AD

Technical Memorandum 9-91

THE EFFECTS OF DISPLAY FAILURES, POLARITY, AND CLUTTER ON VISUAL SEARCH FOR SYMBOLS ON CARTOGRAPHIC IMAGES

Craig J. Dye
Harry L. Snyder

April 1991
AMCMS Code 612716.H700011

Approved for public release;
distribution is unlimited.

U.S. ARMY HUMAN ENGINEERING LABORATORY

Aberdeen Proving Ground, Maryland

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1a. REPORT SECURITY CLASSIFICATION Unclassified	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution unlimited.
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S) Technical Memorandum 9-91	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION Human Engineering Laboratory 6b. OFFICE SYMBOL (If applicable) SLCHE	7a. NAME OF MONITORING ORGANIZATION
6c. ADDRESS (City, State, and ZIP Code) Aberdeen Proving Ground, MD 21005-5001	7b. ADDRESS (City, State, and ZIP Code)
8a. NAME OF FUNDING/SPONSORING ORGANIZATION 8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
8c. ADDRESS (City, State, and ZIP Code)	10. SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO. 6.27.16 PROJECT NO. 1L162716AH70 TASK NO. WORK UNIT ACCESSION NO.

11. TITLE (Include Security Classification)

The Effects of Display Failures, Polarity, and Clutter on Visual Search for Symbols on Cartographic Images

12. PERSONAL AUTHOR(S)

Dye, Craig J.; Snyder, Harry L.

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Final	FROM ________ TO ________	1991, April	48

16. SUPPLEMENTARY NOTATION

17. COSATI CODES

FIELD	GROUP	SUB-GROUP
05	02	
05	08	

18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)

cartographic images, clutter, display failures, human factors, matrix-addressable displays, polarity, visual search

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

Little research has been conducted about human performance in symbol search of cartographic images on matrix-addressable displays. Alphanumeric studies have researched failures and polarity on such displays but not using cartographic images. Two studies were conducted which required subjects to search for symbols on cartographic images under various conditions of failure, polarity, and background clutter. Response time and accuracy were measured. Generally, high clutter impeded performance more than low clutter, and negative contrast produced slightly better performance than positive contrast. Horizontal line failures affected search ability more than vertical line or cell failures did, particularly when the failures were in the "on" mode rather than the "off" mode. Further, it was discovered that while previous alphanumeric studies found effects at 3% failures or greater, effects were observed for the present studies at 2% or greater, apparently because of the presence of map information on the display. Recommendations for cartographic display design and future research are discussed.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT ☐ UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION Unclassified
22. NAME OF RESPONSIBLE INDIVIDUAL Technical Reports Office	22b. TELEPHONE (Include Area Code) (301) 278-4478 22c. OFFICE SYMBOL SLCHE-SS-TSB

DD Form 1473, JUN 86 — Previous editions are obsolete.

SECURITY CLASSIFICATION OF THIS PAGE

UNCLASSIFIED

AMCMS Code 612716.H700011 Technical Memorandum 9-91

THE EFFECTS OF DISPLAY FAILURES, POLARITY, AND CLUTTER ON VISUAL SEARCH FOR SYMBOLS ON CARTOGRAPHIC IMAGES

Craig J. Dye
Harry L. Snyder

April 1991

APPROVED: JOHN D. WEISZ
Director
Human Engineering Laboratory

U.S. ARMY HUMAN ENGINEERING LABORATORY

Aberdeen Proving Ground, Maryland

ACKNOWLEDGMENTS

The authors thank Mr. Clarence Fry, Dr. Richard A. Monty, and the U.S. Army Human Engineering Laboratory, Aberdeen Proving Ground, Maryland, for their support of this research.

CONTENTS

TABLES

THE EFFECTS OF DISPLAY FAILURES, POLARITY, AND CLUTTER ON VISUAL SEARCH FOR SYMBOLS ON CARTOGRAPHIC IMAGES

INTRODUCTION

Cartographic and symbolic displays refer to non-alphanumeric information displays in the form of maps, graphs, or other pictorial material. To date, almost all cartographic research has been conducted using paper maps. With the increase in viability of matrix-addressable and cathode-ray tube (CRT)-based navigation aids in vehicles, as well as real time monitoring of military situational attributes on CRT and matrix-addressable displays, it is most important to discover the display parameters that affect the operator's performance while using such systems.

The human performance research to date concerning cartographic information has centered around symbol resolution and symbol size (Erickson, 1978; Florence & Gieselman, 1986; Gieselman, Landee, & Christen, 1982). Although these considerations are of some importance to cartographic image quality, variables at least as meaningful as symbol characteristics have gone unresearched.

With matrix-addressable displays, it is possible that cells or even whole lines on the display may fail, in which a pixel(s) may remain "on" or "off" regardless of the intended state. To adequately define a failure situation on a matrix-addressable display, one must consider the type of failure (individual cell or line failure), the mode of the failure (on or off), and the amount of failure present (typically quantified as the percent of pixels failed).

Three failures arise most often with matrix-addressable displays. A cell failure exists when individual elements or pixels turn on or off depending on the mode of the failure. This failure is often described as producing a "salt and pepper" effect. A cell failure may result from a variety of causes depending on the type of flat panel technology being considered. Regardless of the technology, all such displays are susceptible to cell failures. The other two types of display failures are horizontal and vertical line failures. In most instances, a line failure of either type is the result of a faulty electrode or driver. In this case, an entire line on the display fails either on or off.

For any failure type, the pixels comprising the failure may fail in on or off modes. Consider that on any given display, the luminance of the information will be higher or lower than the luminance of the background, depending on the display's polarity. The on mode refers to the case when a failure more nearly resembles the luminance of the information on the display, whereas the off mode refers to the case when a failure more nearly resembles the luminance of the background on the display. Therefore, any failure will either turn off the pixels comprising the information (matching the background) or turn on the pixels comprising the background (matching the information). Obviously, when a failure matches the luminance of the information (on mode), it can be readily seen where it intrudes upon the background of the display, but not where a character or symbol is commanded on the display. Likewise, an off failed line or cell is not seen in the background portion of a displayed image.

The amount of failure is typically defined by the percentage of pixels failed or failure percent. This is easily calculated by dividing the number of pixels failed by the total number of addressable pixels on the display and multiplying by 100.

Display Failure Research

Research has examined the effects of failure mode on individual symbols as well as on entire display images. Riley and Barbato (1978) examined the relationship between five fonts and discrete element degradations (a cell failure). A set of 5 x 7 dot-matrix characters was drawn in the center of a 7 x 9 matrix. Dots in the 7 x 9 matrix were then turned on or off, and individuals were asked to identify the character. Riley and Barbato found no difference between the failure modes.

A similar study was conducted by Pastor and Uphaus (1982), who examined the confusability of 7 x 9 American standard code for information interchange (ASCII) numerals under varying percentages of dot loss (or cell failure). They found that a linear relationship existed between specific dot losses and reading errors. Note that this study only examined the effects of off failures.

These studies, while informative, investigated only cell failures and failure mode for individual characters. It is also important to examine the effects of all types of failures (both cell and line), as well as failure mode and percent failure, on an image displaying a more complex field of characters that more closely resembles many operational applications.

Laycock (1985) suggested that line failures were disruptive when they aligned with major components of the characters comprising the text and that off cell failures were less disruptive than on cell failures. Laycock also believed that less than 0.01% of on cell failures and as much as 1.0% of off cell failures are tolerable.

When considering this last point (that of percent failure), it is important to note that Laycock defined percent failure as the percentage of pixels failed that comprise the text, rather than the percentage of pixels failed on the entire display. While the locations of failures on any matrix-addressable display will be random, these locations are not likely to vary from image to image as will the text that is presented on the display. Therefore, it seems more meaningful to quantify percent failure as the percentage of pixels failed on the entire display. Only three studies were found that quantify failure percent in this manner.

Abramson and Snyder (1984) examined failure type, failure mode, and percent failure for a speed of reading task. Cell failures generally resulted in slower reading speeds and a higher number of errors than did horizontal or vertical line failures. This was particularly evident when the failures were on. When the failures were off, line failures resulted in poorer performance than did cell failures. Overall, off failures resulted in better performance than on failures did. It was further shown that as failures increased beyond 2%, reading speed decreased and errors increased. At or below 2%, the failures had little effect on performance. These three failure variables showed markedly similar results for random search tasks (Decker, Dye, Lloyd, & Snyder, 1991; Lloyd, Decker, Kurokawa, & Snyder, 1988).

All three studies involved images of text passages or random patterns of alphanumerics and/or symbols. To date, there are no data examining the effects of these failure variables on the search for symbols on more complex, cartographic images. The present study includes the failure variables and levels examined in the above research, except for percent failure. While levels of 1% to 12% were employed in the previous alphanumeric studies, it was believed that because of the addition of map information to the images, effects should become evident at or below 3% failure.

Polarity Research

Polarity refers to whether the information is of greater or lesser luminance than the background (low luminance information on a high luminance background [negative contrast] or high luminance information on a low luminance background [positive contrast]).

In a review of the literature, Rupp (1981) indicated that the preference in Europe for negative contrast is based almost entirely on an assumption that the pupillary response that results from the switch in eye fixation from text on paper (negative contrast) to text on a positive contrast display induces visual fatigue and irritation. Alluding to the results of an informal experiment he conducted, Rupp contends that there is no difference in magnitude of pupillary response (attributable solely to the positive contrast display) above that which occurs naturally during normal steady fixation or pupillary hippus.

An experiment by Bauer and Cavonious (1980) examined the effects of polarity on people's ability to identify four-letter nonsense words. Error rates indicated that negative contrast produced better identification than positive contrast did.

Semple, Heapy, Conway, and Burnett (1971) reviewed the literature and concluded that display polarity has no interpretable impact on symbol recognition and that claimed differences are primarily attributable to inaccurate reporting of experimental methodology.

The older literature suggests that there is no consistent difference in performance because of display polarity. The 1988 American National Standards Institute (ANSI) standard for video display terminal work stations (Human Factors Society, 1988) indicates that either polarity is acceptable, provided it meets the requirements for resolution, luminance, and contrast. This qualification proves to be quite important.

The previous articles offer little mention of display parameters such as luminance levels, modulation, or stroke widths, even though these parameters all have well-documented effects upon legibility of characters (Snyder, 1980). To accurately compare positive to negative contrast, modulation and stroke width must be kept constant for both polarity conditions. Regarding display luminance, there has been some disagreement about whether the eye adapts to the peak luminance of an image or to the space average luminance of an image. Recent evidence (Knox & Beaton, 1985) supports the concept that the eye adapts to the average luminance, not to the maximum luminance on the display. This distinction is pertinent since it is nearly impossible to establish equal stroke width and modulation as well as equal peak and space average luminances for images of both polarities on the same display. Snyder (1988) stated "... it appears...that the adaptive state of the visual system is driven by the amount of light entering either the fovea (for a positive [contrast] display) or the parafoveal region (driven by the background in a negative [contrast] display)." It is apparent, then, that if adequate modulation, stroke width, and minimum luminance are maintained, there should be no adverse problems in visual adaptation for images of either positive or negative contrast.

These criteria were met, and polarity was examined in three alphanumeric studies (Decker, Kelly, Kurokawa & Snyder, 1991; Lloyd, Decker, & Snyder, 1991; Lloyd et al., 1988). Negative contrast produced significantly shorter response times and proved more accurate than positive contrast for random search tasks. Thus, the most recent and well-controlled data indicate an advantage to negative contrast.

Polarity has not been investigated in cartographic research. While it has been the natural practice of map makers to employ only negative contrast, there has been no empirical research to date to validate this stance.

Background Clutter Research

For the purposes of this research, background clutter may be defined as image information that should not be attended to or is irrelevant at a given time. Any information about a cartographic image that is not the target or does not directly contribute to the acquisition of the target can be considered background clutter, since it should not be attended to. This clutter would include non-target symbols as well as map information that is unnecessary for the acquisition of the target.

Of those studies (Florence & Gieselman, 1986; Silbernagel, 1982; Williges & North, 1973) that investigated the effect of background clutter (sometimes termed "density") on visual search for targets on CRT displays, all found that an increase in the number of non-target symbols increased response time and decreased accuracy, as would be expected. Only one of these studies (Williges & North, 1973) involved the search for targets on cartographic images, but the topographic maps were presented by filming paper maps and displaying them to the subject via closed circuit television, that is, the maps were not computer generated.

