueness Criterion)

We would like you to rank various arrows acording to how unique or distinct you feel each arrow is. Order the sheets from the one you feel is most unique compared to directional arrows you might ordinarily see to the one you feel is most typical of the sorts of arrows you might ordinarily see. The most unique arrow should be placed on the top of the stack, and the least unique should be placed at the bottom, with the others ordered in between.

NBS-114A (REV. 2-80)

U.S. DEPT. OF COMM. BIBLIOGRAPHIC DATA SHEET (See instructions)	1. PUBLICATION OR REPORT NO. NBSIR 81-2268	2. Performing Organ. Report No.	3. Publication Date May 1981

4. TITLE AND SUBTITLE

Evaluation of exit directional symbols

5. AUTHOR(S)

Neil D. Lerner

6. PERFORMING ORGANIZATION (If joint or other than NBS, see instructions)

NATIONAL BUREAU OF STANDARDS
DEPARTMENT OF COMMERCE
WASHINGTON, D.C. 20234

7. Contract/Grant No.

8. Type of Report & Period Covered

Final

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

National Bureau of Standards
Department of Commerce
Washington, D.C. 20234

10. SUPPLEMENTARY NOTES

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

11. 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 discusses visibility considerations for exit symbols and the relationship between understandability and visibility concerns. Two experiments evaluated directional indicators (arrows) in the context of building exit signage. The first experiment compared the visibility of 32 arrows under degraded visual conditions that were comparable to a smoke environment. This experiment had two objectives: (1) the development of a methodology for assessing exit pictogram visibility; and (2) a comparison of the visibilities of the specific arrows tested. A second experiment obtained subjective rankings of the arrow types on the basis of several criteria of concern for exit signage. These criteria included connotative meaning, uniqueness from other directional indicators, and appropriateness.

The visibility procedure proved to be statistically sensitive and demonstrated differences in the visibility of different arrows. An analysis of the type of confusions that occurred in errors for each arrow suggested certain relations between graphic features and errors in detection. The second experiment indicated substantial agreement between participants in ranking the arrows and revealed strong relationships between the several criteria. Together the results of the two experiments were used to evaluate the set of arrows for appropriateness for use with exit designators.

Methodological issues in evaluating symbol visibility were considered and other issues of concern in testing exit signage, such as special user groups, were also discussed.

12. KEY WORDS (Six to twelve entries; alphabetical order; capitalize only proper names; and separate key words by semicolons)

Exit signs; directional arrows; fire safety; pictograms; symbols; understandability; visibility; visual alerting

13. AVAILABILITY

[X] Unlimited

☐ For Official Distribution. Do Not Release to NTIS

☐ Order From Superintendent of Documents, U.S. Government Printing Office, Washington, D.C. 20402.

[X] Order From National Technical Information Service (NTIS), Springfield, VA. 22161

14. NO. OF PRINTED PAGES

51

15. Price

$8.00