te technical note techn

Visual and Auditory Symbols: A Literature Review

October 1994
DOT/FAA/CT-TN94/37

This document is available to the public through the National Technical Information Service, Springfield, Virginia 22161

U.S. Department of Transportation
Federal Aviation Administration

Technical Center
Atlantic City Airport, NJ 08405

NOTICE

This document is disseminated under the sponsorship of the U.S. Department of Transportation in the interest of information exchange. The United States Government assumes no liability for the contents or use thereof.

The United States Government does not endorse products or manufacturers. Trade or manufacturers' names appear herein solely because they are considered essential to the objective of this report.

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
DOT/FAA/CT-TN94/37		

4. Title and Subtitle	5. Report Date
Visual and Auditory Symbols: A Literature Review	October 1994
	6. Performing Organization Code
	8. Performing Organization Report No.
7. Author(s)	DOT/FAA/CT-TN94/37
James P. Duncanson, Ph.D	
9. Performing Organization Name and Address	10. Work Unit No. (TRAIS)
CTA INCORPORATED 2500 English Creek Avenne, Suite 1000 Pleasantville, NJ 08232	11. Contract or Grant No. F2203A
12. Sponsoring Agency Name and Address	13. Type of Report and Period Covered
Federal Aviation Administration Technical Center Advanced Systems Technology Branch Atlantic City International Airport, NJ 08405	Technical Note
	14. Sponsoring Agency Code ACD-350

15. Supplementary Notes

This work was accomplished as part of the FAA's Airways Facility RE&D Program, managed by Brenda Boone, ALM-100 and George Booth, ARD-200. Dan Wagner, ACD-350, is the project manager.

16. Abstract

This report is a review of the human factors literature on visual and auditory symbols. The review is the first step in an effort to develop a set of standard symbols for use throughout Airway Facilities.

The topics included in the review are: general principles, experimental findings, and standard guidelines. The report also includes an annotated bibliography and a list of possible research topics suggested by gaps in the experimental findings.

17. Key Words	18. Distribution Statement
Symbols, Icons, Auditory Symbols, Guidelines, Human Factors	This document is available to the public through the National Technical Information Service, Springfield, Virginia 22161

20. Security Classif. (of this page)	19. Security Classif. (of this report)	21. No. of Pages	22. Price
Unclassified	Unclassified	68	

Form DOT F 1700.7 (8_72)

Reproduction of completed page authorized

TABLE OF CONTENTS

		Page
EXECUTIVE SUMMARY		v
1.	INTRODUCTION	1
2.	GENERAL PRINCIPLES	1
2.1	"Figures"	2
2.2	Figure vs. Ground	2
2.3	Gestalt Laws of Perception	2
2.4	Use of Distinguishing Features	3
2.5	Use of Metaphors	3
2.6	"Earcon" Construction	3
2.7	Conclusions	4
3.	EXPERIMENTAL FINDINGS	4
	Visual Icons	4
3.1.1	Conclusions	6
	Auditory Icons and Earcons	7
3.2.1	Conclusions	8
	Visual Icons vs. Text	9
3.3.1	Conclusions	10
	Alerting Value of Symbols	11
3.4.1	Conclusions	12
4.	STANDARDS AND GUIDELINES	13
4.1	Visual Icons	13
4.1.1	When to Use Visual Icons	16
4.1.2	Summary	17
	Auditory Icons and Earcons	17
4.2.1	When to Use Auditory Icons or Earcons	18
4.2.2	Summary	20
	Alarms and Alerts	20
4.3.1	General	20
4.3.2	Auditory Alarms and Alerts	21

5. POSSIBLE RESEARCH TOPICS 25

GLOSSARY 27

REFERENCES 29

APPENDIX

A - Annotated Bibliography

LIST OF ILLUSTRATIONS

Table Page

4.2.1-1 Characteristics and Ratings of Audio Signals for Various Uses 19

EXECUTIVE SUMMARY

As the monitoring and control of systems and equipment become increasingly centralized, technicians responsible for maintaining these systems and equipment are exposed to an increasing variety of visual and auditory symbols. Because of a lack of standardization in the past, they may see or hear different symbols used to represent the same thing and the same (or similar) symbols used to represent different things. The adoption of a set of standard symbols to represent common systems, facilities, and states can be expected to reduce:

a. the time required to learn, interpret, and respond to symbols;

b. the risk of errors in interpreting and responding to symbols; and,

c. the time and expense required to design symbols, and the amount and cost of documentation.

This report is a review of the human factors literature on visual and auditory symbols. It is the first step in a planned effort to develop a set of standard symbols for use in Airway Facilities. The topics included in the review are: general principles, experimental findings, and standards and guidelines. An annotated bibliography is included as appendix A.

The general principles reviewed might be summarized as "a good symbol is simple, unitary, identifiable, readily associated with the thing it represents, and makes appropriate use of metaphors."

Conclusions derived from the reviewed research on visual symbols are:

a. A good symbol is simple rather than complex, concrete or abstract rather than arbitrary, larger rather than smaller, and filled or solid rather than hollow.

b. Promising methods for evaluating individual symbols include both objective measures, such as response times and time to learn, and subjective measures, such as preferences and ratings of familiarity, associability, communicativeness, and appropriateness.

c. Promising methods for evaluating symbols in the context of a set of symbols include the Discriminability Index, the Componential Model, and the Semantic Differential.

Conclusions derived from the reviewed research on auditory symbols are:

a. Auditory symbols can improve performance, and some users want them, even if they are completely redundant with visual symbols. However, since increasing noise levels in work places tends to produce adverse effects on workers, the use of auditory symbols should be approached with caution.

b. Auditory symbols can be a natural and intuitive means for conveying information.

c. The more successful auditory symbols tend to be complex sounds or "caricatures" of natural sounds rather than simple tones or unstructured bursts of noise.

d. Auditory symbols should conform to traditional human factors requirements with respect to minimum duration and minimum and maximum loudness levels to ensure that they are detectable and identifiable and that they minimize possible annoyance or harm.

A conclusion derived from the reviewed research comparing visual symbols with text is that sometimes symbols are better; sometimes text is better. The conditions under which one is better than the other are not clearly defined.

Conclusions derived from the reviewed research on the alerting value of symbols are:

a. Highlighting visual symbols with flicker or, to a lesser degree, increased luminance, increases the perceived urgency of the symbol.

b. The pulse rate of an auditory symbol has the greatest affect on its perceived urgency. Other parameters affecting perceived urgency are fundamental frequency, harmonic series, amplitude envelope shape, and rhythm.

c. Adding visual symbols to text does not necessarily increase the perceived urgency of the text, and adding visual symbols to instructions does not necessarily increase compliance with them.

This review includes summaries of standards and guidelines relevant to visual and auditory symbols. A list of topics for possible additional research suggested by gaps in the experimental findings is also included.

1. INTRODUCTION

More and more systems and facilities are being monitored and controlled remotely by technicians using visual display terminals in centralized locations. Consequently, these technicians become responsible for many different systems and facilities. Because of a lack of standardization, these technicians are likely to encounter different symbolic representations of the same object and the use of the same or similar symbols to represent different objects. Some probable consequences are increased training time to learn the symbols, increased time to interpret and respond to the symbols, and increased risk of error in interpreting and responding to the symbols.

The adoption of standard symbols to represent common conditions, systems, facilities, and states would alleviate these undesirable consequences. In addition to improved human performance, the adoption of standard symbols would yield savings in the design of symbols and the documentation of their use.

The objectives of this review are to summarize:

a. Theoretical and experimental answers to the question "What makes a good symbol"?

b. Existing standards and guidelines for designing symbols

An earlier literature review (Advanced Systems Technology Branch (ASTB), ACD-350, 1992) covered visual and auditory coding methods extensively, but had comparatively little to say about the design and evaluation of visual and auditory symbols or icons.

The present review is concentrated on nonverbal, nonalphanumeric visual and auditory icons. In the case of visual icons, the primary concern is with icons that can be presented on visual display terminals. This review does not cover coding, and it does not cover verbal labels as such, although it does include experimental studies comparing symbols and text. For any report, article, or book cited, the relevant findings or information are included in the body of this review; complete abstracts or other descriptions of the documents are given in appendix A. The terms "symbol" and "icon" are used synonymously in this review, the choice in any particular instance being the term that was used by the author of the work cited.

2. GENERAL PRINCIPLES.

This section summarizes general principles relevant to the perception of visual symbols as distinct objects, or "figures."

2.1 "FIGURES"

Presumably a "good" symbol must have the properties of a "good" figure. Goldstein (1989) lists the following properties that influence the perception of shapes or collections of shapes as figures:

a. Symmetry. Symmetric shapes tend to be seen as figures.

b. Convexity. Convex (outwardly bulging) shapes tend to be seen as figures. In fact, convexity usually has a stronger influence on perception than does symmetry.

c. Area. Objects having a comparatively smaller area are more likely to be perceived as figures.

d. Orientation. Objects or collections of objects having a vertical or horizontal orientation are more likely to be seen as figures than objects or collections of objects having other orientations.

2.2 FIGURE VS. GROUND.

Above all, a figure (or symbol) must be distinct from its background. Goldstein (1989) lists the following properties of perceptions of figures and grounds:

a. A figure is more "thinglike" and more memorable than the ground.

b. A figure is seen as being in front of the ground

c. The ground is seen as unformed material and as extending behind the figure.

d. The contour that separates the figure from the ground is seen as belonging to the figure.

2.3 GESTALT LAWS OF PERCEPTION.

In the early part of this century, Gestalt psychologists proposed several "laws" governing human perception of objects (Goldstein, 1989). These laws were thought to explain when a collection of points or objects would be perceived as a figure (or symbol). The laws are:

a. Good Figure or Simplicity. A stimulus pattern is seen so that the resulting structure is as simple as possible.

b. Similarity. Similar things appear to be seen as a group or unit. Similarity can be based on shape, hue or shade, orientation, and size.

c. Good Continuation. Points that would result in straight or smoothly curving lines if connected are seen as belonging together, and lines tend to be seen as following the smoothest path.

d. Proximity. Objects that are near to each other tend to be seen as a group or unit.

e. Common Fate. Objects that are moving in the same direction tend to be seen as a group or unit.

f. Meaningfulness or Familiarity. Objects are more likely to be seen as a group or unit if the group or unit is familiar or meaningful.

2.4 USE OF DISTINGUISHING FEATURES.

Kolers (1969) argued that symbols or icons that emphasize the distinguishing features of an object, as is done in caricatures, yield better recognition than do symbols that provide photographic or detailed representations of objects, in which all details are treated equally.

2.5 USE OF METAPHORS.

Waterworth, Chignell & Zhai (1993) and Gittins (1986) hypothesized that icons can have associative or metaphorical properties that provide users with information beyond simple identification, that is, information about the function or use of the objects they represent.

(Gittins, 1986) argued that a graphical metaphor can reinforce a mapping process in ways that would be difficult or impossible with other coding methods. However, there are limitations to the use of metaphors; some objects and functions are difficult to portray visually, and metaphors may break down, that is, they may not apply to all the symbols needed for an application.

2.6 "EARCON" CONSTRUCTION.

Blattner, Sumikawa & Greenberg (1989) proposed a set of principles and methods for constructing both representational and abstract earcons, which they define as the aural counterparts of icons. They asserted that the characteristics of a good figure (closure, continuity, symmetry, simplicity, and unity) apply to earcons as well. They suggested that representational earcons be constructed using natural sounds, analogously to representational (pictorial) icons, and that abstract earcons be constructed from elements such as pitches, timbres, and rhythms to form brief, distinctive patterns or melodies. They strongly recommended that such patterns be designed by experts in sound, perhaps trained composers.

2.7 CONCLUSIONS.

According to the principles in this section, a good symbol is simple, unitary (has the properties of a "figure"), identifiable, readily associated with the thing it represents, distinct from other symbols with which it may appear, and makes appropriate use of metaphors.

3. EXPERIMENTAL FINDINGS.

3.1 VISUAL ICONS.

Green & Pew (1978) evaluated 19 pictographic symbols for use in labeling automotive controls and displays. They found that the relative goodness of individual symbols was not consistent across five measures (reaction time, learning time, association norms, ratings of communicativeness, and ratings of familiarity). Specifically, they found that:

a. Reaction time decreased markedly with learning.

b. Association norms were not correlated with ease of learning or with the intuitions of the authors.

c. Ratings of communicativeness are correlated with associative strength and can be used to estimate it.

d. Prior familiarity does not seem to be important

Hovey & Berson (1987) compared performance and preference data on three sets of symbols for use in an antisubmarine warfare display. One set was abstract, geometric shapes; one was circles with identification information presented as letter codes adjacent to the circles; and one set was simple diagrams (of a boat, airplane, and submarine) accompanied by a letter code. They found that the abstract symbols yielded the fastest reaction times. Color coding the symbols also improved reaction time. The subjects also preferred the abstract set and said that color coding was "helpful."

Geiselman, Landee & Christen (1990) proposed an index of discriminability based on user ratings of intersymbol similarity as a basis for selecting among alternative symbols. They derived a formula for calculating a discriminability index and tested it experimentally, using it to predict search times for symbols embedded in an array. They concluded that the index would be useful as one basis for choosing among alternate symbols intended to be included in an existing set of symbols.

Stammers & Hoffman (1991) studied the usability of icons, examining transfer of learning between two sets of icons to see if there is a benefit attributable to common elements, and comparing concrete with abstract icons. They found little

transfer from one set to another, but they found that, in general, concrete icons were identified faster and with fewer errors than abstract icons. However, some abstract icons performed well, and some concrete icons did not. They found that subjects' ratings of the "appropriateness" of an icon was a good predictor of the performance of the icon.

Blackwell & Cuomo (1991) studied the discriminability of symbols and coding techniques of a set of symbols proposed as standards for space and missile warning systems. They found that:

a. Subjects found larger symbols faster and with fewer errors than smaller symbols.

b. Subjects found filled shapes fastest, hollow shapes slower, and half-filled shapes the slowest.

c. Subjects preferred simplified, less complex shapes and found them faster.

d. Subjects found red and green symbols equally fast

e. If markers were added to symbols (in this case, alphabetic letters adjacent to the symbol), subjects took longer to locate symbols and made more errors.