All three of these studies defined background clutter as the number of non-target symbols present on the image. This definition of background clutter is not adequate for cartographic images on matrix-addressable displays because, as indicated in the operational definition of background clutter, it does not take into consideration the distracting effect of the nonessential map information in the image. A better, more quantitative definition of background clutter would be the percentage of pixels on a display that comprise the nonessential information, both map information and non-target symbols. Quantifying the relationship between user performance and the amount of background clutter will be increasingly important as technology provides avenues for such displays to develop in complexity and sophistication.

Research Rationale and Objectives

Although matrix-addressable displays of cartographic images are becoming increasingly important and available, they seem to have been largely ignored by researchers of human performance.

The effects of failures have been examined for random search of alphanumerics and symbols but not for symbol search on cartographic images.

As a display intended for cartographic information will undoubtedly display alphanumerical information as well, the effects of failures on cartographic images must be ascertained to effectively define the acceptable limits of these failures for displays intended for both cartographic and alphanumeric presentation. One purpose of this research effort is to define such limits.

As mentioned earlier, it has been the natural practice of map makers to employ only negative contrast. Thus, the small amount of cartographic research regarding human performance has limited itself to negative contrast presuming a population stereotype as justification. This research will test the validity of this presumption. Further, since recent alphanumeric studies (Decker, Kelly, Kurokawa, & Snyder, 1991; Lloyd et al., 1988; Lloyd et al., 1991) have shown a slight advantage for performance with negative contrast, the present study tries to determine whether this is the case for cartographic images.

With the increasing sophistication and affordability of matrix-addressable displays and cartographic software, the amount of background clutter will eventually become limited only by the courtesy of the software programmer. While discovering the acceptable limits of background clutter on a cartographic image is beyond the scope of this work, the present study provides initial information regarding this effect and its limits.

It is not known whether results may generalize from alphanumeric search tasks to cartographic search tasks and vice versa, although this information is needed to design an optimum cartographic and alphanumeric display. This research produces information for comparing these two types of images.

Accordingly, two experiments were conducted that were identical in general procedure but differed in the task required of the subjects. One study involved a random search task, and the other involved an information-extraction task.

The information-extraction task required subjects to find a target indicated before each trial solely by contextual information presented on the CRT. For instance, the user of such a display might be asked to find the enemy tank that is closest to the water. This task requires the user to be able to make a quick distinction between friendly versus enemy symbols as well as between the different types of fixed information. This task is similar to those that occur in tactical use of such a display.

Both studies included the variables of failure type, failure mode, percent failure, polarity, and background clutter.

METHOD

Subjects

The subjects for both experiments were college students at Virginia Polytechnic Institute and State University who were paid for their participation in the study. For each experiment, there were 12 subjects (six males, six females) who ranged in age from 18 to 26. The subjects were tested for natural or corrected 20/22 near and far point visual acuity as well as lateral and vertical phorias using a Bausch and Lomb Orthorater. Subjects were also tested for normal near and far contrast sensitivity using a Vistech chart system.

Apparatus

The stimuli were presented on a Video Monitors Incorporated (VMI) high resolution monochromatic CRT with a 48-centimeter (cm) diagonal screen. The area of the screen used for these studies was 27.94 cm^2 (1024 x 1024 pixels) because of bandwidth constraints of the graphics controller.

An 8-bit plane PEPE graphics controller by Vectrix Corporation was installed on an IBM personal computer (PC-AT). The PC controlled the generation and presentation of the stimuli, as well as data collection. A three-button Mouse Systems mouse was used for the subjects' responses. Responses were timed using the built-in clock of the PC, which has a resolution of ±55 milliseconds (ms).

Subjects were seated in a dentist's hydraulic chair adjustable in height and distance from the CRT. Subjects were positioned so that their eyes were 50.8 cm from the CRT. Their heads were made stationary through the use of the headrest on the hydraulic chair, and the line of sight to the center of the CRT was 15° below horizontal. The CRT was tilted back 15° so that the line of sight was normal to the display.

Photometric Measurements

Before these experiments, the display luminance and modulation were set using a photometric system which consists of a Model GS-2110 scanning telemicroscope by Gamma Scientific, with a 10- x 3000-micron slit aperture and a 1X objective lens, a photomultiplier tube (Gamma Scientific, Model D-46), and an intelligent radiometer (Gamma Scientific, Model GS-4100). The photometric system is controlled by an IBM PC-XT.

Calibration

The display luminance was set using the display brightness control so that the luminance level of an all-on field (255 bits) was 39.5 candelas per square meter (cd/m^2). This display brightness setting was kept constant, and screen luminance was varied by changing bit levels.

The background luminance was set as closely as possible to 35 cd/m^2 by making vertical scans across several columns of pixels. A zero bit line was then displayed against this background and the line was scanned. The bit level for the line was adjusted and scanned repeatedly until a modulation of 0.65 was reached. Because of monitor constraints, this modulation was as high as was possible to achieve and still maintain comparable maximum luminances and stroke widths for both polarity conditions. This procedure set the bit levels for the negative contrast condition.

To set the bit levels for the positive contrast condition, an all-on (255-bit) line was displayed on an all-off background, and the bit level of the background was increased until values for modulation, maximum luminance, and stroke width were achieved which were comparable to those achieved for the negative contrast condition. Thus, the maximum luminance and stroke width for both polarity conditions were set as close as possible while maintaining nearly equal modulation levels.

The bit levels for the luminances of the background and symbols were programmed into the experimental software.

Independent Variables

The simulated failures consisted of lines of pixels or individual pixels that either matched the luminance of the information on the display or matched the luminance of the background on the display.

Failure type includes no failure, a cell failure (in which individual pixels are affected), or a line failure (in which entire lines of pixels across the display are affected, either horizontally or vertically). Thus, there were four levels of failure type: none, horizontal, vertical, and cell. The locations of the cell and line failures were randomly selected for each trial by the experimental software.

Failure mode, as previously described, has two levels: on or off. On failures matched the luminance of the information on the image, and off failures matched the luminance of the background of the image.

Failure percent varied from 1% of the display's pixels failed to 3% of the pixels failed at 1% increments.

Background clutter was set at two levels, high and low. It was quantified much the same as percent failure was quantified, in terms of the percentage of pixels associated with information displayed on the CRT or the percentage of pixels <u>not</u> matching the background luminance of the display. The high background clutter condition consisted of approximately 3.8% of the displayed pixels on and the low condition consisted of approximately 2.9% of the displayed pixels on. While these percentages do not appear to be very high, they entail more than 30,000 and 40,000 pixels on for low and high clutter levels, respectively. The clutter consisted of all map information, including fixed items as well as symbols; it did not include on pixel failures. Clutter was adjusted by increasing the number of non-target symbols on the display. The non-target symbols consisted of the other symbols in the symbol set as well as the capital letters A through H in a Huddleston 11 x 15 font. Of the non-target symbols, there was approximately a 2:1 ratio of symbols to letters on the images.

The last independent variable is polarity, with two levels: negative (dark symbols on a light background) and positive (light symbols on a dark background).

Maps

The six maps used for this study were adapted from Army maps. Because of the lack of color coding (as the display used in this research was monochromatic), and because too much detail might have unnecessarily complicated the interpretation of the effects of the failure variables, the adaptations did not include all the information on the source maps. Also, the information was enlarged so that the different information on the maps could be readily identified. The types of information (excluding symbols) that were included are roads, rivers, trees, railroads, and buildings.

Symbols

The symbol set was selected from a set of 26 Army symbols used in the previous alphanumeric research in the Virginia Polytechnic Institute (VPI) laboratory. Stimuli were drawn within an 11 x 15 dot matrix subtending 19 x

26 minutes of visual angle. The stimuli are standard Army symbols redrawn as dot-matrix symbols. The symbols were presented at random locations on the CRT without overlapping other symbols. Overlapping may occur, however, between fixed information on the cartographic image and the symbols, as this phenomenon will undoubtedly occur in an actual system.

The eight symbols for these experiments (see Figure 1) were selected by averaging the response times and response accuracy for each symbol across two studies (Decker, Kelly, Kurokawa, & Snyder, 1991; Decker, Lloyd, Kurokawa, & Snyder, 1991) and selecting a representative sample from the ensuing distributions. The selection process was not entirely random because it contained four pairs of symbols, both symbols in each pair identical except for a single horizontal line that runs through or above one of the symbols in each pair to aid the subject in distinguishing between friendly and enemy symbols in the information extraction task.

The number of symbols (eight) was chosen to facilitate long-term retention of the symbol set before the information-extraction task and yet require the subject to distinguish among symbols. Long-term retention of the symbols was deemed to be a requirement of an operator in an application of such a display system.

Dependent Variables

The dependent measures used for both studies are response time and response accuracy. Response times were recorded by capturing the elapsed time between the depressions of the two buttons on the mouse that initiated and completed each trial. The response time and response accuracy for each trial were written, along with information detailing the trial, to a data file on the IBM PC-AT.

The response times were averaged across the 12 repetitions for each of the 1,152 cells (96 conditions x 12 subjects), and these means were used in the ensuing analyses.

Accuracy was defined as the percentage of correct responses (0% to 100%), averaged across the 12 repetitions for each of the 1,152 cells.

Design and Stimuli

The experimental design is identical for both experiments. The design was a 2 x 2 x 4 x 3 x 2 within-subjects full factorial combining polarity by background clutter level by failure type by percent failure by failure mode, respectively. There were 12 repetitions of each of the 96 cells for a total of 1,152 trials per subject. These were spread across 4 days with 288 trials per day. The days were blocked by failure type.

It is reasonable to assume that the type of failure on any given matrix-addressable display in a work place setting will remain constant throughout a certain portion of the operator's tasks. Presenting one failure type for each of the 4 days reproduces this situation in the experimental setting. This blocking provides greater validity as well as data about the ability of the operator to adapt to such a situation. The order of presentation of failure type was balanced through use of a Latin square and randomly assigned to each subject to minimize any effects from order of presentation.

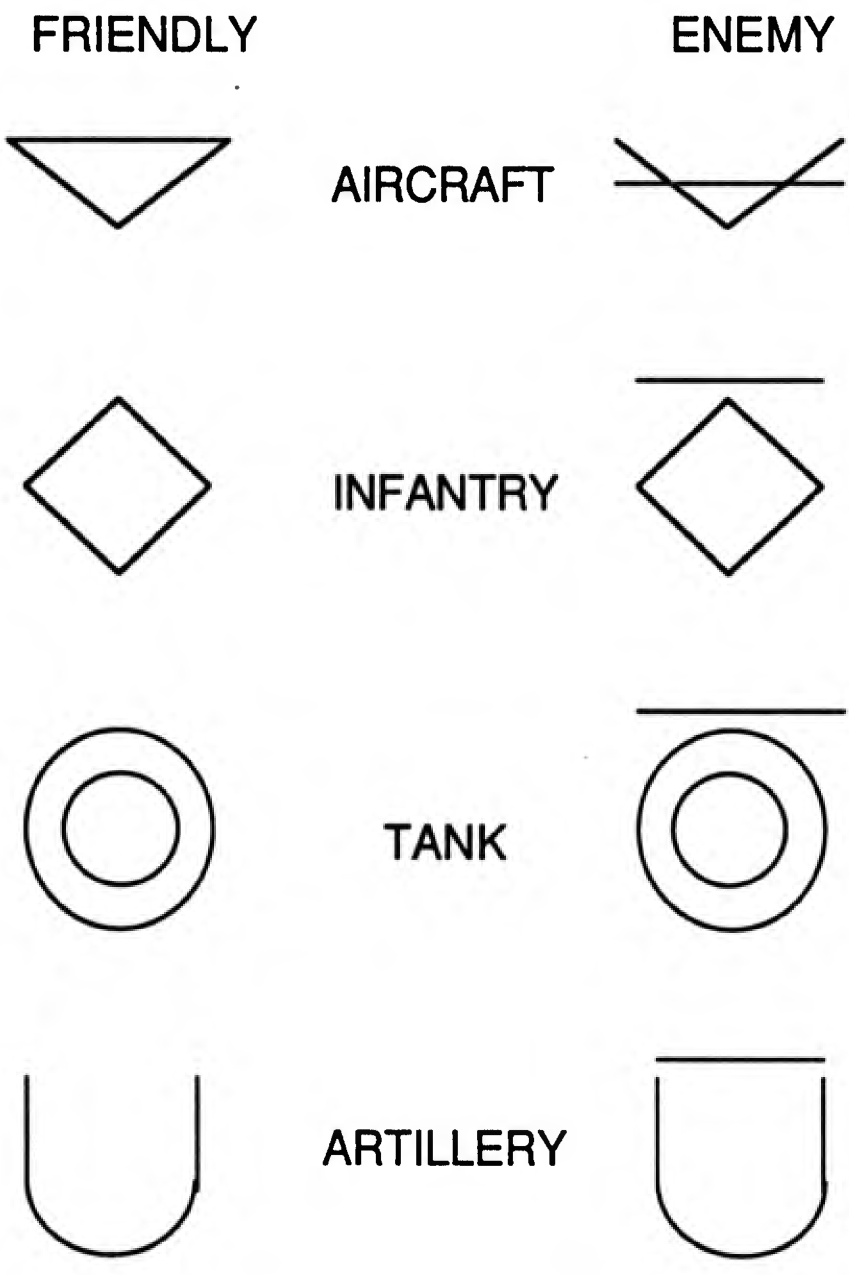

Figure 1. Symbols used for both tasks.

Albert (1975) found that performance differs for contextual versus noncontextual word tasks. If this effect generalizes to symbols, it is potentially confounding to allow a subject to place in context, or attach meaning to the symbols used for the random search task because the subject would no longer be searching for a cognitive template but for a meaningful and contextual representation. To avoid any potential confounding of this nature, a different group of subjects was used for each of the two tasks.

The six maps were used an equal number of times but not across all conditions. That is, they were not treated as a factor in the present studies.

Procedure

At the beginning of each experimental session, the CRT was warmed up for at least 30 minutes to achieve luminance stability (based on prior research). The CRT was adjusted to a luminance of 39.5 cd/m^2 with an all-on field (255 bits). Ambient illumination was set to provide a luminance of 15 cd/m^2 on the wall directly behind the CRT. The ambient did not illuminate the subject's display.

The two tasks performed were a random search task and an information-extraction task. The random search task consisted of the subject finding a target symbol on the cartographic image and identifying its location. The target was presented to the subject graphically on the CRT before each trial.

The subject was seated in the dentist's chair and asked to read the instructions for the particular study in which s/he participated. Following this, the subject was positioned the appropriate distance from the center of the CRT. At this point, the procedure for the two studies diverged.

Random Search Study

For the random search study, the subject was given 30 practice trials that were identical to the actual trials in terms of actions performed. The subject first saw a message screen with the following message appearing at the center of the CRT: "The next symbol is" followed by the target symbol. The symbol appearing with the message was identical in every respect to the actual target symbol (except for pixels of the target symbol that were affected by the failures of the ensuing trial) which appeared at a random position on the subsequent cartographic image of each trial. For each experimental trial, the target symbol was randomly selected and located on the map, while the other symbols and alphanumerics were also randomly placed to achieve the specified background clutter level, as previously noted.

Depressing the right button on the mouse (a) erased the message, (b) started the timer resident in the PC, and (c) displayed that trial's cartographic image with symbols and failures. Once the subject believed s/he had found the target symbol, the subject depressed the left button on the mouse, which stopped the timer. Immediately, the map and failures disappeared, and all symbols on the display were replaced with numerical identification tags. The subject then reported aloud the number of the identification tag that replaced the perceived target symbol. This number was recorded by the investigator.

The next trial's message screen then appeared and the subject proceeded as before.

Following the completion of the final experimental session on the fourth day, the subject was given a debriefing sheet describing in detail the purpose and goals of the study. All questions regarding the study were answered at that time.

Information-Extraction Study

The procedure for the information-extraction study began the same as for the random search study, with the subject seated in the hydraulic chair and positioned with respect to distance from the center of the CRT. The subject was then given a sheet of paper containing the eight symbols and their assigned meanings. The subject was given as much time as needed to study the reference sheet and learn the symbols and their respective meanings. After indicating that s/he was ready to begin, the subject was tested with another sheet of paper containing only the meanings and was told to draw the appropriate symbol. If the subject did not get all eight correct, s/he was given the reference sheet again and, when ready, was given another test sheet. This was repeated as many times as necessary for the subject to get all eight correct. No subject needed more than two tests to get all eight correct.

After the subject passed the learning test, s/he began 30 practice trials. As with the random search procedure, the practice trials were identical to the experimental trials. A message appeared on the screen in the form "Find the <symbol name> closest to the <fixed information>" in which <symbol name> was the learned meaning of the particular symbol to search for, and <fixed information> was the item on the cartographic image to search in relation to, such as woods, water, and so forth.

When ready to proceed, the subject depressed the right button on the mouse key pad, which started the trial, whereby the subject proceeded as in the random search study.

After completing the 30 practice trials, each subject began the experimental trials. There were three groups of trials per session with 16 sections per group for a total of 48 sections per experimental session. Each section consisted of six trials at the same levels of polarity, background clutter, percent failure, and failure mode. Thus, each of the 24 possible combinations of the above variables was repeated 12 times per experimental session: six times in each of two sections. A 5-minute rest break was allowed after each group of 96 trials.

Following the completion of the final experimental session on the fourth day, the subject was given a debriefing sheet describing in detail the purpose and goals of the study. All questions regarding the study were answered at that time.

RESULTS

Response time and response accuracy (percentage of correct responses) were analyzed for both experiments. A five-way analysis of variance (ANOVA) was performed on the entire design (polarity by background clutter level by failure mode by failure type by failure percent) for each dependent variable in each experiment. In this and subsequent analyses, significant ($p < .05$) within-subject sources of variance were checked against violation of the sphericity assumption using minimum (worst case) degrees of freedom (Winer, 1971). When the minimum degrees of freedom calculation resulted in a nonsignificant result, Greenhouse and Geisser (1959) ε calculations were performed and the degrees of freedom were adjusted accordingly. Such calculations are noted on the tables.

Simple effect F tests and Newman-Keuls tests were performed for all significant interactions that were deemed meaningful.

Random Search Task

Response Time

The ANOVA summary for these data is given in Table 1. As indicated, two main effects and one interaction are statistically significant ($p < 0.05$).

Negative contrast produces significantly faster (6.13 seconds) responses than positive contrast (7.29 seconds). Response times are significantly shorter for low background clutter (5.56 seconds) than for high clutter (7.89 seconds).

The percent failure effect ($p = 0.06$) with the conservative Greenhouse and Geisser (1959) correction for violations of sphericity failed to reach traditional ($p < 0.05$) statistical significance. However, because this variable has consistently been demonstrated to have a significant effect on performance, the means for the different levels were investigated. As shown in Table 2 and Figure 2, failures at 1% led to faster responses (6.38 seconds) than at either 2% (6.95 seconds) or 3% (6.84 seconds).

Of greater interest than the main effects, however, are the interactions between the variables. Failure mode significantly interacts with failure type. The simple effect F tests indicate that the effect of failure mode is significant for cell failures only (see Table 3 and Figure 3), with off failures producing significantly faster responses (5.99 seconds) than on failures (7.44 seconds).

Response Accuracy

The five-way ANOVA for accuracy in the random search task found all main effects but percent failure to be statistically significant (see Table 4).

As with response time, polarity and background clutter level are both significant, with performance being slightly more accurate for negative contrast (98.1%) than for positive (97.5%) and more accurate for low clutter (98.1%) than for high clutter (97.6%). Two effects not found to be significant for response time but significant for accuracy are failure mode and failure type. The on mode produces significantly more accurate responses (98.1%) than the off mode (97.4%), and horizontal line failures yield significantly less accuracy than any of the other three failure types (see Table 5 and Figure 4). Cell and vertical line failures do not reduce accuracy below that obtained with no failures.

Again, there is a significant failure mode by failure type interaction. Failure mode shows an effect only for horizontal line failures (see Table 6 and Figure 5), performance for the off mode being more accurate (97.2%) than for the on (95.1%).

Table 1

ANOVA Summary Table for Random Search Task,
Dependent Variable = Response Time

SOURCE	*df*	*MS*	*F*	*p*
Subjects (SUB)	11	75.26		
Polarity (POL)	1	371.32	26.18	0.0003
SUB*POL	11	14.18		
Background Clutter Level (BCL)	1	1556.46	163.18	0.0001
SUB*BCL	11	9.54		
MODE	1	49.42	4.17	0.0658
SUB*MODE	11	11.84		
TYPE	3	59.95	1.95	0.1415
SUB*TYPE	33	30.82		
Percent (PCT)	2	35.52	3.73	0.06*
SUB*PCT	22	9.52		
POL*BCL	1	2.60	0.18	0.6801
SUB*POL*BCL	11	14.50		
POL*MODE	1	8.98	1.32	0.2750
SUB*POL*MODE	11	6.80		
POL*PCT	2	0.95	0.11	0.8952
SUB*POL*PCT	22	8.49		
POL*TYPE	3	3.54	0.32	0.8084
SUB*POL*TYPE	33	10.94		
BCL*MODE	1	5.22	0.39	0.5464
SUB*BCL*MODE	11	13.48		
BCL*PCT	2	2.24	0.34	0.7184
SUB*BCL*PCT	22	6.66		
BCL*TYPE	3	2.14	0.18	0.9111
SUB*BCL*TYPE	33	12.11		
MODE*PCT	2	12.11	1.60	0.2242
SUB*MODE*PCT	22	7.56		
MODE*TYPE	3	41.82	3.95	0.0163
SUB*MODE*TYPE	33	10.58		
PCT*TYPE	6	20.67	2.09	<0.05**
SUB*PCT*TYPE	66	9.88		

Table 1 (continued)

POL*BCL*MODE	1	0.73	0.26	0.6210
SUB*POL*BCL*MODE	11	2.81		
POL*BCL*PCT	2	1.54	0.26	0.7726
SUB*POL*BCL*PCT	22	5.88		
POL*BCL*TYPE	3	16.90	1.87	0.1546
SUB*POL*BCL*TYPE	33	9.06		
POL*MODE*TYPE	3	17.87	1.94	0.1424
SUB*POL*MODE*TYPE	33	9.21		
POL*MODE*PCT	2	0.88	0.10	0.9098
SUB*POL*MODE*PCT	22	9.29		
POL*PCT*TYPE	6	1.67	0.21	0.9724
SUB*POL*PCT*TYPE	66	7.92		
BCL*MODE*PCT	2	23.16	2.12	0.1434
SUB*BCL*MODE*PCT	22	10.90		
BCL*PCT*TYPE	6	4.89	0.69	0.6586
SUB*BCL*PCT*TYPE	66	7.09		
BCL*MODE*TYPE	3	0.85	0.07	0.9772
SUB*BCL*MODE*TYPE	33	12.76		
MODE*PCT*TYPE	6	7.68	0.97	0.4498
SUB*MODE*PCT*TYPE	66	7.89		
POL*BCL*MODE*PCT	2	0.02	0.00	0.9972
SUB*POL*BCL*MODE*PCT	22	6.07		
POL*BCL*MODE*TYPE	3	6.02	0.68	0.5711
SUB*POL*BCL*MODE*TYPE	33	8.87		
POL*MODE*PCT*TYPE	6	6.75	0.57	0.7525
SUB*POL*MODE*PCT*TYPE	66	11.84		
POL*BCL*PCT*TYPE	6	1.28	0.15	0.9881
SUB*POL*BCL*PCT*TYPE	66	8.43		
BCL*MODE*PCT*TYPE	6	9.57	1.40	0.2273
SUB*BCL*MODE*PCT*TYPE	66	6.83		
POL*BCL*MODE*PCT*TYPE	6	13.36	1.60	0.1610
SUB*POL*BCL*MODE*PCT*TYPE	66	8.35		
	1151			

*Greenhouse and Geisser (1959) ε = 0.8357
*Greenhouse and Geisser (1959) ε = 0.7723

Table 2

Results of Newman-Keuls Test for Percent Failure, Random Search Task, Dependent Variable = Response Time