Blankenberger & Hahn (1991) studied the relation of "articulatory distance" (that is, the immediacy with which an icon is associated with what it represents as determined empirically) and performance. They constructed four sets of icons: a "near" set consisted of icons that were all identified correctly in pretesting; a "far" set that were identified correctly 70 to 90 percent of the time; an "arbitrary" set that were never identified correctly; and a "word" set that consisted of verbal labels. The authors found that articulatory distance was correlated with reaction time for inexperienced users, but that this correlation disappeared as users gained experience with the icons.

Fisher & Tanner (1992) proposed a model of visual search (based on components of symbols) that permits selection of the optimal set of symbols (the set that will yield the lowest average search time) from a larger set of candidate symbols. In an experimental evaluation, the componential model conformed to subject performance more closely than did an earlier discriminability model.

Kirkpatrick, Dutra, Lyons, Osga & Pucci (1992) studied the effects of adding color to symbols. They found that subjects preferred the use of color, that color added information, and that filling the symbol with color was more effective than using color to outline the symbol.

Lin (1992) conducted a study to derive and validate the cognitive factors that affect icon designs and examined use of the semantic differential as a tool to evaluate the icons as they are being designed. He concluded that the relevant dimensions are associable, symbolic, meaningful, concise, eye-catching, and identifiable, and that these dimensions are related to comprehension of and preferences for icons.

Byrne (1993) compared user performance on simple, complex, and blank icons for document retrieval in a human-computer interface. He found that simple icons (those permitting discrimination based on a few features) clearly outperformed blank (empty rectangles) or complex icons. Complex icons were the worst, worse even than the blank icons.

Waterworth, Chignell & Zhai (1993) were interested in how well different human-computer interface features could be interpreted in iconic form. They had 9 graduate students independently design icons for 61 concepts for a hypermedia interface; then each student tried to identify the icons designed by the other students. The authors found that classes of concepts varied in difficulty. Arranged from easiest to most difficult they were structure, media, user, information management, and content. The easiest concepts tended to be concrete objects existing in daily life, geometrical or spatial icons, and well-accepted standard symbols. The most difficult concepts tended to be inherently abstract or lacking in physical characteristics, or concepts that share a physical form with other concepts.

3.1.1 Conclusions.

The experimental studies of visual icons suggest that a good icon is:

a. simple rather than complex;

b. concrete or abstract rather than arbitrary;

c. larger rather than smaller; and

d. filled or solid rather than hollow.

Promising evaluation methods for individual symbols include both objective measures, such as response times and time to learn, and subjective measures, such as ratings of familiarity, associability (of the symbol with its object), communicativeness, appropriateness, and user preferences.

Promising methods for evaluating symbols in sets of symbols include the discriminability index developed by Geiselman, et al (1990), the componential model developed by Fisher, et al (1992), and use of the semantic differential as suggested by Lin (1992).

3.2 AUDITORY ICONS AND EARCONS.

Woodson & Conover (1964) cited evidence that any increase in noise level above threshold in a work place tends to increase muscular tension and, as a result, increases the expenditure of energy. Tasks that require intense mental concentration are adversely affected by sound. Operators in a noisy environment are more quickly fatigued, more nervous, and more irritable than operators in a quiet working environment. High-frequency sound is more annoying than low-frequency sound, and irregularly variable sounds are more annoying than steady or periodically changing sounds.

Gaver (1986) proposed that sound can be used in human-computer interfaces in a way analogous to the use of visual icons. He argued that using dimensions of sound such as pitch, loudness, and duration to represent dimensions of data, yields artificial-sounding tones that are more analogous to unlabeled graphs than to icons. An auditory icon, on the other hand, is a caricature of a naturally-occurring sound that suggests its source rather than drawing attention to the sound itself. Such caricatures need only capture the essential features of the sonic events they represent.

He distinguished types of representation as symbolic (the sound of a siren to represent an ambulance, or the sound of applause to represent approval), metaphoric (a hissing sound to represent a snake, or a tone with descending pitch to represent falling), and nomic (a "mailbox" sound to represent arriving mail).

He concluded that the use of auditory icons based on natural sounds provides a natural and intuitive way to represent dimensional data and conceptual objects in a computer system. Furthermore, he asserted that single sounds can permit the categorization of data into distinct families.

Sorkin (1987) stated that the minimum duration signal burst that can be reliably detected by humans is 100 millisecond (msec) plus 25 msec rise and fall times. Sounds of less than 200 to 500 msec in duration do not sound loud and often are not perceived over normal workplace ambient noise. Detectability increases as duration approaches and surpasses 500 msec.

Gaver (1989) described SonicFinder, an auditory interface for Apple computers. The auditory icons it contains are almost entirely redundant with visual information. It had not been formally evaluated, but there was anecdotal evidence that users like it and complain about its absence when they use computers on which it is not available. The following examples illustrate SonicFinder: the selection of a file is accompanied by a tapping sound, with the pitch of the sound corresponding to the size of the file, and dragging the file to the trash can is accompanied by a scraping sound. In a later paper, Gaver (1993) proposed

some synthesis algorithms for creating and manipulating auditory icons.

Folds (1990) hypothesized the feasibility of using multiple continuous background sounds to represent the status of systems or processes, sounds that would be perceptible but that would not be annoying and would not interfere with voice communications. In an experiment, he compared a condition in which subjects received only visual indications of the status of four processes with another in which they received the visual indicators and, in addition, continuous audio signals. Subjects in the audio plus visual condition consistently responded faster and more accurately, committed fewer false alarms, and rated their workload lower than subjects in the visual only condition.

Brewster, Wright & Edwards (1993) studied the effectiveness of earcons as a means of communicating information in sound. They concluded that earcons were more effective than unstructured bursts of sound, and that musical timbres were more effective than simple tones.

Gerth (1993) studied the identification of complex sounds created by simultaneously playing two or more component sounds having different timbres and patterns. He concluded that complex sounds can convey information such as a change in state of a sound source.

King & Corso (1993) cited reports that pilots often turn off auditory displays intended to improve their performance, and suggested, as probable causes, that pilots are simply trying to reduce the intensity of the auditory display and their subjective workload. They studied user performance and preferences for auditory displays of different intensities under varying workloads in a visual search task. An auditory signal indicated the part of the visual display in which the target was located. They found a nonmonotonic relationship between auditory display intensity and display usage, a positive relationship between auditory display and subjective workload, and improved performance with the display on, but a preference on the part of the participants to turn it off.

3.2.1 Conclusions.

The experimental studies of auditory icons suggest the following conclusions:

a. Since increasing noise levels in work places tend to produce adverse effects on workers, the use of auditory icons should be approached cautiously. Still, while some users turn off auditory signals, others want them, even if the signals are completely redundant with visual icons, and auditory icons can improve performance.

b. Auditory icons can be a natural and intuitive means conveying information.

c. The more successful auditory icons tend to be complex sounds rather than simple tones or unstructured bursts of sound. Caricatures of natural sounds (that is, sounds using only essential features) can be effective.

d. In order to ensure detectability, identifiability, and minimization of annoyance and possible harm, auditory icons should conform to traditional human factors requirements for such characteristics as minimum duration and minimum and maximum loudness levels.

3.3 VISUAL ICONS VS. TEXT.

Remington & Williams (1986) compared visual search time for numbers and graphic symbols in a simulated helicopter situation display. They found that numbers were located faster than graphic symbols and that the addition of a marker, such as a flashing dot or a surrounding square, impaired performance more for the graphic symbols than for the numbers.

Guastello, Traut, & Korienek (1989) compared user ratings of meaningfulness for four types of icons, short abbreviation, long abbreviation, pictorial, and pictorial plus abbreviated label. They found that:

a. The combined icons (pictorial plus label were rated the most meaningful.

b. Familiarity with an icon sometimes increased its meaningfulness.

c. Long abbreviations were more meaningful than short ones.

Kline, Ghali, & Kline (1990) compared visibility distances of textual and iconic highway signs for young, middle-aged, and elderly observers during daylight and dusk. They found that iconic signs were better for all age groups and at both times of day. In a similar study, Kline & Fuchs (1993) studied the visibility and comprehensibility of three versions of highway signs, one verbal, one "standard," and one "improved," for young, middle-aged, and elderly observers. They found that, on average, symbolic signs were identifiable at approximately twice the distance as text signs for all age groups. The size of letters on the text signs varied with text length; for example, the letters spelling "HILL" were quite large compared with those spelling "DIVIDED HIGHWAY."

Howard, O'Boyle, Eastman, Andre & Motoyama (1991) conducted two experiments comparing the effectiveness of symbols and keywords representing photocopier instructions for inexperienced users.

They concluded that symbols can perform at least as effectively as keywords.

Blankenberger & Hahn (1991) compared icons with text labels for editing operations (such as delete one character or move the cursor one word to the right). They found that icons yielded faster reaction times and fewer errors than did the text labels

MacGregor (1992) compared response time and errors for verbal labels, labels plus text descriptors, and labels plus icons as designators of menu selections in videotex menus. He found that adding either text or icons to labels reduced errors, but did not reduce response times. He cautions that one particular type of error was reduced, and that the results may not generalize to other sets of icons and descriptors.

Osborne & Huntley (1992) compared text with icons as a means for giving pilots instructions about missed landing approaches. They found that iconic instructions were comprehended more quickly and accurately than text instructions presented in the font style and size used by the National Ocean Service, and that the pilots tested preferred the iconic instructions.

Benbasat & Todd (1993) discussed the claimed advantages and disadvantages of the use of icons, distinguishing between the inherent properties of icons and specific implementations. They also summarized empirical investigations of iconic interfaces, concluding that none has found a definite advantage for iconic representation. They compared iconic and text-based representations of objects and actions in an electronic mail system and again found no advantage for iconic over text-based representations.

Vukelich & Whitaker (1993) studied the effects of different amounts of verbal context (two words and two sentences) on the comprehension of graphic symbols. They found that context did increase comprehension.

3.3.1 Conclusions.

Studies comparing user performance with icons and with text have yielded mixed results; in some, icons perform better; in others, text performs better. As Benbasat and Todd point out, it seems to be the implementation that matters. It does seem that adding labels to icons helps. To some extent, this effort seems misguided; it may well be that it is not that either icons or words are better in general, but that some objects, concepts, and actions can be expressed more adequately as icons, and others, more adequately as text.

3.4 ALERTING VALUE OF SYMBOLS

Hakkinen & Williges (1984) examined the effectiveness of preceding synthesized voice warning messages with alerting cues These cues consisted of the simultaneous presentation of an audible tone and a light in a high-workload, human-computer interaction task. They found that if there is only one type of message, the alerting cues served no purpose, and, in fact, lengthened response times. However, if the cues signalled a particular type of message, the absence of a cue resulted in a decreased detection of the message.

Friedmann (1988) studied the effect of adding symbols to written warnings on consumer products. The symbols did not significantly increase compliance with the warning.

Edworthy, Loxley & Dennis (1991) studied the effects of individual sound parameters on perceived urgency. They varied parameters affecting the temporal characteristics of sound (onset, duration, and offset), overall level, and spectrum (fundamental frequency and harmonics). They found that fundamental frequency, harmonic series, amplitude envelope shape, delayed harmonics, and melodic parameters, such as speed, rhythm, and pitch range, all have clear and consistent effects on perceived urgency.

Baber & Wankling (1992) studied the perceived urgency of different formats (symbols only, symbols plus title, symbols plus an action instruction, and symbol plus title plus an action instruction) for presenting warnings to automobile drivers. They found that the symbol plus action format yielded the highest perceived urgency, and that the symbol plus title plus action format yielded the fastest reaction times.

Erlichman (1992) studied the effects of adding elements and properties to pictograms on the alerting value of the combined stimulus in an in-vehicle safety advisory and warning system. The elements and properties added were color, blinking, tones, text messages, and speech. He concluded that the most effective combination was pictogram plus color, text, audio tone, and speech. He found that some of the pictograms were confusing and called for standardization of the symbols. He also concluded that the audio tone would be more meaningful if it represented the sounds associated with emergency vehicles.

Haas (1992) studied the effect of various combinations of pulse format, pulse duration, and time between pulses on the perceived urgency of auditory warning signals. Pulse format and time between pulses were found to affect perceived urgency; sequential pulses seemed less urgent than other formats, and signals with shorter interpulse intervals were rated more urgent.

Haas & Casali (1993) studied the perceived urgency of auditory signals in a variety of pulse parameters (format, level, and interpulse interval). Their findings about perceived urgency were:

a. The higher the pulse level, the greater was the perceived urgency and the shorter the detection time.

b. Sequential signals were perceived as less urgent and took longer to detect.

c. Time between pulses affected perceived urgency but not detection time.

Hellier, Edworthy & Dennis (1993) determined equal units of perceived urgency for four parameters (pulse rate, fundamental frequency, repetition units, and inharmonicity) of auditory warning sounds using Steven's power law scaling. They then predicted the perceived urgency of combinations of these parameters, and found a high correlation with subjects' ratings In spite of the equating, some parameters (pulse rate, in particular) still contributed more to perceived urgency than others (inharmonicity, in particular).

Jensen (1993) described the selection of an auditory alarm to be added to an existing set of alarms for use in Space Station Freedom. The new alarm, a "chirp" alarm to indicate toxic atmosphere, was the best of four candidate alarms and yielded elicited response times and correct identifications comparable to the existing alarms (fire or smoke, rapid change in pressure, master alarm, and system management alarm).

Van Orden & DiVita (1993) surveyed previous literature and concluded that highlighting color coded symbols using flicker results in reduced search times without impairing search for nonhighlighted symbols. They conducted a series of experiments that confirmed this conclusion.

Another series of experiments conducted by Van Orden, DiVita & Shim (1993) led to the conclusion that using either an increase in luminance or flashing to highlight symbols can reduce search times for those symbols without producing a concomitant increase in search times for nonhighlighted symbols.

3.4.1 Conclusions.

The research summarized in this section seems to permit the following conclusions:

a. Highlighting symbols with flicker or, to a lesser degree increased luminance increases the perceived urgency of the symbol.

b. The pulse rate of an auditory signal has the greatest effect on its perceived urgency. Other parameters of auditory signals affecting perceived urgency are fundamental frequency, harmonic series, amplitude envelope shape, and rhythm.

c. Adding visual symbols to text or instructions does not necessarily increase the perceived urgency of the text or increase compliance with the text.