Percent failure	Means, seconds
2	6.95
3	6.84
1	6.38

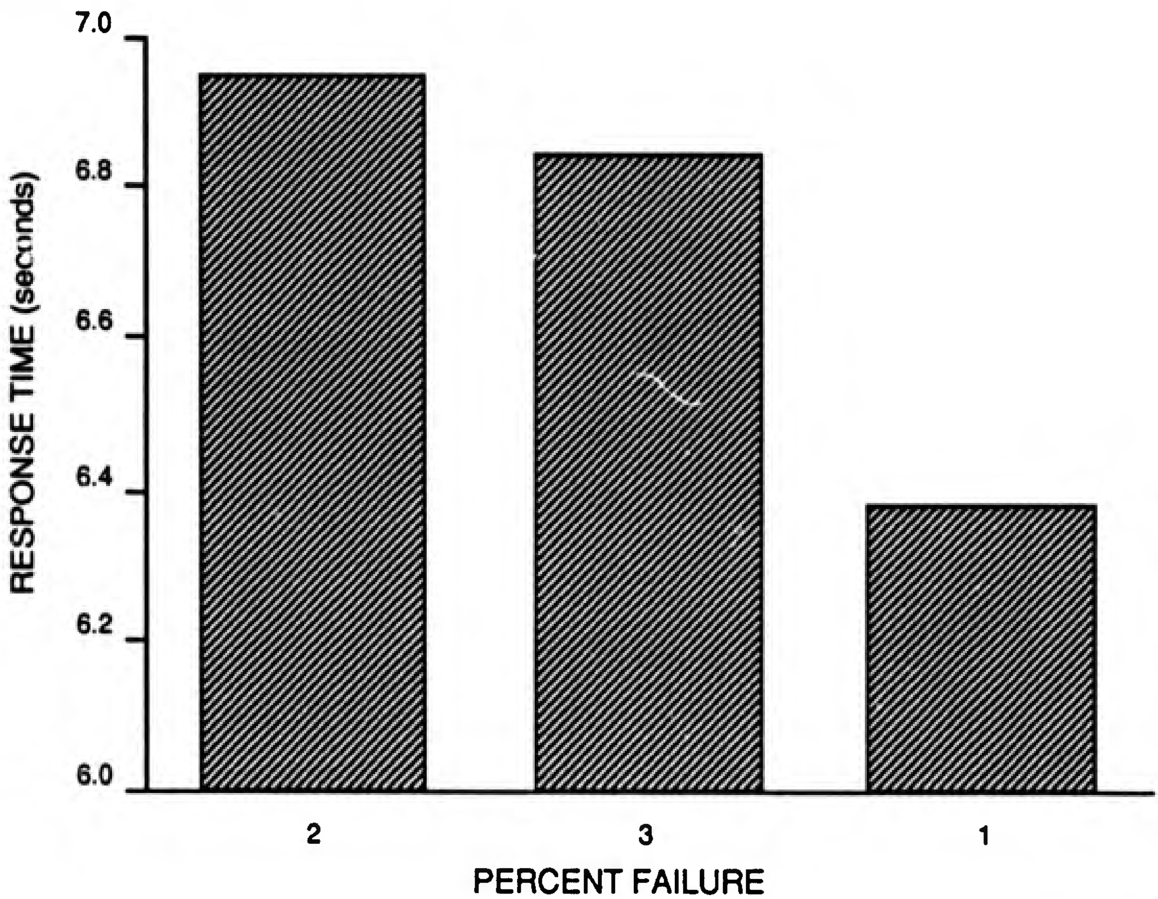

Figure 2. The effect of percent failure on response time for the random search task.

Table 3

Results of Simple Effect F Tests on Failure Mode for Each Failure Type, Random Search Task, Dependent Variable = Response Time

Failure type	MS_{Mode}	*F*	*p*
None	3.40	0.32	> 0.05
Horizontal line	0.65	0.06	> 0.05
Vertical line	19.34	1.83	> 0.05
Cell	151.51	14.32	< 0.01

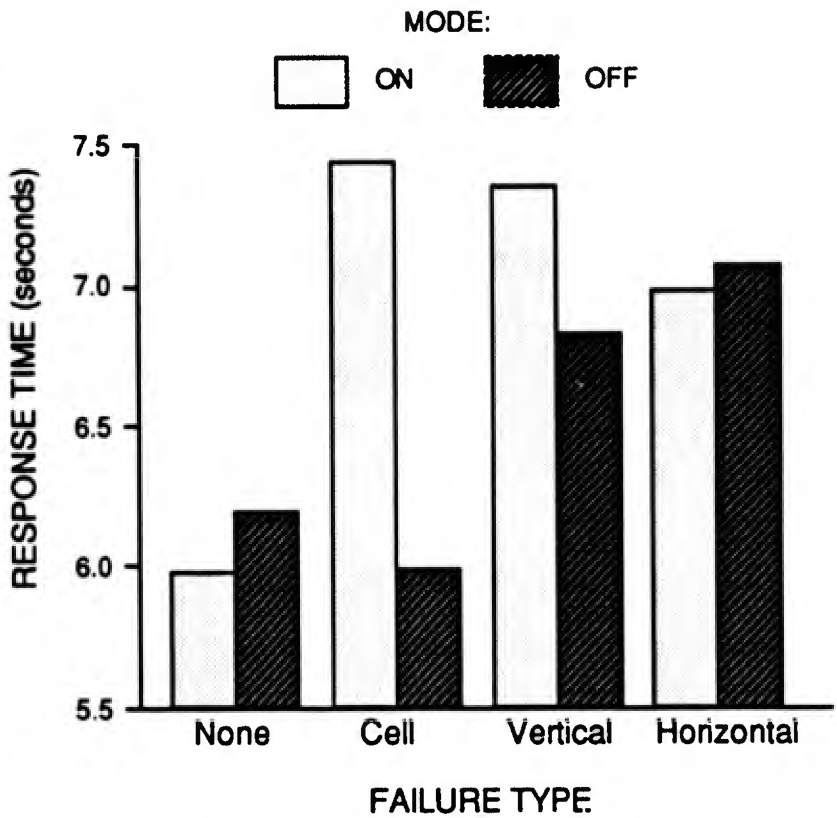

Figure 3. The effect of the failure type x failure mode interaction on response time for the random search task.

Table 4

ANOVA Summary Table for Random Search Task, Dependent Variable = Percent Correct Responses

SOURCE	*df*	*MS*	*F*	*p*
Subjects (SUB)	11	0.0111		
Polarity (POL)	1	0.0097	5.34	0.0419
SUB*POL	11	0.0018		
Background Clutter Level (BCL)	1	0.0070	5.27	0.0426
SUB*BCL	11	0.0013		
MODE	1	0.0151	8.65	0.0135
SUB*MODE	11	0.0017		
TYPE	3	0.0370	14.39	0.0001
SUB*TYPE	33	0.0026		
Percent (PCT)	2	0.0027	1.08	0.3564
SUB*PCT	22	0.0025		
POL*BCL	1	0.0041	2.13	0.1743
SUB*POL*BCL	11	0.0019		
POL*MODE	1	0.0029	1.92	0.1920
SUB*POL*MODE	11	0.0015		
POL*PCT	2	0.0024	1.21	0.3205
SUB*POL*PCT	22	0.0020		
POL*TYPE	3	0.0030	1.30	0.2925
SUB*POL*TYPE	33	0.0023		
BCL*MODE	1	0.0015	0.43	0.5190
SUB*BCL*MODE	11	0.0035		
BCL*PCT	2	0.0022	0.98	0.3934
SUB*BCL*PCT	22	0.0023		
BCL*TYPE	3	0.0029	2.04	0.1280
SUB*BCL*TYPE	33	0.0014		
MODE*PCT	2	0.0002	0.13	0.8642
SUB*MODE*PCT	22	0.0012		
MODE*TYPE	3	0.0063	3.29	< 0.05*
SUB*MODE*TYPE	33	0.0019		
PCT*TYPE	6	0.0082	6.08	0.0001
SUB*PCT*TYPE	66	0.0013		

Table 4 (continued)

POL*BCL*MODE	1	0.0047	3.17	0.1012
SUB*POL*BCL*MODE	11	0.0015		
POL*BCL*PCT	2	0.0007	0.39	0.6897
SUB*POL*BCL*PCT	22	0.0018		
POL*BCL*TYPE	3	0.0002	0.09	0.9694
SUB*POL*BCL*TYPE	33	0.0027		
POL*MODE*TYPE	3	0.0024	1.07	0.3742
SUB*POL*MODE*TYPE	33	0.0022		
POL*MODE*PCT	2	0.0047	3.59	>0.05**
SUB*POL*MODE*PCT	22	0.0013		
POL*PCT*TYPE	6	0.0012	0.49	0.8116
SUB*POL*PCT*TYPE	66	0.0024		
BCL*MODE*PCT	2	0.0060	2.49	0.1053
SUB*BCL*MODE*PCT	22	0.0024		
BCL*PCT*TYPE	6	0.0019	1.51	0.1901
SUB*BCL*PCT*TYPE	66	0.0013		
BCL*MODE*TYPE	3	0.0034	1.08	0.3702
SUB*BCL*MODE*TYPE	33	0.0031		
MODE*PCT*TYPE	6	0.0001	0.09	0.9973
SUB*MODE*PCT*TYPE	66	0.0015		
POL*BCL*MODE*PCT	2	0.0022	1.08	0.3589
SUB*POL*BCL*MODE*PCT	22	0.0020		
POL*BCL*MODE*TYPE	3	0.0002	0.10	0.9625
SUB*POL*BCL*MODE*TYPE	33	0.0017		
POL*MODE*PCT*TYPE	6	0.0007	0.56	0.7642
SUB*POL*MODE*PCT*TYPE	66	0.0012		
POL*BCL*PCT*TYPE	6	0.0036	2.41	>0.05***
SUB*POL*BCL*PCT*TYPE	66	0.0015		
BCL*MODE*PCT*TYPE	6	0.0020	1.05	0.4064
SUB*BCL*MODE*PCT*TYPE	66	0.0019		
POL*BCL*MODE*PCT*TYPE	6	0.0031	2.13	0.0607
SUB*POL*BCL*MODE*PCT*TYPE	66	0.0014		
	1151			

* Greenhouse and Geisser (1959) $\varepsilon = 0.7927$
** Greenhouse and Geisser (1959) $\varepsilon = 0.8202$
***Greenhouse and Geisser (1959) $\varepsilon = 0.7929$

Table 5

Results of Newman-Keuls Tests on Failure Type, Random Search Task, Dependent Variable = Percent Correct Responses

Failure type	Means, seconds
None	98.6 (A)
Cell	98.3 (A)
Vertical line	98.2 (A)
Horizontal line	96.1 (B)

Note. Means for failure types sharing a common letter in parentheses are not significantly different, $p > .05$

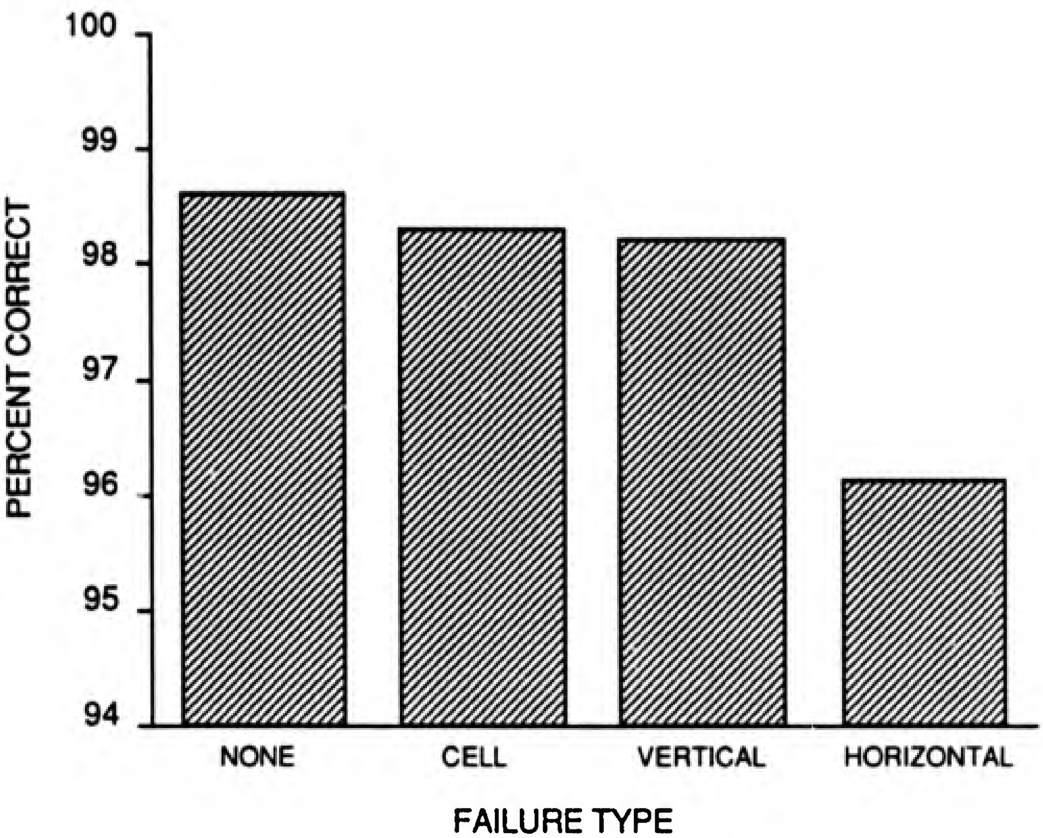

Figure 4. The effect of failure type on response accuracy for the random search task.

Table 6

Results of Simple Effect F Tests on Mode for Each Failure Type, Random Search Task, Dependent Variable = Percent Correct Responses

Failure type	MS_{Mode}	F	p
None	0.00	0.00	> 0.05
Horizontal line	0.0313	16.47	< 0.01
Vertical line	0.0024	1.26	> 0.05
Cell	0.0004	0.21	> 0.05

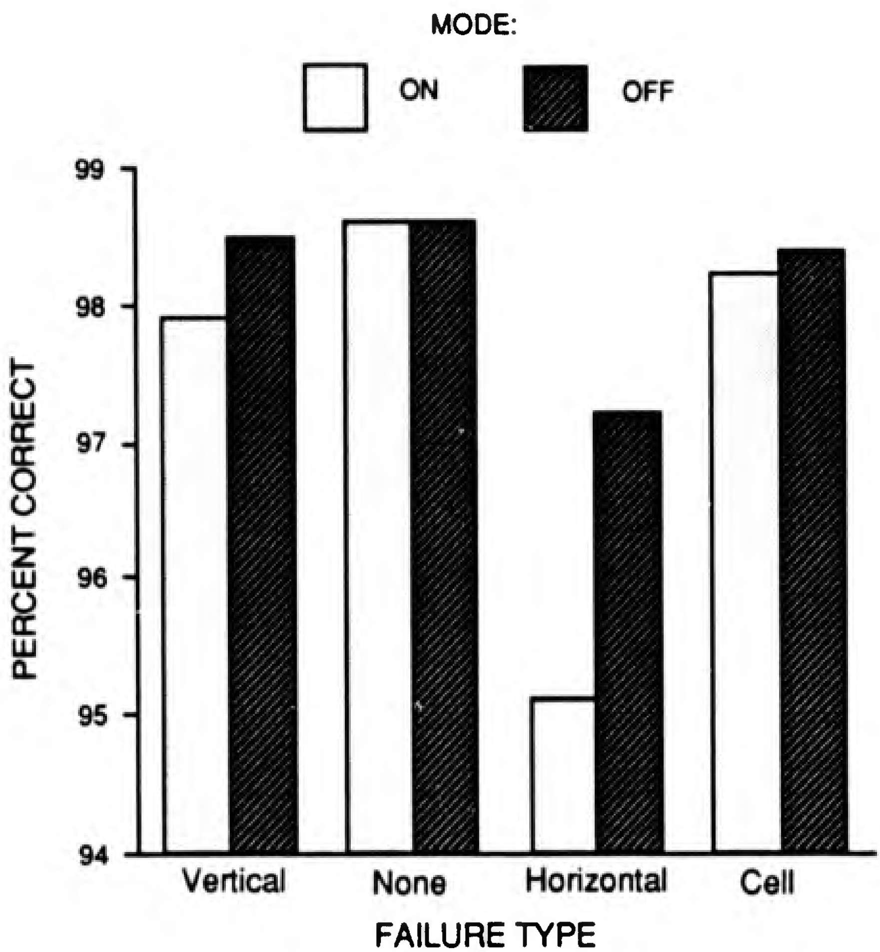

Figure 5. The effect of the failure type x failure mode interaction on response accuracy for the random search task.

Percent failure also interacts significantly with failure type and, as with response times, percent only shows an effect for horizontal line failures (see Table 7 and Figure 6). The Newman-Keuls test reveals that all three failure percents differ significantly within this failure type (see Table 8).

Information-Extraction Task

Response Time

Polarity, background clutter level, and percent failure are statistically significant, as they were in the random search task ANOVA for response time, and all display the same trends (see Table 9). Performance with negative contrast is faster (4.13 seconds) than with positive (4.45 seconds); performance with low clutter is faster (3.83 seconds) than with high (4.74 seconds). A failure rate of 1% produces significantly faster responses (4.14 seconds) than a failure rate of either 2% (4.31 seconds) or 3% (4.41 seconds, see Table 10 and Figure 7).

In addition, the effect of failure mode shows significance with the off mode producing faster (4.16 seconds) times than the on mode (4.41 seconds).

The failure mode by failure type interaction is again significant. Failure mode displays an effect for vertical line (off mode = 4.12 seconds; on mode = 4.65 seconds) and cell (off mode = 4.00 seconds; on mode = 4.51 seconds) failure types, off mode producing consistently faster responses (see Table 11 and Figure 10).

The three-way interaction among background clutter level, failure percent, and failure type is also statistically significant. Simple effect F tests show that percent failure is significant for both types of line failures at high clutter and for horizontal line failures at low clutter (see Table 12 and Figures 9 through 12). For the horizontal line failure with low clutter and the vertical line failure with high clutter combinations, the Newman-Keuls test found that a 3% failure rate causes significantly slower responses than both 1% and 2% failures (see Table 13). For the horizontal line, high clutter combination, the 2% failure rate causes slower responses than a failure of 1%, but not 3% (see Table 13).

Table 7

Results of Simple Effect F Tests on Percent Failure for Each Failure Type, Random Search Task, Dependent Variable = Percent Correct Responses

Failure type	$MS_{Percent}$	F	p
None	0.002	1.54	> 0.05
Horizontal line	0.0241	18.54	< 0.01
Vertical line	0.00002	0.02	> 0.05
Cell	0.0012	0.92	> 0.05

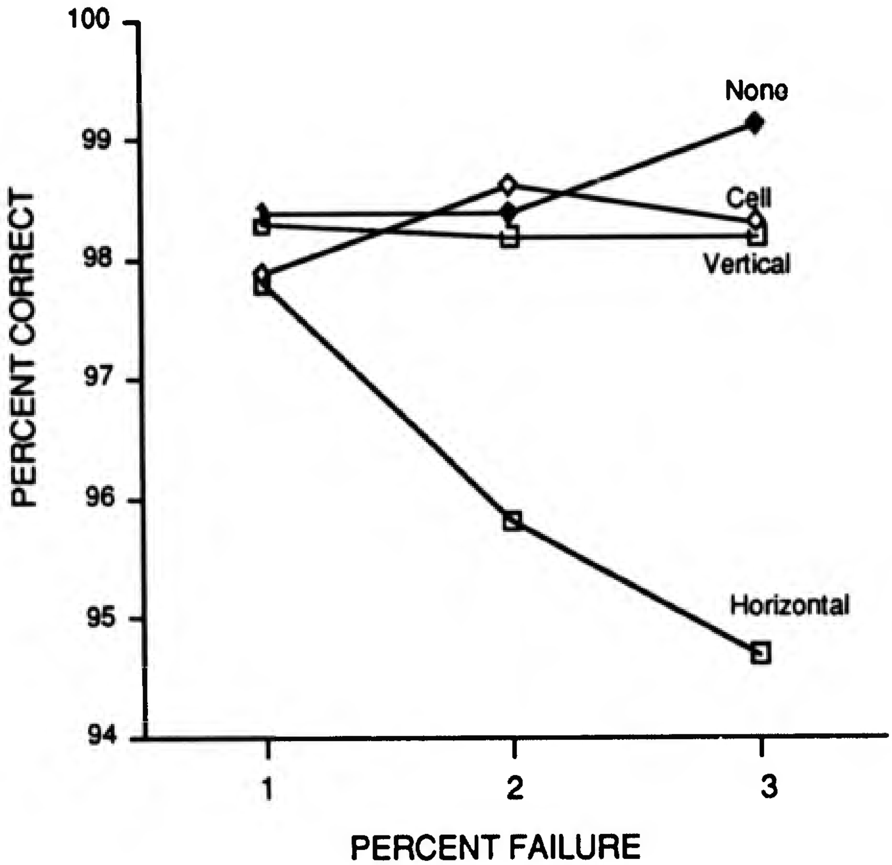

Figure 6. The effect of the failure type x percent failure interaction on response accuracy for the random search task.

Table 8

Results of Newman-Keuls Tests on Percent Failure for Horizontal Line Failure, Random Search Task, Dependent Variable = Percent Correct Responses

Failure percent	Means, seconds
1	97.8 (A)
2	95.8 (B)
3	94.7 (C)

Note. Means sharing a common letter in parentheses are not significantly different, $p > 0.05$.

Table 9

ANOVA Summary Table for Information-Extraction Task, Dependent Variable = Response Time

SOURCE	*df*	*MS*	*F*	*p*
Subjects (SUB)	11	231.27		
Polarity (POL)	1	28.89	48.52	0.0001
SUB*POL	11	0.60		
Background Clutter Level (BCL)		238.55	73.61	0.0001
SUB*BCL	11	3.24		
MODE	1	18.21	26.60	0.0003
SUB*MODE	11	0.68		
TYPE	3	2.51	0.27	0.0851
SUB*TYPE	33	9.23		
Percent (PCT)	2	7.33	6.91	0.0047
SUB*PCT	22	1.06		
POL*BCL	1	0.04	0.06	0.8196
SUB*POL*BCL	11	0.69		
POL*MODE	1	0.09	0.24	0.6354
SUB*POL*MODE	11	0.38		
POL*PCT	2	1.08	1.63	0.2189
SUB*POL*PCT	22	0.66		
POL*TYPE	3	0.16	0.26	0.8547
SUB*POL*TYPE	33	0.61		
BCL*MODE	1	1.54	4.60	0.0551
SUB*BCL*MODE	11	0.33		
BCL*PCT	2	0.02	0.05	0.9573
SUB*BCL*PCT	22	0.28		
BCL*TYPE	3	1.07	1.05	0.3824
SUB*BCL*TYPE	33	1.01		
MODE*PCT	2	1.21	1.41	0.2660
SUB*MODE*PCT	22	0.86		
MODE*TYPE	3	7.16	5.02	0.0056
SUB*MODE*TYPE	33	1.43		
PCT*TYPE	6	1.54	1.17	0.3357
SUB*PCT*TYPE	66	1.32		

Table 9 (continued)

POL*BCL*MODE	1	0.06	0.09	0.7755
SUB*POL*BCL*MODE	11	0.65		
POL*BCL*PCT	2	0.03	0.05	0.9479
SUB*POL*BCL*PCT	22	0.57		
POL*BCL*TYPE	3	1.16	0.92	0.4440
SUB*POL*BCL*TYPE	33	1.26		
POL*MODE*TYPE	3	0.51	0.84	0.4795
SUB*POL*MODE*TYPE	33	0.61		
POL*MODE*PCT	2	1.78	2.41	0.1135
SUB*POL*MODE*PCT	22	0.74		
POL*PCT*TYPE	6	1.08	1.85	0.1023
SUB*POL*PCT*TYPE	66	0.58		
BCL*MODE*PCT	2	1.58	2.07	0.1507
SUB*BCL*MODE*PCT	22	0.76		
BCL*PCT*TYPE	6	2.66	3.06	< 0.05*
SUB*BCL*PCT*TYPE	66	0.87		
BCL*MODE*TYPE	3	0.27	0.50	0.6869
SUB*BCL*MODE*TYPE	33	0.55		
MODE*PCT*TYPE	6	0.37	0.45	0.8426
SUB*MODE*PCT*TYPE	66	0.81		
POL*BCL*MODE*PCT	2	2.36	2.38	0.1161
SUB*POL*BCL*MODE*PCT	22	0.99		
POL*BCL*MODE*TYPE	3	0.53	0.79	0.5101
SUB*POL*BCL*MODE*TYPE	33	0.68		
POL*MODE*PCT*TYPE	6	1.00	1.33	0.1211
SUB*POL*MODE*PCT*TYPE	66	0.75		
POL*BCL*PCT*TYPE	6	1.41	1.76	0.2555
SUB*POL*BCL*PCT*TYPE	66	0.80		
BCL*MODE*PCT*TYPE	6	0.46	0.52	0.7899
SUB*BCL*MODE*PCT*TYPE	66	0.94		
POL*BCL*MODE*PCT*TYPE	6	0.58	1.09	0.3796
SUB*POL*BCL*MODE*PCT*TYPE	66	0.53		
	1151			

*Greenhouse and Geisser (1959) $\varepsilon = 0.7823$

Table 10

Results of Newman-Keuls Test for Percent Failure, Information-Extraction Task, Dependent Variable = Response Time

Failure percent	Means, seconds
3	4.41 (A)
2	4.31 (A)
1	4.14 (B)

Note. Means sharing a common letter in parentheses are not significantly different, $p > 0.05$.