4. STANDARDS AND GUIDELINES.

4.1 VISUAL ICONS

Gittins (1986) suggests the following guidelines:

a. The ground form of a figure should be clear and stable

b. The boundary around an icon should be solid, closed, and contrast-bounded, with the corners as smooth as possible. Color alone should not be used as the boundary.

c. Technological icons should be used rather than natural objects to reduce differences in interpretation due to cultural variations. To the extent possible, graphic primitives and boundary shapes should be consistent for all icons.

d. Use color judiciously and for a specific purpose.

e. If appropriate, provide users the ability to migrate to other sets of icons, other metaphors, and other media.

f. Avoid metaphors if self-contained icons of equal usability can be designed.

Schmidt & Kern (1991) provide the following guidelines for luminance (symbol-background) contrast, resolution, and symbol size:

a. Luminance Contrast. The ratio of luminance between the symbol and the surrounding background should be at least 3:1, with 7:1 preferred. For cathode ray tube (CRT) presentation, either the symbol or its background, whichever is of higher luminance, must be 10 foot lamberts or more.

b. Resolution. Resolution is considered "good" if the separation between elements is less than the resolving power of the human eye. When this quality is matched with the guidance for symbol size (described below), a symbol should be presented with not less than 10 lines or resolution elements.

c. Size. Generally, symbol size should be sufficient for visual acuity and color differentiation (if a color code is used) but small enough to minimize clutter and overlap when many

symbols are on the display surface. In terms of presentation size, the literature consensus is that at least one square centimeter of symbol area and at least 20 to 24 minutes of visual arc is recommended (at least 15 minutes of visual arc is required for the recognition of nonredundant color).

Lin & Kreifeldt (1992) suggest the following guidelines for choosing between pictorial and abstract icons:

a. If the information or function appears to have strong ; with an object, use a pictorial icon.

b. If an icon will be used frequently, of if it is strongly associated with a common concept, use an abstract icon.

Marcus (1992 suggests the following steps for developing icons:

a. Analyze the verbal contents and the display environment.

b. Design the initial icons by creating quick sketches.

c. Sort the icons into styles

d. Design a layout grid that organizes all major elements of icons.

e. Use large objects, bold lines, and simple areas to distinguish icons.

f Simplify their appearance.

g. Use color with discretion.

h Evaluate the designs by showing them to potential users

Strijland 1992 asserts that a good icon:

a. is based on an appropriate metaphor;

b. works in the computer's context; and

c. looks good on the screen.

She recommends strongly that icons be prepared by professional graphics illustrators, and that all icons in a set have a consistent style.

ASTB, ACD-350 1993a) provides the following guidelines:

a. Ensure that geometric shapes vary widely

b. Keep the number of shape symbols to a minimum.

c. Exploit the associative meanings of commonly-used geometric shapes, for example, the association of an octagon with "stop" and the association of an inverted triangle with "caution."

d. Use common graphic features to represent shared features.

e. Clearly define the boundaries of graphic symbols.

f. Contrast background and boundary.

g. Close graphic symbols.

h. Simplify graphic symbols.

i. Unify graphic symbols.

j. Incorporate only the basic elements of graphic symbols for identification.

ASTB, ACD-350 (1993b) provides the following standards and guidelines:

a. Size. If a target of complex shape is to be distinguished from other objects having complex shapes, the target shape should subtend a visual angle of at least 20 minutes. An angle of 40 minutes is preferred. At a viewing distance of 18 inches (in), these angles represent values of 0.06 in and 0.12 in, respectively. The complex shape should span at least 10 lines or resolution elements. Image quality should be consistent with the maintainer's needs.

b. Resolution. CRTs for displaying complex symbols and graphic detail should have at least 100 resolution elements per in.

c. Pictorial symbols. If pictorial symbols are used in place of or in addition to word labels [on controls], they shall be completely unambiguous in the expected visual operating environment. They shall not be used on a control that may rotate and thus position the symbol so that it may be confusing. Pictorial symbols shall be approved by the acquisition program office.

d. Icon design. To the extent possible, icons should be simple line drawings that suggest the object or operation they represent. Humorous representations should be avoided. Icons may be designed to represent a process or operation literally (a drawing of an aircraft), functionally (a figure representing a network), or operationally (a drawing of a pen in hand on paper).

e. Size of color-coded symbols. A symbol that is color coded shall subtend a visual angle of at least 20 minutes. For

example, at a viewing distance of 800 mm (31.5 in), a symbol would be at least 5 millimeter (mm) (0.2 in) high.

f. Brightness of color-coded symbols. Color-coded symbols shall have a minimum brightness of 1 foot lambert.

g. Refresh rate of color-coded symbols. Color-coded symbols shall have a refresh rate that provides no perceptible flicker.

h. Design of symbols. To the extent possible, a symbol should be:

1. an analog of the object it represents;

2. in general use and well known to the users; or

3. based on established standards or conventional meanings

i. Consistent use of symbols. Symbols, if used, shall be assigned unique meanings and used consistently throughout an application and related applications.

4.1.1 When to Use Visual Icons.

Strijland (1992) recommends using visual icons if:

a. they are intended for international applications;

b. they must be recognizable at a distance;

c. space is limited; and

d. their use is supported by the contextual environment.

She recommends not using visual icons if:

a. they represent actions or concepts that are difficult or impossible to portray in icons; or

b. their use is not supported by the contextual environment.

ASTB, ACD-350 1993a) recommends that visual icons be used:

a. to represent a lot of information in a small space;

b. to represent discrete information;

c. if an established icon already exists.

4.1.2 Summary.

The major points covered in these guidelines might be summarized as follows:

a. A pictorial or representational symbol is preferred over an abstract symbol unless there is a strong association between an abstract shape and a meaning, as exists in traffic signs, for example. The picture or representation should be a simplified portrayal that emphasizes the distinguishing features of the thing portrayed.

b. A symbol should be a clear, perceptually stable shape with a closed border or outline. Preferably, it will be solid or filled, providing a good contrast with its background.

c. A symbol must be large enough to be recognizable under its normal conditions of use - at least 20 minutes of visual angle.

d. Visual symbols are particularly useful if language might be a problem and if information must be conveyed at a distance.

4.2 AUDITORY ICONS AND EARCONS.

Gerth (1991) suggests the following guidelines for the design of auditory icons:

a. If there will be more than one auditory signal, its source should be unambiguous, both when other auditory signals are presented simultaneously and nonsimultaneously with the signal.

b. The auditory signals within a system or application should be easily distinguishable from each other. For suggestions on designing easily discriminated sounds, see Gerth, 1992.

ASTB, ACD-350 (1993a) provides the following guidelines:

a. Auditory signals should be easily perceived and identified.

b. Auditory signals should be compatible with environmental sounds.

c. The total number of auditory signals should be minimized.

d. Auditory signals should have both alerting and message components.

e. Auditory signals should not be less than 6 decibel (dB)(A) and should not exceed 90 dB(A).

Brewster, Wright & Edwards (1993) suggest the following guidelines for the creation of earcons:

a. Timbre. Use synthesized musical instrument timbres. If possible, use timbres with multiple harmonics. The timbres used should be easy to distinguish from each other, for example, "brass" and "organ."

b. Pitch. Do not use pitch alone unless the differences are large.

c. Register. Do not use register alone unless the differences are large.

d. Rhythm. Rhythms should be as different as possible. Using different numbers of notes in the different rhythms is effective.

e. Intensity. The suggested range is from a minimum of 10dB above threshold to a maximum of 20 dB above threshold, but with overall sound level under the control of the user.

f. Separation. Sequential earcons should be separated by at least 0.1 second (sec).

ASTB, ACD-350 (1993b) provides the following standards and guidelines:

All nonverbal audio signals shall be accompanied by a visual signal that defines the condition.

4.2.1 When to Use Auditory Icons or Earcons.

Gerth (1991) suggests the following guidelines on when to use auditory displays:

a. Use auditory displays only for alarm and alerting signals.

b. Auditory displays should be redundant with visual displays.

ASTB, ACD-350 (1993a) recommends that auditory signals be used to represent information requiring special user attention.

ASTB, ACD-350 (1993b) provides the following guidelines:

a. Signal type. If an audio display is used, the particular type of signal (tone, complex sound, or speech) should be the best for the intended use as indicated in table 4.2.1-1.

TABLE 4.2.1-1. CHARACTERISTICS AND RATINGS OF AUDIO SIGNALS FOR VARIOUS USES

Use	Tones	Complex Sounds	Speech
Quantitative indication	(Poor) Maximum of 5 to 6 tones absolutely recognizable.	(Poor) Interpolation between signals inaccurate.	(Good) Minimum time and error in obtaining exact value in terms compatible with response.
Qualitative indication	(Poor to fair) Difficult to judge approximate value and direction of deviation from null setting unless presented in close temporal sequence.	(Poor) Difficult to judge approximate deviation from desired value.	(Good) Information concerning displacement, direction, rate presented in form compatible with required response.
Status indication	(Good) Start and stop timing; continuous information if rate of change of input is low.	(Good) Especially suitable for irregularly occurring signals, such as alarms.	(Poor) Inefficient; more easily masked; problem of repeatability.
Tracking	(Fair) Null position easily monitored; problem of signal-response compatibility.	(Poor) Required qualitative indications difficult to provide.	(Good) Meaning intrinsic in signal.
General	Good for automatic communication of limited information; meaning must be learned; easily generated.	Some sounds available with common meaning, for example, a fire bell; easily communicated.	Most effective for rapid, but not automatic, communication of complex, multidimen-sional information; meaning intrinsic in signal and context, if standardized; minimum learning required.

b. When to provide. Audio signals should be provided, as necessary, to: (1) warn personnel of impending danger, (2) alert a maintainer to a critical change in system or equipment status, and (3) remind a maintainer of critical actions that must be taken.

4.2.2 Summary.

The major points covered in these guidelines might be summarized as follows:

a. Auditory symbols should be easily perceived and easily distinguished from each other.

b. Auditory symbols should be compatible with other sounds in the environment in which they will be used.

c. The intensity of auditory symbols should be not less than 6 dB(A) and not more than 90 dB(A). The preferred presentation is to allow the user to control the level within the range 10 dB above threshold to 20 dB above threshold.

d. Auditory symbols should be constructed from complex sounds, for example, musical timbres, rather than from pure tones.

e. Use of an alerting signal preceding an auditory symbol should be considered.

f. Use of a redundant visual symbol, or text accompanying an auditory symbol, should be considered.

g. The restriction of auditory symbols to use as alerting signals should be considered.

4.3 ALARMS AND ALERTS.

4.3.1 General.

ASTB, ACD-350 1993b) provides the following standards and guidelines:

a. Distinctive and consistent alarms. Alarm signals and messages shall be distinctive and consistent for each class of event. For example, a signal alerting a user to an incoming message would be different from a signal alerting a user to a hazardous condition.

b. Acknowledging and terminating alarms. A system or application shall provide users a means of acknowledging critical and noncritical alarms. It shall also provide users a means of turning off alarm signals once the alarms have been acknowledged or the condition generating the alarm has been corrected.

Procedures for acknowledgment and termination shall not decrease the speed and accuracy of operator reaction to the alerting situation.

c. Feedback about alarms and alerts. Users shall be provided informative feedback for actions that trigger alarms and alerting signals. If necessary, users shall be able to request help and related information for the operation and processing of critical and noncritical alarms, messages, and signals.

d. Special acknowledgment of critical alarms. If a user must acknowledge a special or critical alarm in a unique way, for example, with a special combination of key strokes, this special acknowledgment shall not inhibit or slow the response to the condition initiating the alarm.

e. User setting of alarm parameters. If appropriate to the task, a system or application should allow a user to set the parameter or condition that results in a software-generated alarm, alert, or status message. Some examples of parameters or conditions are priorities, percentages, and absolute values or ranges of values. User setting of parameters should not be allowed if:

1. the settings by one user might affect the reception of alarms by another user;

2. the settings might affect the safety of systems, equipment, or personnel; or

3. alarm parameters are determined by functional procedural, or legal requirements.

4.3.2 Auditory Alarms and Alerts

ASTB, ACD-350 1993b) provides the following standards and guidelines:

a. Alerting and warning system. An alerting and warning system or signal shall provide the maintainer with a greater probability of detecting the triggering condition than his or her normal observation would provide in the absence of the alerting or warning system or signal.

b. Audible coding of priority levels. If audible signals accompany visual alarms, they, too, shall be coded by priority. Acceptable coding methods for audible signals include pulse coding, frequency modulation, and frequency. Intensity shall not be used as a coding method. If pulse coding is used, the number of levels shall not exceed three. If frequency coding is used, the number of levels shall not exceed five.

c. Nature of signals. Audio warning signals should consist of two elements, an alerting signal and an identifying or action signal.

d. Two-element signals. If reaction time is critical and a two-element signal is used, an alerting signal of 0.5 sec duration shall be provided followed by an identifying or action signal. All essential information shall be transmitted in the first 2.0 sec of the identifying or action signal.

e. Single-element signal. If reaction time is critical, signals shall be of short duration. If a single-element signal is used, all essential information shall be transmitted in the first 0.5 sec.

f. Caution signals. Caution signals shall be readily distinguishable from warning signals and shall be used to indicate conditions requiring awareness, but not necessarily immediate action.

g. Relation to visual displays. If used in conjunction with a visual display, an audio warning device shall be supplementary or supportive. The audio signal shall be used to alert and direct the maintainer's attention to the appropriate visual display.

h. Frequency range. The frequency range of a warning signal shall be between 200 and 5,000 Hertz (Hz). The preferred range is between 500 and 3,000 Hz. If the signal must be audible at a distance of 300 meters (m) (985 feet (ft)) or more, the frequency shall be below 1,000 Hz. If the signal must be heard around obstacles or through partitions, the frequency shall be below 500 Hz. The selected frequency band shall differ from the most intense background frequencies and shall be in accordance with the other criteria in this section.

i. Audibility. The intensity, duration, and source location of audio alarms and signals shall be compatible with the acoustical environment of the intended receiver.

j. Compatibility with clothing and equipment. If applicable, audio signals shall be loud enough to be heard and understood through equipment or garments such as parka hoods and hearing protective devices covering the ears of a listener. If the audio attenuation characteristics of the garments are known, the intensity of the signal shall be increased to compensate for the attenuation so that it meets the audibility requirements of item k, but does not exceed the levels specified in item l.

k. Audibility. Alarms shall exceed the prevailing ambient noise level by at least 10 dB(A) or any maximum sound level with a duration of 30 sec by at least 5 dB(A), whichever is louder.