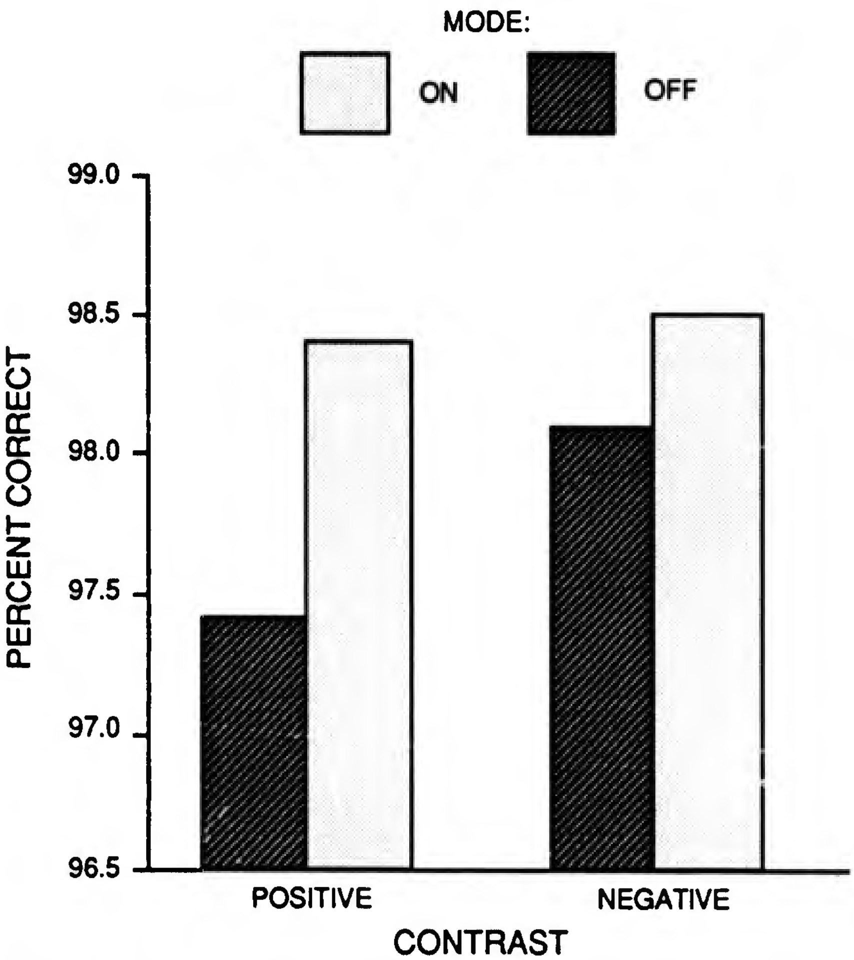

Figure 7. The effect of percent failure on response time for the information-extraction task.

Table 11

Results of Simple Effect F Tests on Failure Mode for Each Failure Type, Information-Extraction Task, Dependent Variable = Reaction Time

Failure type	MS_{Mode}	F	p
None	0.58	0.41	> 0.05
Horizontal line	0.23	0.16	> 0.05
Vertical line	20.22	14.14	< 0.01
Cell	18.67	13.06	< 0.01

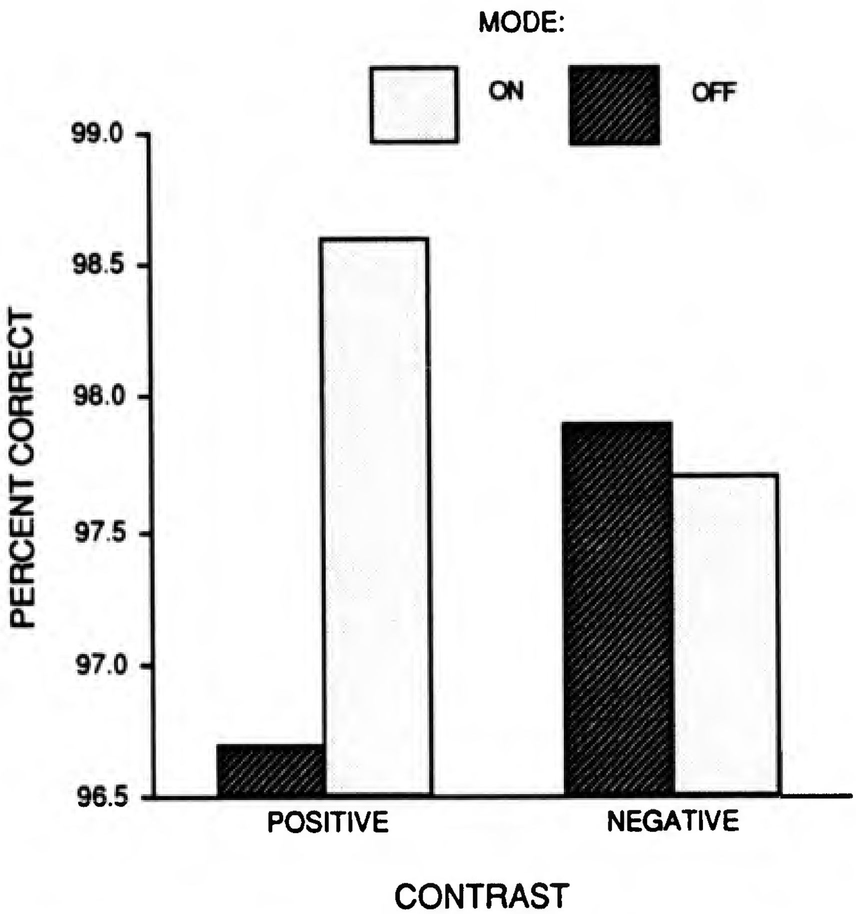

Figure 8. The effect of the failure mode x failure type interaction on response time for the information-extraction task.

Table 12

Results of Simple Effect F Tests on Percent Failure for all Combinations of Failure Type and Background Clutter Level, Information-Extraction Task, Dependent Variable = Reaction Time

Failure type	BCL	$MS_{Percent}$	F	p
None	Low	0.02	0.03	> 0.05
None	High	0.73	0.84	> 0.05
Horizontal line	Low	5.21	5.99	< 0.01
Horizontal line	High	4.17	4.79	< 0.05
Vertical line	Low	2.05	2.36	> 0.05
Vertical line	High	6.57	7.55	< 0.01
Cell	Low	0.62	0.72	> 0.05
Cell	High	0.55	0.63	> 0.05

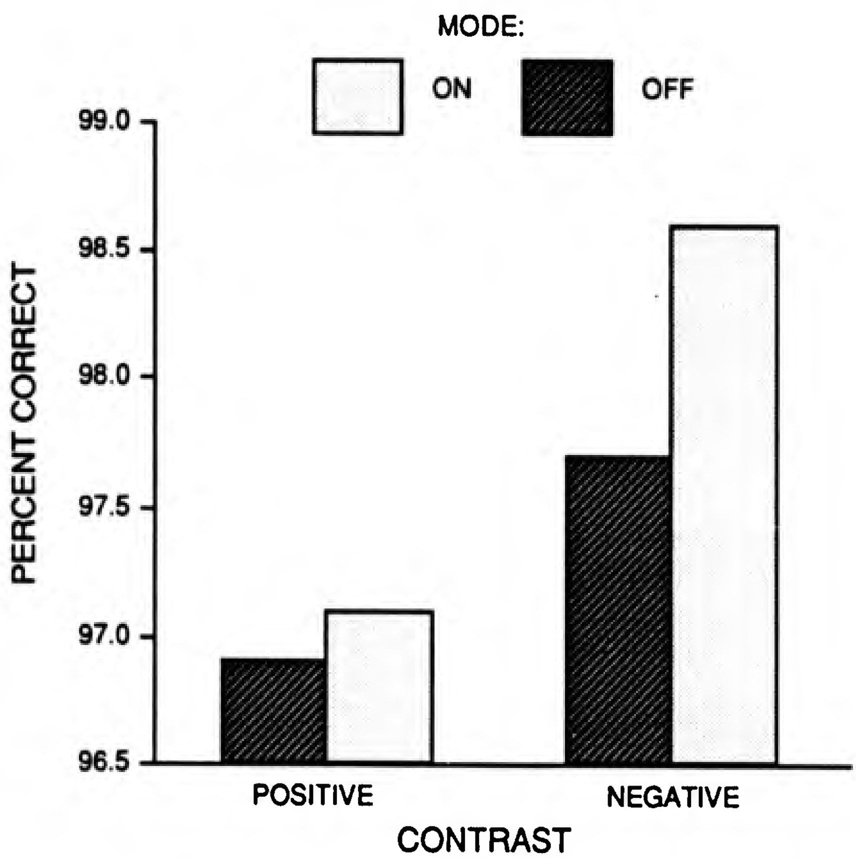

Figure 9. The effect of the percent failure x background clutter level interaction on response time for the information-extraction task, no failure condition.

Table 13

Results of Newman-Keuls Tests on Percent Failure for Background Clutter Levels and Vertical and Horizontal Line Failures, Information-Extraction Task, Dependent Variable = Response Time

Failure type	BCL	Percent	Means, seconds
Vertical line	High	3	5.282 (A)
		1	4.651 (B)
		2	4.632 (B)
Horizontal line	Low	3	4.196 (A)
		2	3.673 (B)
		1	3.588 (B)
Horizontal line	High	2	5.118 (A)
		3	4.900 (AB)
		1	4.535 (B)

Note. Means for each background clutter level sharing the same letter in parentheses are not significantly different, $p > 0.01$.

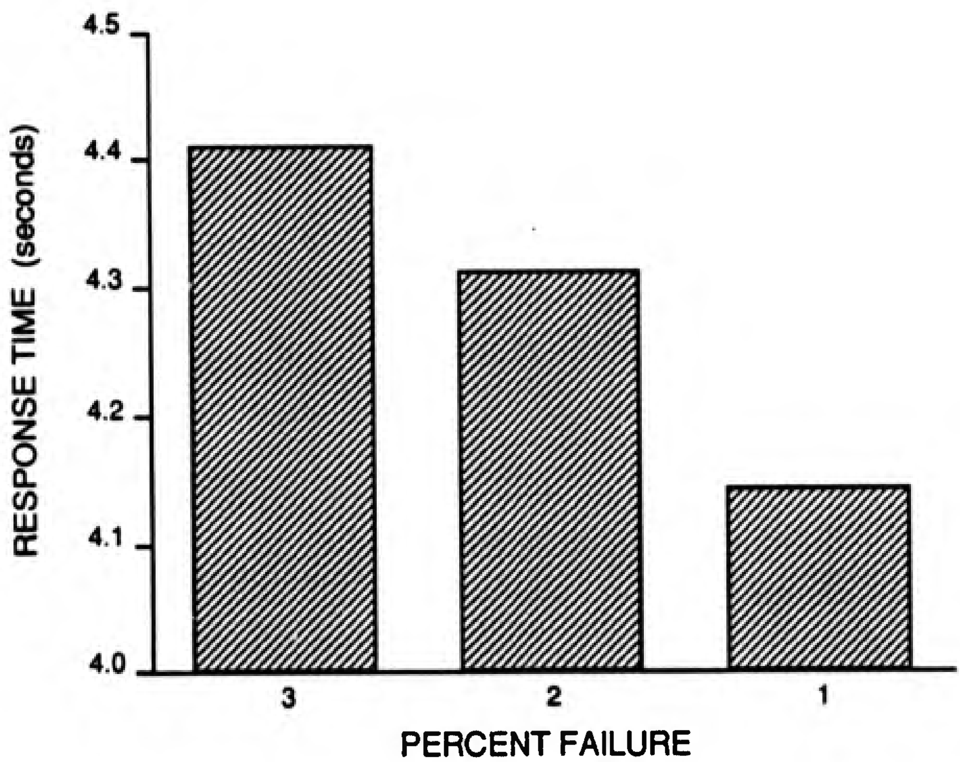

Figure 10. The effect of the percent failure x background clutter level interaction on response time for the information-extraction task, horizontal line failures only.

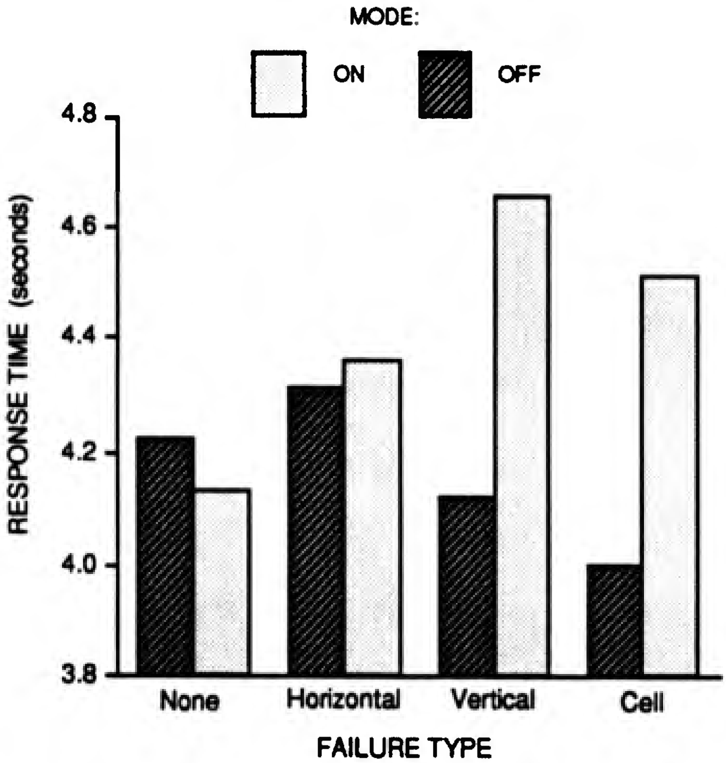

Figure 11. The effect of the percent failure x background clutter level interaction on response time for the information-extraction task, vertical line failures only.

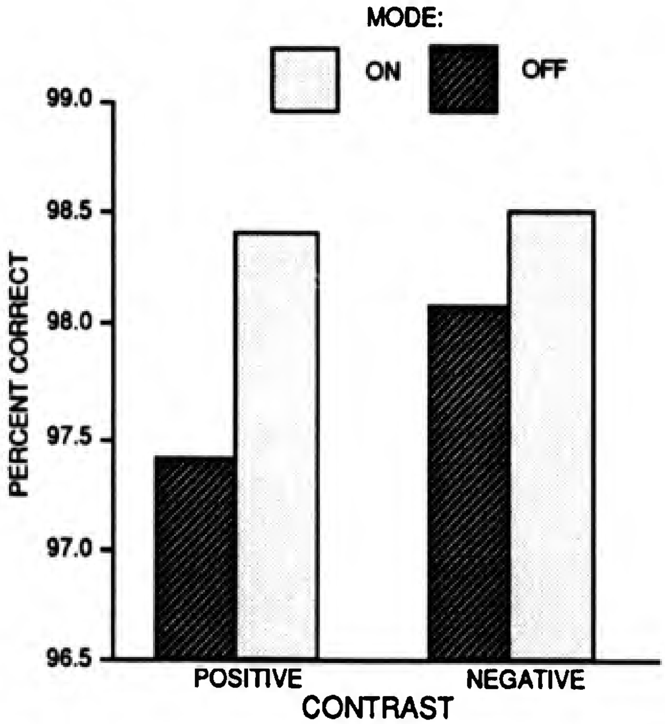

Figure 12. The effect of the percent failure x background clutter level interaction on response time for the information-extraction task, cell failures only.

Response Accuracy

For the dependent measure of response accuracy, only the main effect of background clutter level ($p < 0.0001$) is significant (see Table 14). Low clutter produces significantly more accurate (90.3%) responses than high clutter (86.1%).

The only significant higher order interaction is that among polarity, background clutter level, and failure mode. The simple effect F tests found that polarity is significant for low clutter and on mode, and for high clutter and off mode combinations (see Table 15, Figures 13 and 14), with responses for negative contrast being more accurate than for positive contrast for both combinations (91.7% > 88.5% and 86.7% > 84.1%, respectively).

It should be noted that this is the only one of the four ANOVAs in which the failure mode by failure type interaction is not significant, albeit by a small margin ($p = 0.0612$).

DISCUSSION

One of the less obvious, yet important, results of the study is the difference in response time and response accuracy between the two experiments. Performance was slower yet more accurate for the random search task than for the information-extraction task. This result is really not too surprising since there is a smaller potential target area in the information-extraction task. That is, in this task, the subject is told that the target is near some reference information on the map, while in the random search task, the target may be located anywhere on the screen. Thus, response times should be shorter in the information-extraction task.

Further, accuracy should be better for the random search task since there is a unique target symbol to search for, whereas in the information-extraction task, four to seven symbols are identical in appearance to the target, but only one is closest to the indicated reference on the map. Thus, the subject would quickly find a symbol that looks like the target, but it simply would not be the closest to the target reference.

Background Clutter Level

Background clutter level shows expected results, with a high clutter condition requiring more search time and producing less accuracy than a low clutter condition for both tasks. Search time seems to suffer more than accuracy as clutter is increased. The differences were not large for accuracy, but response times were 20% and 30% faster for low clutter than for high clutter in the information-extraction and random search tasks, respectively. It should be noted that the displays were really not as cluttered as they might be (approximately 2.9% of the display pixels on for low clutter and approximately 3.8% of the display pixels on for high clutter). With such seemingly low overall levels of clutter as were present in this study, asymptotic degradation of performance certainly was not attained. It is reasonable to assume, then, that performance would be increasingly degraded if clutter were increased above the modest levels used in these experiments.

Table 14

ANOVA Summary Table for Information-Extraction Task, Dependent Variable = Percent Correct Responses

SOURCE	*df*	*MS*	*F*	*p*
Subjects (SUB)	11	0.0777		
Polarity (POL)	1	0.0467	2.07	0.1779
SUB*POL	11	0.0225		
Background Clutter Level (BCL)	1	0.4931	32.31	0.0001
SUB*BCL	11	0.0153		
MODE	1	0.0087	0.85	0.3761
SUB*MODE	11	0.0102		
TYPE	3	0.0390	2.94	>0.05*
SUB*TYPE	33	0.0133		
Percent (PCT)	2	0.0133	1.34	0.2819
SUB*PCT	22	0.0099		
POL*BCL	1	0.0024	0.29	0.5984
SUB*POL*BCL	11	0.0082		
POL*MODE	1	0.0000**	0.00	1.0000
SUB*POL*MODE	11	0.0053		
POL*PCT	2	0.0142	1.61	0.2221
SUB*POL*PCT	22	0.0088		
POL*TYPE	3	0.0042	0.52	0.6716
SUB*POL*TYPE	33	0.0081		
BCL*MODE	1	0.0232	2.97	0.1129
SUB*BCL*MODE	11	0.0078		
BCL*PCT	2	0.0110	1.21	0.3186
SUB*BCL*PCT	22	0.0091		
BCL*TYPE	3	0.0105	1.34	0.2779
SUB*BCL*TYPE	33	0.0078		
MODE*PCT	2	0.0214	2.08	0.1493
SUB*MODE*PCT	22	0.0103		
MODE*TYPE	3	0.0142	2.71	0.0612
SUB*MODE*TYPE	33	0.0053		
PCT*TYPE	6	0.0087	0.75	0.6116
SUB*PCT*TYPE	66	0.0116		

Table 14 (continued)

POL*BCL*MODE	1	0.0703	7.91	0.0169
SUB*POL*BCL*MODE	11	0.0089		
POL*BCL*PCT	2	0.0084	0.94	0.4084
SUB*POL*BCL*PCT	22	0.0089		
POL*BCL*TYPE	3	0.0010	0.09	0.9663
SUB*POL*BCL*TYPE	33	0.0112		
POL*MODE*TYPE	3	0.0099	1.52	0.2273
SUB*POL*MODE*TYPE	33	0.0065		
POL*MODE*PCT	2	0.0167	2.40	0.1137
SUB*POL*MODE*PCT	22	0.0069		
POL*PCT*TYPE	6	0.0065	0.80	0.5747
SUB*POL*PCT*TYPE	66	0.0082		
BCL*MODE*PCT	2	0.0151	1.67	0.2107
SUB*BCL*MODE*PCT	22	0.0090		
BCL*PCT*TYPE	6	0.0075	0.92	0.4905
SUB*BCL*PCT*TYPE	66	0.0081		
BCL*MODE*TYPE	3	0.0102	1.65	0.1979
SUB*BCL*MODE*TYPE	33	0.0062		
MODE*PCT*TYPE	6	0.0069	0.60	0.7332
SUB*MODE*PCT*TYPE	66	0.0116		
POL*BCL*MODE*PCT	2	0.0079	1.34	0.2828
SUB*POL*BCL*MODE*PCT	22	0.0059		
POL*BCL*MODE*TYPE	3	0.0021	0.20	0.8951
SUB*POL*BCL*MODE*TYPE	33	0.0107		
POL*MODE*PCT*TYPE	6	0.0051	0.88	0.3856
SUB*POL*MODE*PCT*TYPE	66	0.0058		
POL*BCL*PCT*TYPE	6	0.0079	1.08	0.5167
SUB*POL*BCL*PCT*TYPE	66	0.0073		
BCL*MODE*PCT*TYPE	6	0.0049	0.52	0.7939
SUB*BCL*MODE*PCT*TYPE	66	0.0094		
POL*BCL*MODE*PCT*TYPE	6	0.0062	0.64	0.6974
SUB*POL*BCL*MODE*PCT*TYPE	66	0.0096		
	1151			

* Greenhouse and Geisser (1959) ε = 0.8672
**This unlikely value was calculated independently to assure its validity.

Table 15

Results of Simple Effect F Tests on Polarity for all Combinations of Background Clutter Level and Failure Mode, Information-Extraction Task, Dependent Variable = Percent Correct Responses

BCL	Mode	$MS_{Polarity}$	F	p
Low	Off	0.000	0.00	> 0.05
Low	On	0.070	7.87	< 0.05
High	Off	0.047	5.25	< 0.05
High	On	0.002	0.27	> 0.05

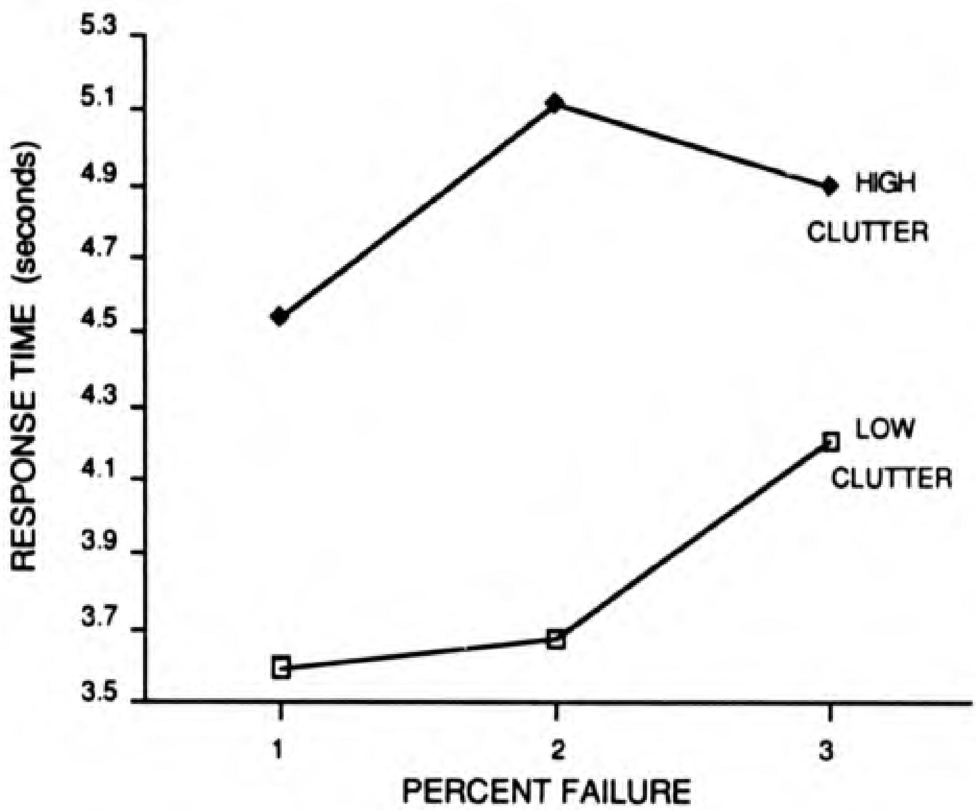

Figure 13. The effect of the background clutter x polarity interaction on response time for the information-extraction task, on mode only.