l. Maximum intensity. The intensity of evacuation and emergency signals shall not exceed 115 dB(A). The intensity of other signals shall not exceed 90 dB(A). If meeting the requirement of item j would result in levels higher than these, the ambient noise level shall be decreased so that these levels are not exceeded.

m. Onset and sound pressure level. The onset of critical alerting signals should be sudden, and a relatively high sound pressure level should be provided, as specified in item i.

n. Dichotic presentation. If earphones will be worn in an operational environment, a dichotic presentation should be used whenever feasible, with the signal alternating from one ear to the other by means of a dual-channel headset.

o. Headset. If the maintainer will wear earphones covering both ears during normal equipment operation, the audio warning signal shall be directed to the maintainer's headset as well as to the work area.

p. When not to use headsets. Binaural headsets should not be used if both of the following conditions apply:

1. The ambient noise level in the operational environment is less than 85 dB(A).

2. The operational environment contains sounds that provides useful information to the maintainer and that information cannot be directed to the maintainer's headset. Such sounds may include voices, machine noise that indicates wear or malfunctions, and other auditory indications of system performance or mission status.

q. Multiple audio signals. If several different audio signals will be used to alert a maintainer to different conditions, the signals shall differ discriminably in intensity, pitch, or use of beats and harmonics. If absolute identification is required, the number of signals to be identified shall not exceed four.

r. Coding. If discrimination of warning signals from each other will be critical to personnel safety or system performance audio signals shall be appropriately coded. Alarms that are perceptibly different shall correspond to different conditions that require critically different maintainer responses, such as maintenance, emergency conditions, and health hazards. These signals shall be sufficiently different to minimize the maintainer's search of visual displays.

s. Differentiation from routine signals. Audio alarms intended to attract the maintainer's attention to a malfunction or failure shall be different from routine signals, such as

bells, buzzers, random noises generated by air conditioning other equipment, and normal operation noises.

t. Consistent signals. The meaning of audio warning signals selected for a particular function in a system should be consistent with warning signal meanings already established for that function.

u. Acoustic environment. Established signals shall be used, provided they are compatible with the acoustic environment and the requirements in this standard for voice communication systems. Standard signals shall not be used to convey new meanings.

v. Noninterference. Audio warning signals shall not interfere with any other critical functions or warning signals or mask any other critical audio signals.

w. Separate channels. If feasible, a warning signal delivered to a headset that might mask another essential audio signal should be delivered to one ear and the other signal to the other ear. If this is done, and if it is warranted by the operating conditions, this dichotic presentation should provide for alternation of the two signals from one ear to the other.

x. Automatic and manual shutoff. If an audio signal is designed to persist as long as it contributes useful information, a shutoff switch controllable by the maintainer, the sensing mechanism, or both, shall be provided consistent with the operational situation and personnel safety.

y. Alarm reset. A system or application shall provide users with a simple means for turning off an auditory alarm without erasing any displayed message that accompanies the auditory signal.

z. Automatic reset. An automatic reset function for audio signals shall be provided, whether the signals are designed to terminate automatically, manually, or both. The automatic reset function shall recycle the signal system to a specified condition as a function of time or the state of the signalling system so that the warning device can sound again if the condition reappears.

aa. Control of volume. The volume (loudness) of an audio warning signal shall be designed to be controlled by the maintainer, the sensing mechanism, or both, depending upon the operational situation and personnel safety. Control movement shall be restricted to prevent reducing the volume to an inaudible level or increasing it to an unacceptable high level.

bb. Caution signal controls. Audio caution signals shall be provided with manual reset and volume controls.

cc. Duration. Audio warning signal duration shall be at least 0.5 sec, and may continue until the appropriate response is made. Completion of a corrective action by the maintainer or by other means shall automatically terminate the signal.

dd. Duration limitations. In an emergency situation, signals that persist or increase progressively in loudness shall not be used if manual shutoff may interfere with the corrective action required.

4.3.2.1 When to use auditory alarms and alerts.

ASTB, ACD-350 1993b) provides the following guideline:

Attention. Signals with high alerting capacities should be provided if the system or equipment imposes a requirement on the maintainer for concentration of attention. Such signals should not be so startling that they preclude appropriate responses or interfere with other functions by diverting attention away from other critical signals.

5. POSSIBLE RESEARCH TOPICS.

The literature review did not provide clear answers to all the questions about symbol design. Some questions with missing or incomplete answers that might benefit from additional research are:

a. What makes a symbol "good" from a user's point of view, and how can this goodness be specified and measured?

b. Which objects, concepts, and actions are better expressed as visual symbols, which as auditory symbols, and which as text? Under what conditions should more than one of these types be used simultaneously, for example, under what conditions should both a visual and auditory symbol be used?

c. To what extent does a symbol depend upon context for its meaning, that is, how independent and context-free can a symbol be? This includes the question of cultural dependence and context, that is, to what extent can symbols be used internationally?

d. To what extent can symbols of the same type (visual or auditory) be used in combinations? Can a visual symbol identifying a facility be combined with a visual symbol specifying its status? If symbols are combined, how should they be presented -- superimposed or adjacent, in the case of visual symbols; simultaneously or sequentially, in the case of auditory symbols?

e. If it is necessary to use alphanumeric information, such as a name or label, what value, if any, can a symbol add? For example, if it is necessary to specify the type and location of a radar system alphanumerically, is there any reason to present a radar symbol in addition?

GLOSSARY

ABSTRACT ICON. An attempt to visualize an object or concept in a way other than with a concrete image of the thing itself, for example, using a sheet of paper to represent a file.

ARBITRARY ICON. An icon that has no obvious reference to the thing it represents.

AUDITORY ICON. A nonverbal, familiar, everyday sound that intuitively conveys information about an object, for example, using a siren sound to represent an emergency vehicle (see also EARCON).

CODING. The application of an attribute (such as size, color, brightness, or flashing) to an object (such as a symbol or a block of text) to convey additional information (such as status) about the object.

EARCON. A simple, nonverbal sound arbitrarily assigned to represent an object; the auditory analogue of a visual or graphic icon (see also AUDITORY ICON).

GLYPH. A pictorial or graphic representation of an object; as used here, synonymous with ICON.

GRAPHIC. Written, printed, drawn, or engraved (not the presentation of data in graphs).

ICON. A simplified, pictorial representation of an object that resembles or suggests the object. The pictorial aspect is necessary; an icon "looks like" something. If unmodified, its meaning is restricted to the visual modality (see AUDITORY ICON and EARCON for the auditory modality). In ordinary English usage, it implies a highly representational image; originally, it had a strong religious connotation. The term "visual symbol" is often used for abstract or semi-abstract symbols. In computer science, the term "icon" has assumed a broad meaning that includes both representational images and visual symbols. "Representational icon" is not redundant in this terminology. Many computer interfaces intermix these types, so a single name for both is desirable. The term "icon" has assumed this meaning among computer users.

PICTOGRAPH. A sign that conveys a message pictorially.

REPRESENTATIONAL ICON. Typically, a simple picture of a familiar object or operation, for example a trash can or a printer.

SYMBOL. Something that represents or suggests something else by association, resemblance, or convention. A symbol may be

completely arbitrary in appearance, for example, use of the American flag to stand for the United States.

SYMBOLOGY. A system or set of symbols intended to support a particular application.

TIMBRE. The quality of a sound determined primarily by the fundamental frequency or frequencies and harmonics it contains and, to a lesser extent, by its amplitude envelope.

REFERENCES

Advanced Systems Technology Branch, ACD-350. (1992). MCS coding scheme literature review. Federal Aviation Administration Technical Center, Atlantic City International Airport, N.J. 08405.

Advanced Systems Technology Branch, ACD-350. (1993a). Guidelines for developing status symbology within Airway Facilities. Federal Aviation Administration Technical Center, Atlantic City International Airport, N.J. 08405.

Advanced Systems Technology Branch, ACD-350. (1993b). Preliminary human factors design standard for Airway Facilities. Federal Aviation Administration Technical Center, Atlantic City International Airport, NJ.

Baber, C., & Wankling, J. (1992). An experimental comparison of text and symbols for in-car reconfigurable displays. Applied Ergonomics, 23(4), 255-262.

Backs, R. W., & Walrath, L. C. (1992). Eye movement and pupillary response indices of mental workload during visual search of symbolic displays. Applied Ergonomics, 23(4), 243-254.

Baecker, R., Small, I., & Mander, R. (1991). Bringing icons to life. CHI '91 Conference Proceedings, 1-6.

Benbasat, I., & Todd, P. (1993). An experimental investigation of interface design alternatives: Icon vs. text and direct manipulation vs. menus. International Journal of Man-Machine Studies, 38, 369-402.

Birkmire, D. P., Karsh, R., Barnette, B. D., Pillalamarri, R., & DiBastiani, S. (1993). Eye movements in search and target acquisition. Proceedings of the Human Factors and Ergonomics Society 37th Annual Meeting, 1305-1309.

Blackwell, J. S., & Cuomo, D. L. (1991). Evaluation of a proposed space and missile warning symbology standard for graphical displays. Proceedings of the Human Factors Society 35th Annual Meeting, 102-106.

Blankenberger, B., & Hahn, K. (1991). Effects of icon design on human-computer interaction. International Journal of Man-Machine Studies, 35, 363-377.

Blattner, M. M., Sumikawa, D. A., & Greenberg, R. M. (1989). Earcons and icons: Their structure and common design principles. Human Computer Interaction, 4, 11-44.

Brewster, S. A., Wright, P. C., & Edwards, A. D. N. (1993). An evaluation of earcons for use in auditory human-computer interfaces. INTERCHI '93 Conference Proceedings, 222-227.

Brown, C. M. (1989). Human-computer interface design guidelines. Norwood, NJ: Ablex Publishing Corporation.

Byrne, M. D. (1993). Using icons to find documents: Simplicity is critical. INTERCHI '93 Conference Proceedings, 446-453.

Calhoun, G. L., Valencia, G., & Janson, W. P. (1988). Effectiveness of three-dimensional auditory directional cues -- in fighter cockpit. Proceedings of the Human Factors Society 32nd Annual Meeting, 68-72.

Commission Electrotechnique Internationale (International Electrotechnical Commission). (1973). Graphical symbols for use on equipment. International Standard CEI IEC 417.

Crawford, R. L., Toms, M. L., & Wilson, D. L. (1991). Effects of display luminance on the recognition of color symbols on similar color backgrounds. Proceedings of the Human Factors Society 35th Annual Meeting, 1466-1470.

Diethelm, W. (1972). Signet, signal, symbol: Handbook of international signs. Zurich: ABC Edition.

Dingus, T. A., Hathaway, J. A., & Hunn, B. P. (1991). A most critical warning variable: Two demonstrations of the powerful effects of cost on warning compliance. Proceedings of the Human Factors Society 35th Annual Meeting, 1034-1038.

Duncanson, J. P. (1994). Symbol Standardization in Airway Facilities, Part One: Investigating the State of the Art. CTA INCORPORATED. Advanced Systems Technology Branch, ACD-350, Federal Aviation Administration Technical Center, Atlantic City International Airport, N.J. 08405.

Edworthy, J., Loxley, S., & Dennis, I. (1991). Improving auditory warning design: Relationship between warning sound parameters and perceived urgency. Human Factors, 33(2), 205-231.

Erlichman, J. (1992). A pilot study of the In-Vehicle Safety Advisory and Warning System (IVSAWS) driver-alert warning system design. Proceedings of the Human Factors Society 36th Annual Meeting, 480-484.

Fisher, D. L., & Tanner, N. S. (1992). Optimal symbol set selection: A semiautomated procedure. Human Factors, 34(1), 79-95.

Folds, D. J. (1990). Advanced audio displays in aerospace systems -- Technology requirements and expected benefits. Proceedings of the IEEE National Aerospace and Electronics Conference, 2, 739-743.

Friedmann, K. (1988). The effect of adding symbols to written warning labels on user behavior and recall. Human Factors, 30(4), 507-515.

Gaver, W. W. (1986). Auditory icons: Using sound in computer interfaces. Human Computer Interaction, 2, 167-177.

----- (1989). The SonicFinder: An interface that uses auditory icons. Human Computer Interaction, 4, 67-94.

- (1993). Synthesizing auditory icons. INTERCHI '93 Conference Proceedings, 228-235.

Geiselman, R. E., Landee, B. M., & Christen, F. G. (1990). Perceptual discriminability as a basis for selecting graphic symbols. In M. Verturino. (ed). Selected readings in human factors. Santa Monica, CA: Human Factors and Ergonomics Society.

Gerth, J. M. (1991). Identifying the temporal patterning confusions in audio displays: Developing guidelines for complex audio displays. Proceedings of the Human Factors Society 35th Annual Meeting, 723-727.

- (1992). Systematic errors occur in the discrimination of complex audio signals. Proceedings of the Human Factors Society 36th Annual Meeting, 263-267.

----- (1993). Identification of sounds with multiple timbres. Proceedings of the Human Factors and Ergonomics Society 37th Annual Meeting, 539-543.

Gittins, D. (1986). Icon-based human-computer interaction. International Journal of Man-Machine Studies, 24, 519-543

Goldstein, E. B. (1989). Sensation and Perception, 3rd Edition. Belmont, CA: Wadsworth Publishing Company.

Green, P., & Pew, R. W. (1978). Evaluating pictographic symbols: An automotive application. Human Factors, 20(1), 103-114.

Guastello, S. J., Traut, M., & Korienek, G. (1989). Verbal versus pictorial representations of objects in a human-computer interface. International Journal of Man-Machine Studies, 31, 99-120.

Haas, E. C. (1992). A pilot study on the perceived urgency of multitone and frequency-modulated warning signals. Proceedings of the Human Factors Society 36th Annual Meeting, 248-252.

Haas, E. C., & Casali, J. G. (1993). The perceived urgency and detection time of multi-tone and frequency-modulated warning signals. Proceedings of the Human Factors and Ergonomics Society 37th Annual Meeting, 544-548.

Hakkinen, M. T., & Williges, B. H. (1984). Synthesized warning messages: Effects of an alerting cue in single- and multiple-function voice synthesis systems. Human Factors, 26(2), 185-195.

Hellier, E. J., Edworthy, J., & Dennis, I. (1993). Improving auditory warning design: Quantifying and predicting the effects of different warning parameters on perceived urgency. Human Factors, 35(4), 693-706.

Hovey, C. K., & Berson, B. L. (1987). Evaluation of ASW tactical symbology. Proceedings of the Human Factors Society 31st Annual Meeting, 1403-1407.

Howard, C., O'Boyle, M. W., Eastman, V., Andre, T., & Motoyama, T. (1991). The relative effectiveness of symbols and words to convey photocopier functions. Applied Ergonomics, 22(4), 218-224.

International Organization for Standardization. (1969). Graphical symbols for use on equipment -- Index and synopsis International Standard ISO 7000.

International Organization for Standardization. (1976). Graphic symbols -- General principles for presentation. International Standard ISO 3461.

Jensen, D. G. (1993). Space Station Freedom aural toxic alarm development. Proceedings of the Human Factors and Ergonomics Society 37th Annual Meeting, 52-54.

Kalsher, M. J., Clarke, S. W., & Wogalter, M. S. (1991). Posted warning placard: Effects on college students knowledge of alcohol facts and hazards. Proceedings of the Human Factors Society 35th Annual Meeting, 456-460.

Kern, J. E. 1994). Unpublished symbols

King, R. A., & Corso, G. M. (1993). Auditory displays: If they are so useful, why are they turned off? Proceedings of the Human Factors and Ergonomics Society 37th Annual Meeting, 549-553.

Kirkpatrick, M., Dutra, L. A., Lyons, R. A., Osga, G. A., & Pucci, J. J. (1992). Tactical symbology standards. Proceedings of the Human Factors Society 36th Annual Meeting 1087-1091.

Kline, D. W., & Fuchs, P. (1993). The visibility of symbolic highway signs can be increased among drivers of all ages. Human Factors, 35(1), 25-34.

Kline, T. J. B., Ghali, L. M., & Kline, D. W. (1990). Visibility distance of highway signs among young, middle-aged, and older observers: Icons are better than text. Human Factors, 32(5), 609-619.

Kolers, P. A. (1969). Some formal characteristics of pictograms. American Scientist, 57, 348-363.

Lin, R. (1992). An application of the semantic differential to icon design. Proceedings of the Human Factors Society 36th Annual Meeting, 336-340.

Lin, R., & Kreifeldt, J. G. (1992). Understanding the image functions for icon design. Proceedings of the Human Factors Society 36th Annual Meeting, 341-345.

MacGregor, J. N. (1992). A comparison of the effects of icons and descriptors in videotex menu retrieval. International Journal of Man-Machine Studies, 37, 767-777.

Marcus, A. (1992). Graphic design for electronic documents and user interfaces. New York, NY: ACM Press.

Osborne, D. W., & Huntley, M. S., Jr. (1992). Design of instrument approach procedure charts: Comprehending missed approach instructions coded in text or icons. Proceedings of the Human Factors Society 36th Annual Meeting, 1455-1459.

Remington, R., & Williams, D. (1986). On the selection and evaluation of visual display symbology: Factors influencing search and identification times. Human Factors, 28(4), 407-420.

Schmidt, J. K., & Kern, J. E. (1991). A literature review on tactical map symbology. Air Vehicle and Crew Systems Technology Department, Naval Air Development Center, Warminster, PA. (Draft)

Sorkin, R. D. (1987). Design of auditory and tactile displays. In G. Salvendy (Ed. , Handbook of human factors, 549-576. New York: Wiley.

Stammers, R. B., & Hoffman, J. (1991). Transfer between icon sets and ratings of icon concreteness and appropriateness. Proceedings of the Human Factors Society 35th Annual Meeting, 354-358.

Strijland, P. (1992). Icon design. ACM Conference on Human Factors in Computing Systems. Tutorial Notes, CHI '92, Monterey, CA.

Thome, R. J. (1991). The effect of type of task, degree of integration, and modality on the performance of concurrent tasks. Proceedings of the Human Factors Society 35th Annual Meeting, 67-71.

Van Orden, K. F., & DiVita, J. (1993). Highlighting with flicker. Proceedings of the Human Factors and Ergonomics Society 37th Annual Meeting, 1300-1304.

Van Orden, K. F., DiVita, J., & Shim, M. J. (1993). Redundant use of luminance and flashing with shape and color as highlighting codes in symbolic displays. Human Factors, 35(2), 195-204.

Vukelich, M., & Whitaker, L. A. (1993). The effects of context on the comprehension of graphic symbols. Proceedings of the Human Factors and Ergonomics Society 37th Annual Meeting, 511-515.

Waterworth, J. A., Chignell, M. H., & Zhai, S. M. (1993). From icons to interface models: Designing hypermedia from the bottom up. International Journal of Man-Machine Studies, 39, 453-472.

Wogalter, M. S., Kalsher, M. J., & Racicot, B. M. (1992). The influence of location and pictorials on behavioral compliance to warnings. Proceedings of the Human Factors Society 36th Annual Meeting, 1029-1033.

Woodson, W. E., & Conover, D. W. (1964). Human engineering guide for equipment designers. Berkeley, CA: University of California Press.

APPENDIX A

ANNOTATED BIBLIOGRAPHY

Advanced Systems Technology Branch, ACD-350. (1992). MCS coding scheme literature review. Federal Aviation Administration Technical Center, Atlantic City International Airport, N.J. 08405.

The focus of the activity presented in this report was the identification of the elements in the MCS Design Model that should be coded. That is, the information which would be more easily and quickly understood by symbolic representation, such as icons, text, color, and auditory symbols, etc., is identified. Elements that are candidates for symbolic representation include: states, facility control information, scales for parameter values and limits, user interface controls, alarm status and control information, facility conditions, and feedback acknowledging user actions.

Advanced Systems Technology Branch, ACD-350. (1993a). Guidelines for developing status symbology within Airway Facilities. Federal Aviation Administration Technical Center, Atlantic City International Airport, N.J. 08405.

The objective of this report is to provide guidelines for developing current and future AF status symbology. These guidelines include a brief discussion of the symbology development process, design recommendations, and a discussion on applying the guidelines.

Advanced Systems Technology Branch, ACD-350. (1993b). Preliminary human factors design standard for Airway Facilities. CTA Incorporated, Pleasantville, NJ. Contract Number DTFA03-86-C-00018. Advanced Systems Technology Branch, ACD-350. Federal Aviation Administration Technical Center, Atlantic City International Airport, NJ.

This document is intended to provide compliance information governing the design of new Federal Aviation Administration systems and equipment that will be acquired and maintained by Airway Facilities.

Baber, C. and Wankling, J. (1992). An experimental comparison of text and symbols for in-car reconfigurable displays. Applied Ergonomics, 23(4), 255-262.

As cars become more sophisticated, there is a corresponding increase in the range of information which needs to be presented to drivers. This increased information load is

beyond the scope of conventional technology; space requirements on the dashboard limit the number of displays which can be employed. The study reported in this paper is based on the premise that centralized information presentation could be employed for in-car warning systems. Forty subjects were tested on a computer-based task, in which they had to rate the urgency of displayed information. The information was presented in one of four formats: symbols only, symbols plus title, symbols plus action, or symbol, title and action. Results show that the most effective type of display for urgency rating was symbol plus action, and that symbol, title and action produced the fastest reaction times. These results are considered in terms of previous research, and implications for the design of in-car warnings are discussed.

Backs, R. W. and Walrath, L. C. (1992). Eye movement and pupillary response indices of mental workload during visual search of symbolic displays. Applied Ergonomics, 23(4), 243-254.

Eye movement and pupillary response measures (in addition to search time and accuracy) were collected as indices of visual workload during two experiments designed to evaluate the addition of color coding to a symbolic tactical display. Displays also varied with regard to symbol density and the type of information participants were required to abstract from the display. These variables were factorially manipulated to examine the effects of color coding in conditions of varying difficulty. In Experiment 1 (n = 8), search time and the number of eye fixations were affected by all variables and in a similar manner; fixation dwell time and the pupillary response dissociated from the other measures. Compared to monochrome displays, color coding facilitated search (reduced search time, but not accuracy) during exhaustive search, but had no effect during self-terminating search. Experiment 2 (n = 8) was a replication of Experiment 1 with a pseudo-search control condition added to examine further the pupillary response measures : In particular, to assess the effects of the physical parameters of the displays, and to verify the findings of Experiment 1. Pupillary response measures were sensitive to the information processing demands of the search task, not merely to the physical parameters of the display. Further, the search time, accuracy, and eye movement results from the active search condition generally replicated Experiment 1, but the fixation dwell time data did not. These between-study differences were interpreted as indicating the importance of participant search strategy.

Baecker, R., Small, I., and Mander, R. (1991). Bringing icons to life. CHI '91 Conference Proceedings, 1-6.

Icons are used increasingly in interfaces because they are compact "universal" pictographic representations of computer functionality and processing. *Animated icons* can bring to life symbols representing complete applications or functions within an application, thereby clarifying their meaning, demonstrating their capabilities, and even explaining their method of use. To test this hypothesis, we carried out an iterative design of a set of animated painting icons that appear in the HyperCard tool palette. The design discipline restricted the animations to 10 to 20 second sequences of 22x20 pixel bit maps. User testing was carried out on two inferfaces - one with static icons, one with the animated icons. The results showed significant benefit from the animations in clarifying the purpose and functionality of the icons.

Benbasat, I. and Todd, P. (1993). An experimental investigation of interface design alternatives: Icon vs. text and direct manipulation vs. menus. International Journal of Man-Machine Studies, 38, 369-402.

This paper reports on two experiments which examine the effects of iconic and direct manipulation interfaces on the performance of causal users using an electronic mail system. There are two key aspects to these experiments. First, they have been carefully designed to separate the effect of iconic representation from that of direct manipulation in order to examine the independent effect of each as well as their joint effect. Second, subjects performed the same experimental task three different times over 1 week, thus allowing for the effects of icons and direct manipulation interfaces to be assessed over repeated trials. Each experiment measured time taken and errors made in task completion as dependent variables.

Results indicate that there were no advantages associated with iconic representations compared to text-based representations of actions and objects. Subjects working with direct manipulation interfaces completed the task faster than those with menu-based interfaces. However, this difference in time was not significant when the task was repeated for a third time, indicating that the benefits to direct manipulation might diminish after a learning period No interface was better than others in terms of reducing error rates when interacting with the computer system.

Birkmire, D. P., Karsh, R., Barnette, B. D., Pillalamarri, R., and DiBastiani, S. (1993). Eye movements in search and target acquisition. Proceedings of the Human Factors and Ergonomics Society 37th Annual Meeting, 1305-1309.

The frequency distribution of eye fixations and fixation durations during a search and target acquisition task was examined to determine if the allocation of visual attention was related to target, scene, and/or observer characteristics. Ninety computer-generated scenes simulating infrared imagery and containing different levels of clutter and zero, one, two, or three targets were produced. Targets were embedded in these scenes counterbalancing for range and position. Global and local clutter were measured using both statistical variance and probability of edge metrics. Thirty-three aviators, tankers, and infantry soldiers were shown still video images of the 90 scenes and were instructed to search for targets. Results of multiple regression analyses of global clutter, local clutter, range, number of targets, target dimensions, target complexity, and group membership on eye fixations and fixation durations are given and discussed in terms of search strategies.

Blackwell, J. S. and Cuomo, D. L. (1991). Evaluation of a proposed space and missile warning symbology standard for graphical displays. Proceedings of the Human Factors Society 35th Annual Meeting, 102-106.

A discriminability evaluation was performed on a proposed Space and Missile Warning symbol set. Our analysis focused on the discriminability of the symbols and the application of the information coding techniques. Inconsistent or inappropriate use of coding techniques can affect a user's interpretation of the symbol's intended meaning. Potential problems included the similarity of individual symbols, use of alphanumeric markers, and partially shaded symbols, and the lack of guidance on the minimum size of the symbols. After a lengthy review of previous research, we felt the literature could not provide adequate solutions. A two-part discriminability study was conducted to test the overall effects of the information coding techniques on discriminability, to identify individual symbols with low discriminability, and to determine an appropriate minimum size for these symbols. Search time was used as a measure of symbol discriminability. Size, shape, markers, and shading had significant effects on search time and errors. The experiments confirmed the suspected discriminability problems and modifications were made to the existing symbol set to create three new alternative symbol sets. Testing performed on these new symbol sets revealed that many of the problem areas from the original symbol set had been

improved. Design guidelines and a new modified symbol set were proposed for review by the operational community.

Blankenberger, B. and Hahn, K. (1991). Effects of icon design on human-computer interaction. International Journal of Man-Machine Studies, 35, 363-377

After subjects practiced using a pointing device (two-button mouse) for selecting icons on a computer screen, the effect of "articulatory distance" (i.e. the difference between a picture and its meaning) on performance in menu-selection tasks was analyzed. Three icon sets with different articulatory distances and one text set were constructed, validated and tested in a "search and select" experiment with icon positions randomized on the screen. This was contrasted with an experiment in which icons were to be selected from fixed screen positions. Results indicate that articulatory distance indeed had an effect on reaction time in the first design, but not in the latter. A recognition task was finally given to decide whether articulatory distance could influence memory for icons. The fact that subjects were able to recode icon meanings to screen positions after some training backs the everyday experience that icon design seems to be of little influence on the performance of advanced users. Icon-oriented interfaces are aimed, however, at the computer novice.

Blattner, M. M., Sumikawa, D. A., and Greenberg, R. M. (1989). Earcons and icons: Their structure and common design principles. Human Computer Interaction, 4, 11-44.

In this article, we examine *earcons*, which are audio messages used in the user-computer interface to provide information and feedback to the user about computer entities. (Earcons include messages and functions, as well as states and labels.) We identify some design principles that are common to both visual symbols and auditory messages, and discuss the use of representational and abstract icons and earcons. We give some examples of audio patterns that may be used to design modules for earcons, which then may be assembled into larger groupings called *families*. The modules are single pitches or rhythmicized sequences of pitches called *motives*. The families are constructed about related motives that serve to identify a family of related messages. Issues concerned with learning and remembering earcons are discussed.

Brewster, S. A., Wright, P. C., and Edwards, A. D. N. (1993). An evaluation of earcons for use in auditory human-computer interfaces. INTERCHI '93 Conference Proceedings, 222-227.

An evaluation of earcons was carried out to see whether they are an effective means of communicating information in sound. An initial experiment showed that earcons were better than unstructured bursts of sound and that musical timbres were more effective than simple tones. A second experiment was then carried out which improved upon some of the weaknesses shown up in Experiment 1 to give a significant improvement in recognition. From the results of these experiments some guidelines were drawn up for use in the creation of earcons. Earcons have been shown to be an effective method for communicating information in a human-computer interface.

Byrne, M. D. (1993). Using icons to find documents: Simplicity is critical. INTERCHI '93 Conference Proceedings, 446-453.

A common task at almost any computer interface is that of searching for documents, which GUIs typically represent with icons. Oddly, little research has been done on the processes underlying icon search. This paper outlines the factors involved in icon search and proposes a model of the process. An experiment was conducted which suggests that the proposed model is sound, and that the most important factor in searching for files is the type of icons used. In general, simple icons (those discriminable based on a few features) seem to help users, while complex icons are no better than simple rectangles.

Calhoun, G. L., Valencia, G., and Janson, W. P. (1988). Effectiveness of three-dimensional auditory directional cues -- in fighter cockpit. Proceedings of the Human Factors Society 32nd Annual Meeting, 68-72.

Natural aural directional cuing in the cockpit should relieve the demands placed on the visual modality, reduce display clutter and alleviate cognitive attention needed to process and extract meaning from coded formats. This experiment compared the effectiveness of three-dimensional (3-D) auditory cues to conventional visual and auditory methods of directing visual attention to peripheral targets. Five directional cues were evaluated: visual symbol, coded aural tone, speech cue, 3-D tone (white noise appearing to emanate from peripheral locations) and 3-D speech (speech cue appearing to emanate from peripheral locations). The results showed significant performance differences as a function of directional cue type in peripheral target task completion time, as well as eye and head reaction time.

Results, such as these, will help improve the application of directional sound in operational cockpits.

Commission Electrotechnique Internationale (International Electrotechnical Commission). (1973). Graphical symbols for use on equipment. International Standard CEI IEC 417.

This document illustrates hundreds of symbols and titles for use on equipment.

Crawford, R. L., Toms, M. L., and Wilson, D. L. (1991). Effects of display luminance on the recognition of color symbols on similar color backgrounds. Proceedings of the Human Factors Society 35th Annual Meeting, 1466-1470.

This study examined the effects of display luminance on the ability of human observers to recognize color symbols displayed against similar or color backgrounds. The Signal Detection paradigm was utilized and subject sensitivity, as measured by d', was the primary measure of interest. The symbol colors were red, green, and blue. Background colors were .01 to .07 1976 CIE/UCS units distant from the symbol color. Luminance levels ranged from 11.85 cd/m^2 to 127.25 cd/m^2. The symbols were presented on a cathode ray tube (CRT) under ambient lighting of two lux. Display luminance was found to affect subject sensitivity, d', as a function of symbol-background color combination. The results imply that display luminance for the presentation of blue symbology on bluish backgrounds is optimal at 19 cd/m^2. For the red and green symbol-background conditions, display luminance between 56 and 93 cd/m^2 yields the best performance.

Edworthy, J., Loxley, S., and Dennis, I. (1991). Improving auditory warning design: Relationship between warning sound parameters and perceived urgency. Human Factors, 33(2), 205-231.

This paper presents an experimental study of the effects of individual sound parameters on perceived (psychoacoustic) urgency. Experimental Series 1 showed that fundamental frequency, harmonic series, amplitude envelope shape, and delayed harmonics all have clear and consistent effects on perceived urgency. Experimental Series 2 showed that temporal and melodic parameters such as speed, rhythm, pitch range, and melodic structure also have clear and consistent effects on perceived urgency. The final experiment tested a set of 13 auditory warnings generated by an application of the earlier experimental findings. The urgency rank ordering of this warning set was predicted, and the correlation between the predicted and the obtained order was highly significant. The results of these experiments have a

widespread application in the improvement of existing auditory warning systems and the design of new systems, where the psychoacoustic and psychological appropriateness of warnings could be enhanced.

Erlichman, J. (1992). A pilot study of the In-Vehicle Safety Advisory and Warning System (IVSAWS) driver-alert warning system design. Proceedings of the Human Factors Society 36th Annual Meeting, 480-484.

This pilot study was conducted to obtain preliminary information regarding alternative signalling presentations and symbologies for the Driver-Alert Warning System design within the In-Vehicle Safety Advisory and Warning System Program sponsored by the Federal Highway Administration. Preliminary analysis had been conducted by both Hughes Aircraft Company and The University of Michigan Transportation Research Institute. This pilot study concentrated on the driver attributes of understanding, relative effectiveness and signalling format. Thirteen subjects were exposed to the new pictograms prototyped on a Macintosh computer and were requested to verbalize their understanding and preferences in regard to varying signalling characteristics. These characteristics included, a) monochrome, b) color, c) blink, d) tone, e) text message and f) voice message. The results indicated that, as a group, the combination of color, audio tone, text and voice message was the preferred signalling presentation. Gender differences were noted with the female subjects indicating a preference for the combination that included color and blink. All pictograms were recognizable by the subjects and all subjects agreed that IVSAWS would be a substantial aid to the driver.

Fisher, D. L. and Tanner, N. S. (1992). Optimal symbol set selection: A semiautomated procedure. Human Factors, 34(1), 79-95.

A new model of the visual search process is developed which can improve the design of large symbol sets such as those used by nuclear power plant personnel, air traffic controllers, and battlefield troops. An experiment was conducted to determine whether the new, componential model or an already existing, discriminability model better explains visual search behavior. The results were consistent with the componential model. We show how to use the componential model to help automate selection of the optimal symbol set (i.e., the symbol set that minimizes the average time to find a target).

Folds, D. J. (1990). Advanced audio displays in aerospace systems -- Technology requirements and expected benefits. Proceedings of the IEEE National Aerospace and Electronics Conference, Vol. 2, 739-743.

It is suggested that multiple concurrent audio signals could be used to represent the status of various onboard subsystems and processes, as well as external objects (e.g., threats). Perceptually, these sounds could remain in the background, continuously available to the operator, without causing annoyance or interference with usual voice communications. Research at Georgia Tech is examining the effectiveness of various types of complex, nonspeech sounds for conveying information in the background and its resulting impact on operator performance. The use of steady-state sounds to represent the in-bounds versus out-of-bounds status of one to four concurrent processes was examined. Two conditions were compared: A visual-only condition in which process states were represented solely by visual indicators, and an audiovisual condition in which the visual indicators were supplemented by continuous audio signals. Subjects in the audiovisual condition consistently responded faster and more accurately, committed fewer false alarms, and rated their workload lower than subjects in the visual-only condition.

Friedmann, K. (1988). The effect of adding symbols to written warning labels on user behavior and recall. Human Factors, 30(4), 507-515.

The use of consumer products by 144 college students was studied to determine the effects that (a) adding symbols to written warnings, (b) subjects' familiarity with the product, and (c) type of hazard would have on their noticing, recalling, or complying with the warning. Subjective ratings of perceived confidence, hazardousness, likelihood of injury, and severity of injury were also collected. Across all behavioral measures there was a steady decline in the number of subjects who first noticed (88%), then read (46%), and finally followed the warning (27%). In some conditions, however, compliance levels were as high as 42%. Symbols added to written warning labels did not significantly increase levels of compliance. A significant positive relationship was found between the perceived hazardousness of the product and reading, following, and recalling the warning. Factors affecting the user's motivation to read and follow on-product warnings were also noted.

Gaver, W. W. (1986). Auditory icons: Using sound in computer interfaces. Human Computer Interaction, 2, 167-177.

There is growing interest in the use of sound to convey information in computer interfaces. The strategies employed thus far have been based on an understanding of sound that leads to either an arbitrary or metaphorical relation between the sounds used and the data to be represented. In this article, an alternative approach to the use of sound in computer interfaces is outlined, one that emphasizes the role of sound in conveying information about the world to the listener. According to this approach, auditory icons, caricatures of naturally occurring sounds, could be used to provide information about sources of data. Auditory icons provide a natural way to represent dimensional data as well as conceptual objects in a computer system. They allow categorization of data into distinct families, using a single sound. Perhaps the most important advantage of this strategy is that it is based on the way people listen to the world in their everyday lives.

Gaver, W. W. (1989). The SonicFinder: An interface that uses auditory icons. Human Computer Interaction, 4, 67-94.

The appropriate use of nonspeech sounds has the potential to add a great deal to the functionality of computer interfaces. Sound is a largely unexploited medium of output, even though it plays an integral role in our everyday encounters with the world, a role that is complementary to vision. Sound should be used in computers as it is in the world, where it conveys information about the nature of sound producing events. Such a strategy leads to auditory icons, which are everyday sounds meant to convey information about computer events by analogy with everyday events. Auditory icons are an intuitively accessible way to use sound to provide multidimensional, organized information to users.

These ideas are instantiated in the SonicFinder, which is an auditory interface I developed at Apple Computer, Inc. In this interface, information is conveyed using auditory icons as well as standard graphical feedback. I discuss how events are mapped to auditory icons in the SonicFinder, and illustrate how sound is used by describing a typical interaction with this interface.

Two major gains are associated with using sound in this interface: an increase in direct engagement with the model world of the computer and an added flexibility for users in getting information about that world. These advantages seem to be due to the iconic nature of the mappings used between sound and the information it is to convey. I discuss sound

effects and source metaphors as methods of extending auditory icons beyond the limitations implied by literal mappings, and I speculate on future directions for such interfaces.

Gaver, W. W. (1993). Synthesizing auditory icons. INTERCHI '93 Conference Proceedings, 228-235.

Auditory icons add valuable functionality to computer interfaces, particularly when they are parameterized to convey dimensional information. They are difficult to create and manipulate, however, because they usually rely on digital sampling techniques. This paper suggests that new synthesis algorithms, controlled along dimensions of events rather than those of the sounds themselves, may solve this problem. Several algorithms, developed from research on auditory event perception, are described in enough detail here to permit their implementation. They produce a variety of impact, bouncing, scraping, and machine sounds. By controlling them with attributes of relevant computer events, a wide range of parameterized auditory icons may be created.

Geiselman, R. E., Landee, B. M., and Christen, F. G. (1990). Perceptual discriminability as a basis for selecting graphic symbols. In Verturino, M. (ed). Selected readings in human factors. Santa Monica, CA: Human Factors and Ergonomics Society.

The purpose of this research was to develop a performance-based criterion for selecting among alternative symbols to be used in graphic displays. The specific criterion developed was an index of perceptual discriminability. Through regression analyses of an intersymbol similarity-rating matrix, it was concluded that symbols are judged more or less similar on the basis of the number of shared versus unique configural attributes (an X, a triangle, etc.), as opposed to primitive attributes (number of lines, arcs, etc.). An easy-to-use discriminability-index formula was derived from the regression analysis involving the configural attributes, and this formula was used to predict the results of an experiment involving a search for specific symbols embedded in an array. Indices obtained from a formula such as the one developed here could be used as part of the basis for choosing among alternative candidate symbols for inclusion in an existing symbol domain.

Gerth, J. M. (1991). Identifying the temporal patterning confusions in audio displays: Developing guidelines for complex audio displays. Proceedings of the Human Factors Society 35th Annual Meeting, 723-727.

Current guidelines and recommendations for auditory displays suggest that human auditory discrimination performance is limited and that auditory displays should be used only for alarm and alerting signals. Auditory warnings are likely to be confused even when their spectra are very different. Reducing confusion between warnings should increase the number of auditory signals which can be presented. The present research investigated the ability of human listeners to discriminate sounds varying in temporal patterning in several sound categories.

Although overall accuracy was 92 percent across the 45 dissimilar sound sequences, 7 sequences were found to be easily confused and accounted for 64 percent of the total errors made by listeners, regardless of sound category. According to subject reports, multiple simultaneously presented temporally patterned sounds within each sound category were not perceived as multiple sources but rather were fused into a single complex temporal pattern. Implications for developing complex audio displays by increasing the number and complexity of sounds and planned continuing research are also discussed.

Gerth, J. M. (1992). Systematic errors occur in the discrimination of complex audio signals. Proceedings of the Human Factors Society 36th Annual Meeting, 263-267.

Previous research suggests that the temporal pattern of dissimilar sounds may be a basis for confusion. To extend this research, the present study used complex sounds formed by simultaneously playing components drawn from four sound categories. Four temporal patterns, determined by sound duration and duty cycle were also used, producing a total of 16 basic components. The density (i.e., number of components played simultaneously) ranged from one to four. Subjects heard a sequence of two complex sounds and judged whether they were same or different. For trials in which the sounds differed, there were three possible manipulations: the addition of a component, the deletion of a component, and the substitution of one component for another.

Overall accuracy was 94 percent across the 144 dissimilar sound complexes. As density increased, a significantly greater number of errors occurred for all classes of manipulations. Changes in individual temporal patterns across a variety of manipulations of sounds involving

adding, deleting and substituting components were accurately discriminated. Subjects were least accurate in detecting substitutions of a pattern. A single sound category was identified in error prone sequences which was most often involved as the changing component from first to second sound presentation. Suggestions for the design of easily discriminated sounds are discussed.

Gerth, J. M. (1993). Identification of sounds with multiple timbres. Proceedings of the Human Factors and Ergonomics Society 37th Annual Meeting, 539-543.

The present research examined identification of complex sounds created by simultaneously playing two or more component sounds in various combinations. Sixteen component sounds were used, created by imposing four distinct temporal patterns on four basic timbres, two musical timbres and two complex real-world timbres. In the present experiment, complex sounds were created by simultaneously playing one to four component sounds, each with a different timbre. Subjects heard a complex sound, followed by a second complex sound that always differed from the first by adding a component, deleting a component or substituting a component. Subjects indicated which component had been added, deleted, or substituted. Sound changes were identified with moderate accuracy (above 60 percent). The errors committed varied with temporal pattern, timbre, sound change and density. The analyses of identification confusions indicated that subjects identified the correct timbre of the sound change even when temporal patterning was confused. The finding that temporal patterns were confused largely within the sound category of the correct response limits the previous interpretation of other research, which found that similar temporal patterns are confusable even with differences in spectra. Results of the present investigation suggest that multiple, temporal patterns with varying timbres can be presented from a single physical location to convey a change in state or status of an informative sound source. Design contributions of the present research to auditory information systems such as virtual reality are discussed. For such an application, a combination of physical separation and multiple patterns with varying timbres could provide a coherent, yet informationally complex, auditory display.

Gittins, D. (1986). Icon-based human-computer interaction. International Journal of Man-Machine Studies, 24, 519-543.

This paper is concerned with the use of icons in human-computer interaction (HCI). Icons are pictographic representations of data or processes within a computer system, which have been used to replace commands and menus

as the means by which the computer supports a dialogue with the end-user. They have been applied principally to graphics-based interfaces to operating systems, networks and document-processing software.

The paper attempts to provide a more systematic treatment of icon interfaces than has hitherto been made, and to create a classification which it is hoped will be of use to the dialogue designer. The characteristics, advantages and disadvantages of icon-based dialogues are described. Metaphors, design alternatives, display structures and implementation factors are discussed, and there is a summary of some icon design guidelines drawn from a variety of sources. Some mention is also made of attempts by researchers to measure the effectiveness of icon designs empirically.

Goldstein, E. B. (1989). Sensation and Perception, 3rd Edition. Belmont, CA: Wadsworth Publishing Company.

Discusses the Gestalt laws of organization, figure-ground separation, and properties influencing the perception of figures.

Green, P. and Pew, R. W. (1978). Evaluating pictographic symbols: An automotive application. Human Factors, 20(1), 103-114.

Fifty university students participated in a laboratory experiment which examined 19 pictographic symbols previously used or proposed for labelling automobile controls and displays. Association norms, measures of familiarity, and magnitude estimates of the symbols' communicativeness were collected. Twenty of these subjects also participated in a paired-associate learning task and a two-alternative, forced-choice reaction-time task in which they made same-different judgments in response to verbally presented symbol labels followed by visually presented pictograms. It was found that, in general, the relative order of merit for the individual symbols was not consistent across tasks. Specifically, ratings of communicativeness were found to be well correlated with associative strength and to a lesser extent with reaction time, but associative strength was only weakly correlated with reaction time. Ease of learning was found to be an independent issue.

Guastello, S. J., Traut, M., and Korienek, G. (1989). Verbal versus pictorial representations of objects in a human-computer interface. International Journal of Man-Machine Studies, 31, 99-120.

Computer systems often use icons to represent objects of interest within the system. In this study we tested four hypotheses concerning the relative effectiveness of icon construction: (1) Pictorial icons would be rated as more meaningful than verbal icons for concrete objects. (2) Ratings of meaningfulness would be dependent upon qualities of icons such as long versus short abbreviations, and industry standard versus enhanced pictogram. (3) Ratings would be dependent upon experience with the content domains. (4) Icons composed of both verbal and pictorial elements would be rated as more meaningful than icons composed of verbal or pictorial elements only. Hypotheses were developed from literature on text formatting, command names, human memory functionality, population stereotypes, and brain lateralization. Two experiments were conducted. The first involved icons for objects found in a building automation system (BAS) environment which were rated by 187 system operators. The second experiment involved icons from BAS, engineering, computer systems and finance environments as rated by 139 undergraduates with varying experience in those content domains. Results overall showed that: (1) Mixed modality icons were rated as distinctively more meaningful than alternatives. (2) Ratings were occasionally bolstered by population stereotypes acquired through experience. (3) Long abbreviations are preferable to short ones. (4) It is possible to construct pictograms that are more meaningful than industry standards, and (5) verbal icons are sometimes preferred over pictorial icons when mixed modes are not available.

Haas, E. C. (1992). A pilot study on the perceived urgency of multitone and frequency-modulated warning signals. Proceedings of the Human Factors Society 36th Annual Meeting, 248-252.

In some environments, there is a serious mismatch between the perceived (psychoacoustic) urgency of a warning and its situational urgency. This pilot study investigated effect of pulse format, pulse duration, and time between pulses on the perceived urgency of warning signals. The intent was to determine the best combination of variables and levels of variables to use in a formal study on the perceived urgency of warning signals. The results indicated that only pulse format and time between pulses were significant. Subjects rated sequential pulses as being less urgent than any other format. Signals with shorter inter-pulse intervals were rated as significantly more urgent. Pulse format and time

between pulses were determined to be variables which should be used in future research.

Haas, E. C. and Casali, J. G. (1993). The perceived urgency and detection time of multi-tone and frequency-modulated warning signals. Proceedings of the Human Factors and Ergonomics Society 37th Annual Meeting, 544-548.

The perceived urgency and detectability of auditory warning signals are important safety considerations. When designed correctly, auditory warning signals can improve performance and reduce accidents. However, in some environments, there is a serious mismatch between the perceived (psychoacoustic) urgency of a warning and its situational urgency. In addition, many auditory warnings are not detectable within their environments. This research examined several prominent pulse parameters which affect the perceived urgency and detection time of auditory warning signals. These elements included pulse format (sequential, simultaneous, and sawtooth frequency-modulated pulses), pulse level (65 dBC and 79 dBC), and time between pulses (0 ms, 150 ms, and 300 ms). The environments of interest were those settings with steady-state broadband machinery noise. A loading task presented additional demands during the signal detection task. Free-modulus magnitude estimation and the method of paired comparisons quantified perceived urgency. Simple reaction time measured signal detectability and signal effects were analyzed using a multivariate approach.

Results indicated that detection time decreased as perceived urgency increased. The higher the pulse level, the greater the perceived urgency of the signal and shorter the detection time. Sequential signals were rated as less urgent than the other pulse formats, and subjects took longer to detect their occurrence. Under most conditions, there was no significant difference in the perceived urgency or detection time of simultaneous and frequency-modulated pulses. Time between pulses (inter-pulse interval) affected only perceived urgency, not detection time. The shorter the time between pulses, the greater the perceived urgency of the signal.

Hakkinen, M. T. and Williges, B. H. (1984). Synthesized warning messages: Effects of an alerting cue in single- and multiple-function voice synthesis systems. Human Factors, 26(2), 185-195.

The present study examined the effectiveness of preceding synthesized voice warning messages with an alerting cue as a function of the amount of information presented by the voice synthesizer and the workload level in the primary task.

Subjects performed a simplified air traffic control task in which they were required to monitor two visual displays and to enter commands via a standard keyboard. Emergency messages were always presented by phoneme-based synthesized speech. However, the presence of an alerting cue (light and tone) prior to emergency messages and the presentation mode of noncritical messages (visual or auditory) were varied experimentally. When synthesized speech was used only for emergency messages, the presence of an alerting cue lengthened the response time to the message. However, when computer-generated speech was used for multiple functions, more emergency messages were detected when an alerting cue was used.

Hellier, E. J., Edworthy, J., and Dennis, I. (1993). Improving auditory warning design: Quantifying and predicting the effects of different warning parameters on perceived urgency. Human Factors, 35(4), 693-706.

The effects of four parameters (speed, fundamental frequency, repetition units, and inharmonicity) on perceived urgency were scaled using an application of Steven's power law. From the exponents obtained, equal units of urgency change were calculated for three parameters. The units were combined in a set of stimuli, and the order of urgency was predicted. The obtained and predicted orders of urgency were highly correlated. The results also showed that even when equalized by psychophysical techniques, some parameters contribute more to perceived urgency than do others. This may be attributable to the different types of parameters scaled or the proportion of the usable range of each parameter that represents a unit change in urgency. The implication of the work for the design and improvement of auditory warnings is discussed.

Hovey, C. K. and Berson, B. L. (1987). Evaluation of ASW tactical symbology. Proceedings of the Human Factors Society 31st Annual Meeting, 1403-1407.

Twelve experienced tactical crew members served as subjects in a study which compared three synonymous sets of tactical contact symbols under conditions of color/monochrome and high/low density. Time and accuracy data of responses to stimulus displays were collected and analyzed, and subjective findings were obtained through post-experiment questionnaires and interviews. The National Tactical Data System symbology produced significantly better results than the two other symbologies, and was favored most by subjects. This presentation describes the study and offers an explanation of the results based on Anne Triesman's findings of preattentive visual processing.

Howard, C., O'Boyle, M. W., Eastman, V., Andre, T., and Motoyama, T. (1991). The relative effectiveness of symbols and words to convey photocopier functions. Applied Ergonomics, 22(4), 218-224.

Two experiments have been conducted to examine preference patterns of inexperienced users of symbols intended to communicate photocopier functions and the relative effectiveness with which keywords, subject-generated symbols and industry symbols are matched to verbal descriptions of various photocopier operations. First, inexperienced photocopier users were asked to draw a set of symbols that they believed best represented 16 designated copier functions, and subsequently a second set of subjects compared these generated symbols with those currently used by the photocopier industry, and gave their preference. Inexperienced users generally preferred the subject-generated symbols, except for those involving the fairly complex 'sidedness' functions. Essentially, keywords and generated symbols did not differ in matching accuracy, while both were more accurately matched than industry symbols. However, the complexity of the photocopier function to be conveyed and the familiarity of the user with a given copier operation did produce several exceptions to this overall pattern. The results suggest that careful ergonomics research can produce symbols that are at least as effective as keywords in conveying photocopier functions. The universal nature of such symbols and their possibilities for cross-cultural use are also noted.

International Organization for Standardization. (1969). Graphical symbols for use on equipment -- Index and synopsis. International Standard ISO 7000.

A graphical symbol is defined as a visually perceptible figure used to transmit information independently of language. It may be produced by drawing, printing, or other means. The principles for the creation of graphical symbols for use on equipment are laid down in ISO 3461-1.

To respond to the increasing international interest in the design and use of graphical symbols, this International Standard presents a certain number of graphical symbols, for use in different technical fields....For each graphical symbol registered in this standard, the original (prepared in accordance with the principles stated in ISO-3461-1) exists at the Secretariat of ISO.

This International Standard provides a synopsis of graphical symbols which are placed on equipment or parts of equipment of any kind in order to instruct the persons handling the equipment as to its use and operation.

Jensen, D. G. (1993). Space Station Freedom aural toxic alarm development. Proceedings of the Human Factors and Ergonomics Society 37th Annual Meeting, 52-54.

To minimize training and enhance crew performance, Space Station Freedom will use the same auditory alarms that are in use for the Orbiter (i.e., fire/smoke, rapid delta pressure, master, and systems management alarms). However, it has been determined that an additional toxic-atmosphere alarm is required for Space Station Freedom. The purpose of this study was to select an auditory toxic-atmosphere alarm for Space Station Freedom. Four final toxic alarm candidates were selected to be tested based upon expected performance indicated by previous research, compatibility with existing Orbiter alarms, and human hearing characteristics. The candidate toxic atmosphere and the Orbiter alarms were tested to determine how well they could be remembered and discriminated. All 100 subjects received the four Orbiter alarms. Each group of 25 of these subjects received one candidate toxic alarm. The results of the comparisons between the candidate toxic alarms indicated that the "chirp" alarm ranked first for both response times and errors. The next evaluation looked at how the Orbiter alarms compared with the "chirp" toxic alarm. The results indicated that the "chirp" toxic alarm had the lowest mean response time and the lowest mean errors when compared to the Orbiter alarms. These results indicate that the "chirp" alarm most consistently ranked highest of the four candidate toxic alarms tested and that this alarm also elicited competitive performance, in terms or response times and correct identifications, when compared with the existing Orbiter alarms.

King, R. A. and Corso, G. M. (1993). Auditory displays: If they are so useful, why are they turned off? Proceedings of the Human Factors and Ergonomics Society 37th Annual Meeting, 549-553.

Pilots often turn off the auditory displays which are provided to improve their performance (Weiner, 1977; Veitengruber, Boucek, & Smith, 1977). The intensity of the auditory display is often cited as a possible cause of this behavior (Cooper, 1977). However, the processing of the additional information is a concurrent task demand which may increase subjective workload (Wickens & Yeh, 1983; McCloy, Derrick, & Wickens, 1983). Pilots may attempt to reduce subjective workload at the expense of performance by turning off the auditory display.

Forty undergraduate males performed a visual search task. Three conditions: auditory display on, auditory display off, and subject's choice were run in combination with nine

levels of visual display load. The auditory display, a 4000 Hz tone with a between-subject intensity of 60 dB(A), 70 dB(A), 80 dB(A), and 90 dB(A), indicated that the target letter was in the lower half of the search area. NASA-TLX (Task Load Index) was used to measure the subjective workload of the subjects after each block of trials (Hart & Staveland, 1988).

A nonmonotonic relationship was found between auditory display intensity and auditory display usage. Evidence was found that the auditory display increased some aspects of subjective workload -- physical demands and frustration. Furthermore, there was a dissociation of performance and subjective workload in the manner predicted by Wickens & Yeh (1983). The implications of these results for display design are discussed.

Kirkpatrick, M., Dutra, L. A., Lyons, R. A., Osga, G. A., and Pucci, J. J. (1992). Tactical symbology standards. Proceedings of the Human Factors Society 36th Annual Meeting, 1087-1091.

CRT displays aboard U. S. Navy ships use a standardized monochrome Naval Tactical Data System (NTDS) symbol set to represent properties of symbols such as platform type (e.g., Aircraft Carrier, Combat Air), environment (e.g., air, surface, subsurface), and identification (e.g., hostile, friendly). A color symbol set has been proposed in NATO Standardization Agreement 4420, *Display Symbology and Colors for NATO Maritime Units* (1990). The U. S. Navy is currently considering ratification of this standardization agreement (STANAG). Empirical comparisons of operator performance using the NTDS symbology versus those using the color-filled NATO STANAG symbology were conducted. Two additional experimental symbologies were also created. The first, called NTDS Equated, is a color version of the NTDS symbol set, and the second experimental symbol set, called NATO Outline, is a color outline version of the color filled NATO STANAG symbol set. Test subjects were asked to find (hook) specific symbols during a tactically relevant scenario.

Time to the first correct hook and percentage of correct hooks were subjected to analyses of variance (ANOVA). Experimental results revealed that the NATO STANAG symbol set outperformed all other symbol sets in terms of symbol recognition time, and outperformed the NTDS Standard symbol set for symbol recognition accuracy as well. The results indicated that tactical information can be transferred more quickly and accurately to watch standers through effective use of symbol coding. Test subjects familiar with the NTDS symbology expressed a preference for the color symbol sets in opinion surveys administered after the experiment.

General conclusions resulting from comparisons across symbol sets were that color fill was more effective than color outline. This paper presents the human performance assessment that was conducted, the results, and the implications of the findings for ratification of NATO STANAG 4420.

Kline, D. W. and Fuchs, P. (1993). The visibility of symbolic highway signs can be increased among drivers of all ages. Human Factors, 35(1), 25-34.

Visibility and comprehension of standard text, standard symbolic, and improved symbolic highway signs were compared among young, middle-aged, and elderly observers. The average distance at which standard symbolic signs could be identified was about two times that of text signs for all three age groups. The visibility distances of the improved symbolic signs, which were designed using an optical blur (i.e., low-pass) approach in order to avoid higher spatial frequencies, exceeded those of both text and standard symbolic signs. Visibility distance was decreased significantly among older drivers on some signs but not others. There were no significant age differences in the comprehension of symbolic signs. Acuity, a good predictor of visibility distance of both text and standard symbolic signs, was only weakly related to the visibility distance of the improved symbolic signs. These findings demonstrate that low-pass symbolic signs have significant advantages in visibility over their text counterparts for all drivers.

Kline, T. J. B., Ghali, L. M., and Kline, D. W. (1990). Visibility distance of highway signs among young, middle-aged, and older observers: Icons are better than text. Human Factors, 32(5), 609-619.

The visibility distances for young, middle-aged, and elderly observers of text and icon versions of four different highway signs were compared under day and dusk lighting conditions. No age differences were observed. Icon signs, however, were visible at much greater distances than were text signs for all three age groups, a difference that was more pronounced under dusk conditions. There were no age differences in the comprehension of icon signs, but there was considerable variability from one icon sign to another in the degree to which they were comprehended. Acuity was found to be a better predictor of the visibility distance of text signs in both day and dusk conditions than it was of icon signs. To the degree that they are comprehended, icon signs appear to offer drivers of all ages almost twice as much time in which to respond to them.

Kolers, P. A. (1969). Some formal characteristics of pictograms. American Scientist, 57, 348-363.

Discusses communication needs of international travelers, absence of an international language, possible use of pictograms for "universal" communication. Includes sections on (1) kinds of writing systems, (2) mechanisms of action, (3) elements and compounds, (4) requirements of writing systems, and (5) need for research.

Lin, R. (1992). An application of the semantic differential to icon design. Proceedings of the Human Factors Society 36th Annual Meeting, 336-340.

Previous studies have indicated that the semantic differential was effective in evaluating comprehension of icons. However, the capability of semantic differential ratings depends on whether the underlying rating factors have been chosen properly. It is necessary to find out what cognitive factors affect the evaluation of an icon. Then, these factors can be used as the basis for semantic differential ratings on the proposed icons during the design stage. Most of the studies are focused on the evaluation after the design is completed. Very few have ever mentioned the approaches of icon evaluation at the design stage to ensure the design quality. Therefore, the purpose of this study is intended to derive and validate the cognitive factors that affect icon designs, and to provide designers with a tool having predictive information for evaluating and modifying the proposed icons at an early design stage.

Lin, R. and Kreifeldt, J. G. (1992). Understanding the image functions for icon design. Proceedings of the Human Factors Society 36th Annual Meeting, 341-345.

Because icons vary from very representational to extremely abstract symbols in a user interface, an important issue faced by designers when designing an icon is how to select an appropriate design style for the image. There are no simple rules that can be followed by designers to determine the design style. The present study is intended to help designers to choose a proper design style for the icon at an early design stage. First, a classification of icons is summarized and the levels of stylization are discussed and demonstrated with examples. Then, thirty icons from several drawing packages and generally used symbols are selected, and a matching test is conducted to obtain the correct matching rates. The results are presented with some explanations for icon recognition and confusion, and finally how to select the right image function for icon design is discussed.

MacGregor, J. N. (1992). A comparison of the effects of icons and descriptors in videotex menu retrieval. International Journal of Man-Machine Studies, 37, 767-777.

The paper describes an experiment to test whether icons improve performance with computer menus because of (a) the additional information they provide, or (b) inherent pictorial properties. The experiment used three different versions of menu pages adapted from a national videotex system. In one version, the menus consisted of labels only, in the second, of the labels plus text descriptors, and in the third, of the labels plus icons. The results indicated that adding icons to videotex menus had the same effect as adding equivalent textual descriptors. Neither reduced response times, while both reduced errors by the same amount (40%). Furthermore, the effect of both icons and descriptors was entirely attributable to a reduction in a specific type of error. In the absence of either icons or descriptors, subjects frequently failed to recognize any of the menu options as relevant, including the correct one, and wrongly selected a "none of the above" option. Adding descriptors and icons appeared to specify the contents of categories sufficiently to reduce this type of error.

However, caution should be exercised in interpreting the results. The study used specific sets of icons and descriptors, and the results may not generalize to other sets of icons or descriptors. Similarly, the study used videotex information-retrieval menus and the results may not generalize to software command menus and other applications.

Marcus, A. (1992). Graphic design for electronic documents and user interfaces. New York, NY: ACM Press.

Chapter 3, **Symbolism**, contains the following sections: (3.1) Clarity and consistency in icon design, (3.2) Icon design tips, (3.3) Icon design in a CAD/CAM graphical user interface: A case study, and (3.4) An annotated bibliography of signs, icons, and symbols.

Osborne, D. W. and Huntley, M. S., Jr. (1992). Design of instrument approach procedure charts: Comprehending missed approach instructions coded in text or icons. Proceedings of the Human Factors Society 36th Annual Meeting, 1455-1459

The objectives of this experiment were to determine whether coding missed approach instructions in text or icons would result in more efficient information transfer, and if the information transfer efficiency for either coding technique was dependent upon the level of information content. Twelve pilots currently licensed for instrument (IFR) flight participated as subjects. Text instructions were either

taken directly or developed from instructions found on National Ocean Service (NOS) instrument approach procedure charts. These instructions possessed one of three levels of information content: low, medium, or high. Across the range of information content levels, iconic missed approach instructions were comprehended more quickly and as accurately as instructions coded in text of the font style and size used by NOS. Regardless of coding technique, report accuracy was significantly worse for instructions with a high information content level. Pilots indicated that in single pilot IFR conditions, they would rather have the iconic than the text version of the missed approach instructions.

Remington, R. and Williams, D. (1986). On the selection and evaluation of visual display symbology: Factors influencing search and identification times. Human Factors, 28(4), 407-420.

Three single-target visual search tasks were used to evaluate a set of cathode-ray tube (CRT) symbols for a helicopter situation display. The search tasks were representative of the information extraction required in practice, and reaction time was used to measure the efficiency with which symbols could be located and identified. Familiar numeric symbols were responded to more quickly than graphic symbols. The addition of modifier symbols, such as a nearby flashing dot or surrounding square, had a greater disruptive effect on the graphic symbols than did the numeric characters. The results suggest that a symbol set is, in some respects, like a list that must be learned. Factors that affect the time to identify items in a memory task, such as familiarity and visual discriminability, also affect the time to identify symbols. This analogy has broad implications for the design of symbol sets. An attempt was made to model information access with this class of display.

Schmidt, J. K. and Kern, J. E. (1991). A literature review on tactical map symbology. Air Vehicle and Crew Systems Technology Department, Naval Air Development Center, Warminster, PA. (Draft)

Advances in computer and information display technologies have made it possible to portray information of military tactical significance in real time on digitized terrain databases. This capability has the potential to increase the tactical situation awareness of military personnel. Many applications for this type of capability exist within the context of a military mission. Where land-based, sea-based, and airborne assets are being deployed jointly to deter an enemy threat, an awareness of the tactical

situation is critical. The military requirement is for limited tactical resources to be effectively employed in a timely fashion. Consequently, it is vital that intelligence information, encoded symbolically, on a Tactical Situation Display (TSD) to be designed to ensure its accurate, quick, and reliable extraction. To achieve an understanding of symbol design parameters, the operational, technological, and ergonomic literature was reviewed. Recommendations are then made for encoding tactical information in symbols.

Sorkin, R. D. (1987). Design of auditory and tactile displays. In Salvendy, G. (Ed.), Handbook of human factors, 549-576. New York: Wiley.

Discusses when to use a nonvisual means of presenting information. Describes the auditory and tactile senses and provides recommendations for desirable characteristics of auditory and tactile signals.

Stammers, R. B. and Hoffman, J. (1991). Transfer between icon sets and ratings of icon concreteness and appropriateness. Proceedings of the Human Factors Society 35th Annual Meeting, 354-358.

Two studies are reported on the usability of icons. Firstly, an experiment was carried out to investigate the transfer of performance between sets of icons for different computer functions. The aim was to determine whether users would benefit from common underlying elements within concrete icon sets, compared with unrelated abstract icons. A cross-over study was employed whereby subjects performed an icon identification task with one set of icons (a concrete or abstract learning set). It was hypothesized that transferring from concrete to concrete icons would yield the best results due to the use of repeated elements in the concrete sets. Identification time data did not produce significant results to support the hypothesis. Significant results were yielded from the error data, this demonstrated the superiority of concrete icons, but no advantage came from previous experience with a concrete set A second study examined the rating of the icons used in the first study in terms of concreteness and appropriateness. Appropriateness was found to be a reasonable predictor of icon identification time.

Strijland, P. (1992). Icon design. Tutorial Notes, CHI '92, ACM Conference on Human Factors in Computing Systems, Monterey, CA.

This tutorial is designed for interface designers and programmers that are or will be involved in the design of icons for graphical user interfaces. Many aspects of icon

development are discussed, including concept generation, graphic design, interaction between a graphic designer and an implementing programmer, user testing, and compliance with standards. This class starts with a general discussion of the principles of pictorial versus verbal information. These principles will be applied to the computer interface. Several exercises will be done to come up with concepts for icons, and the good and bad aspects of the generated ideas will be discussed. Several case studies from Macintosh system 7 icon development are used as illustration. Different methods for testing icons, and an overview of different systems' icon solutions are presented. Finally, a brief update of the current status of an ISO standard for icons that is in development will be given, as well as an overview of its implications for icon designers.

Van Orden, K. F. and DiVita, J. (1993). Highlighting with flicker. Proceedings of the Human Factors and Ergonomics Society 37th Annual Meeting, 1300-1304.

Previous research has demonstrated that search times are reduced when flicker is used to highlight color coded symbols, but that flicker is not distracting when subjects must search for nonhighlighted symbols. This prompted an examination of flicker and other stimulus dimensions in a conjunctive search paradigm. In all experiments, at least 15 subjects completed a minimum of 330 trials in which they indicated the presence or absence of target stimuli on a CRT display that contained either 8, 16, or 32 items. In Experiment 1, subjects searched for blue-steady or red-flickering (5.6 Hz) circular targets among blue-flickering and red-steady distractors. Blue-steady targets produced a more efficient search rate (11.6 msec/item) than red-flickering targets (19.3 msec/item). In Experiment 2, a conjunction of flicker and size (large and small filled circles) yielded the opposite results; the search performance for large-flickering targets was unequivocally parallel. In Experiment 3, conjunctions of form and flicker yielded highly serial search performance. The findings are consistent with the response properties of parvo and magnocellular channels of the early visual system, and suggest that search is most efficient when one of these channels can be filtered completely.

Van Orden, K. F., DiVita, J., and Shim, M. J. (1993). Redundant use of luminance and flashing with shape and color as highlighting codes in symbolic displays. Human Factors, 35(2), 195-204.

Three visual search experiments evaluated the benefits and distracting effects of using luminance and flashing to highlight subclasses of symbols coded by shape and color.

Each of three general shape/color classes (circular/blue, diamond/red, square/yellow) was divided into three subclasses by presenting the upper half, lower half, or entire symbol. Increasing the luminance of a subclass by a factor of two did not result in a significant improvement in search performance. Flashing a subclass at a rate of 3 Hz resulted in a significantly shorter mean search time (48% improvement). Increasing the luminance of one subclass (by a factor of five) while simultaneously flashing another significantly improved search times by 31% and 43%, respectively, compared with nonhighlighted search conditions. In each experiment, the search times for nonhighlighted target subclasses were not affected by the presence of brighter and flashing targets. The failure of the initial experiment to find a significant performance improvement caused by increasing symbol luminance suggested that a larger luminance increase was necessary for this code to be effective. The overall results suggest that using luminance and flashing to highlight subclasses of color- and shape-coded symbols can reduce search times for these subclasses without producing a distraction effect by way of a concomitant increase in the search times for unhighlighted symbols.

Vukelich, M. and Whitaker, L. A. (1993). The effects of context on the comprehension of graphic symbols. Proceedings of the Human Factors and Ergonomics Society 37th Annual Meeting, 511-515.

When graphic symbols are used to convey warning information, these symbols must be evaluated for effectiveness prior to their use. In general, the ability of these symbols to convey their intended meaning has been determined in tests which provide no contextual information surrounding the symbols. In the present study, 75 university students were tested to determine their comprehension of twenty different symbols using various context conditions. Verbal context was provided in two forms: full context and partial context. Full context consisted of a two-sentence description of the setting in which the symbol would be presented. Partial context consisted of a more general, two-word description of the use context. The control condition presented the symbols without contextual information. Comprehension was higher when full context was provided with the symbols than when the symbols were presented in isolation. For some symbols, the full context condition resulted in higher comprehension than the partial context condition and the partial context condition resulted in higher comprehension than the no context condition. Comprehension accuracy was also affected by the subject's familiarity with the symbols. Comprehension was higher for symbols rated high in familiarity than for symbols rated

lower in familiarity. On the basis of these findings, a recommendation was made that evaluations should provide some form of contextual information along with the symbols to allow a more realistic test of symbol comprehension.

Waterworth, J. A., Chignell, M. H., and Zhai, S. M. (1993). From icons to interface models: Designing hypermedia from the bottom up. International Journal of Man-Machine Studies, 39, 453-472.

We describe a method to derive design models for hypermedia interfaces from the bottom up. firstly, we compile a list of hypermedia interface features which we classify according to the category of functions they fulfill. We then describe an experiment in which candidate designs for low-level interface features were designed and tested for recognizability. In the experiment, icons for each of 61 hypermedia concepts were generated and then judged. Finally, we outline and illustrate a model induction phase in which low-level features are combined into an overall interface model, via "micro-models" that take account of the types of icons that worked best for each class of interface feature. We suggest that, at least for hypermedia systems, a bottom-up approach to interface design based on the functions of low-level features is preferable to the dominant, top-down approach based around one or more metaphors.

•U.S. GOVERNMENT PRINTING OFFICE: 1994-604-081-00208