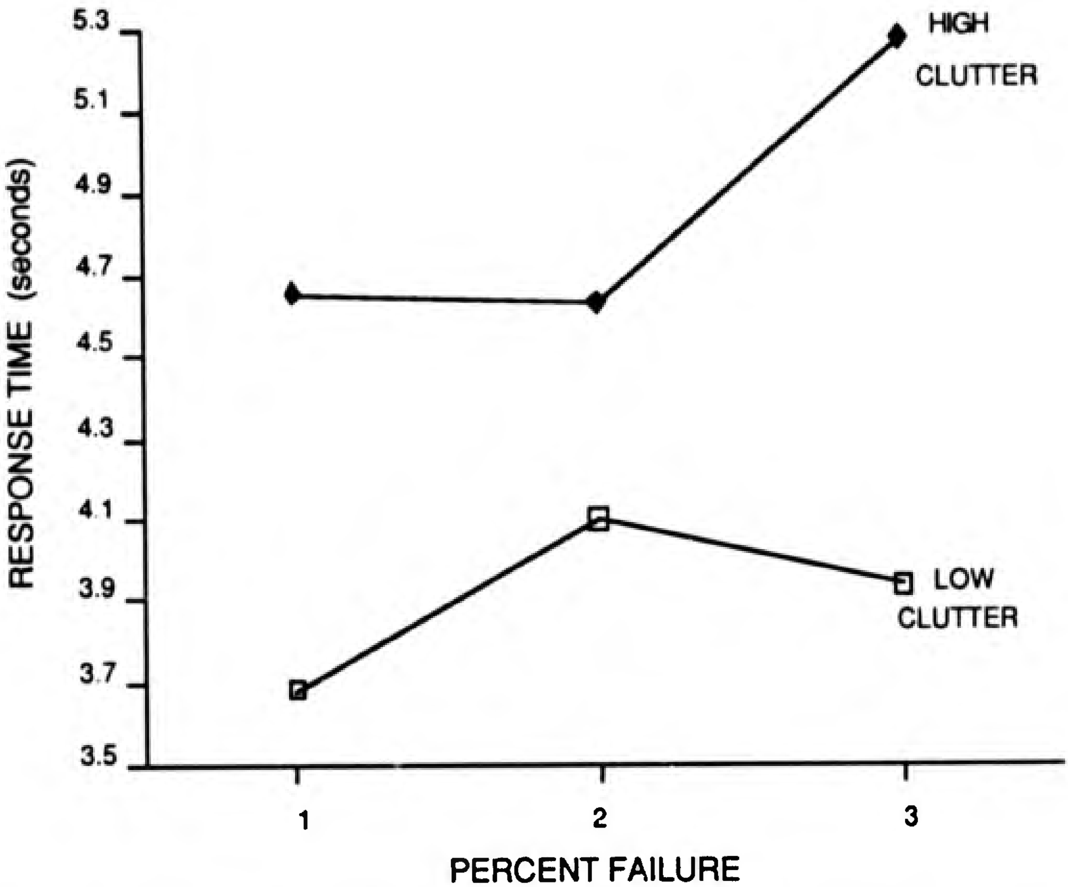

Figure 14. The effect of the background clutter x polarity interaction on response time for the information-extraction task, off mode only.

Polarity

As in the earlier polarity studies (Decker, Dye, Lloyd, & Snyder, 1991; Decker, Kelly, Kurokawa, & Snyder, 1991; Lloyd et al., 1988), negative contrast produces better visual task performance than does positive contrast. Response times were 16% faster in the random search task and 7% faster in the information extraction task. The difference between the two polarities for accuracy in the random search task was quite small (0.6%).

Snyder (1980) indicated that increasing space average luminance increases contrast sensitivity. It is possible that the higher space average luminance that exists in the negative contrast condition increases contrast sensitivity beyond that in the positive contrast condition. This difference in sensitivity may lead to greater legibility for the images appearing in negative contrast, thereby increasing the task performance for that condition.

Certainly more research about the issue of polarity needs to be performed, but the results of this and previous studies seem to give a consistent nod to negative contrast. Those performing future research in this area are again cautioned that stroke width, modulation, and luminances for both polarities must be set as equal as possible to provide meaningful and unconfounded results.

Failure Variables

Failures affect only the symbols for the random search task, but they affect both the symbols and the target references for the information extraction task. Thus, one might expect the effects of failure mode and failure type to be greater for the information-extraction task.

Horizontal line failures seem to have the greatest effect on performance overall. The cause of this effect can be explained by referring to the symbol set in Figure 1. Half the symbols have a horizontal bar over or through them. This bar is used to distinguish the enemy symbols from the friendly ones. It should be readily apparent that of the three types of failures (excluding no failure), the type that presents the most potential interference to the greatest number of symbols is the horizontal line failure.

Interpreting the effects of failure mode is somewhat difficult. The results show that on failures produced slower response times than off failures in the information-extraction task. This result is not surprising and it agrees with the results of the recent alphanumeric search studies (Abramson & Snyder, 1984; Decker, Dye, Lloyd, & Snyder, 1991; Lloyd et al., 1988).

In the random search task, there was greater accuracy for on failures than for off failures. This difference is not consistent with previous research. However, while the results for accuracy were statistically significant, for all practical purposes there was very little difference in accuracy for the two modes. The on mode averaged 98.1% accuracy and the off mode averaged 97.4% accuracy. In an operational task of this nature, either of these modes is probably acceptable.

It is of interest to note that the previous alphanumeric studies (Decker, Dye, Lloyd, & Snyder, 1991; Lloyd et al., 1988) found off failures to be slightly more accurate than on failures. The overall accuracy was fairly high in these studies, as well. The averages of the two experiments were approximately 91% and 88% for the off and on modes, respectively. There does not appear to be a logical explanation for the slight advantage of on failures in the present search task.

Generally, as percent failure increased from 1% to 2% to 3%, response time increased, with the greatest effects at 2% and 3% and 1% failure rates producing faster responses than either 2% or 3% failure rates. These results support the concept that performance would be sensitive to lower percentages of failure for cartographic images than for alphanumeric images because of the added interference and complexity of map information. Previous studies (Abramson & Snyder, 1984; Decker, Dye, Lloyd, & Snyder, 1991; Lloyd et al., 1988) found that failure percents of three or greater significantly degraded response time and accuracy in a random search for alphanumerics and symbols. However, these studies required the subjects to search for targets on a plain background. It was anticipated that because of the more complex background inherent in the cartographic images employed in this research, effects would become evident at lower failure rates. This is the case, as effects were realized between 1% and 2% for this study, as opposed to between 2% and 3% for the previous studies.

To summarize the effects of the failure variables, there are no significant differences in performance during the no-failure condition, and there should not have been. The lack of any effects for this condition confirms the integrity of the data and of the design. It is intuitively obvious that if the individual is working with a display in good operating condition, there should be no measurable change in performance, except to the extent that the image characteristics change (e.g., because of background clutter).

Horizontal line failures on cartographic images generally impede accuracy more than vertical line or cell failures do and are particularly less accurate in the on failure mode than in the off failure mode. Also, performance tends to be less accurate as the percentage of horizontal failures on the display increases. Yet, horizontal line failures seem to have less effect on response time. It should be realized that the results of this study regarding the effects of horizontal line failures may not be generalizable to other display formats because of the specific nature of the symbols used in this research.

Vertical line failures seem to have a greater effect on the information-extraction task than on the random search task. Specifically, in the information-extraction task, vertical line failures slow response time significantly more in the on mode than in the off mode. It is possible that the vertical line failures interfere with the map information that was used as target references in the information-extraction task. Since there was no relevance of the map information in the random search task, the vertical lines would not have interfered with it.

Finally, cell failures produced slower response times when on than when off in both tasks. In previous studies (Decker, Dye, Lloyd, & Snyder, 1991; Lloyd et al., 1988), it was found that cell failures reduced accuracy and produced significantly slower response times than either type of line failure. Cell failures are not significantly injurious to performance in this study, but horizontal line failures are significant. At the failure rates used in the present study, the cell failures do not necessarily make one symbol look much like another symbol as do horizontal (and sometimes vertical) line failures. The cell failures tend to distort the outline of the symbol by either adding or subtracting elements depending on the failure mode. This distortion, while significant to some degree, does not evidently produce the same effects that line failures do when they block whole attributes of symbols.

Additional Analyses

It was desired to find the symbols that were most readily confused with other symbols. Since the eight symbols were not used an equal number of times, only the chi-square test was applicable. However, the distribution of errors did not meet the minimum requirements for the chi-square test, since more than 5% of the cells had zero values. Thus, a traditional confusion analysis could not be made.

There are differences in the response times and response accuracies among the various maps and target references. These results, however, are situationally dependent. That is, search time and accuracy vary depending on the relative complexity of the cartographic information present in the image and the specific location of that information. It is reasonable to assume, for instance, that images of topographic maps will place different demands on the operator than, for example, images of maps used for road navigation in the downtown area of a large city.

It could be argued that the data from these experiments are only appropriate for the particular maps employed in these studies; that is, the results cannot be generalized to other map images because of the undefinable effects of the maps on performance. This argument can be circumvented by using maps that have generalizable characteristics. For instance, all six maps used in this study had to show much detail because otherwise, the various types of information on the monochrome monitor could not be coded. The maps varied in the orientation of the information, but all showed approximately equal amounts of each type of detail. Thus, the maps were varied intentionally to provide for more generalizable results, but confounds resulting from the amount of information on each map were avoided. To provide greater generalization of the research results, it will be necessary to do similar research with additional types of cartographic images.

SUMMARY AND RECOMMENDATIONS

Relatively low levels of background clutter impede performance. Only two levels of clutter were studied in these experiments. Future research about cartographic images should establish definitions and limits of the amount of clutter that is tolerable. Further, the research should employ quantitative levels of clutter, as was provided in this research, rather than subjective levels. The percentage of a display's pixels that are on (comprising the information) would be one way of quantifying this variable.

Negative contrast consistently produces better performance than positive contrast for both alphanumeric and cartographic images. This conclusion and recommendation assumes that the stroke width, modulation, and luminance values are equal or nearly equal for both polarities. Different results may be evident for the two polarities if these parameters are not held constant, as some research has shown.

Cartographic images should not be used if more than 1% of the image's pixels fail; otherwise, performance will be significantly reduced. Stated otherwise, displays with more than 1% failed cells will yield reduced performance with cartographic images.

Regarding symbol design, one should avoid the use of symbols with essential markers or components that are susceptible to full obscuration by a display's failures. This recommendation is supported by Laycock (1985), who found that line failures were particularly disruptive when they aligned with major components of the characters.

As mentioned earlier, half the symbols used in this study had an essential component that was susceptible to a horizontal line failure. A study similar to the present one should be conducted using all 26 symbols used in the previous alphanumeric research (Decker, Dye, Lloyd, & Snyder, 1991; Lloyd et al., 1988). Such a study would be able to draw more complete comparisons of the findings of the alphanumeric research previously mentioned.

Finally, in this study, the luminance of the failures was equal to the luminance of the information and symbols. Thus, these results are directly applicable to one bit level displays or images on gray scale displays that only employ two intensity levels per image.

It is likely that, for gray scale displays, the failures may be of different luminances than the information. If the software program assigned different bit levels during its execution, it is possible that the failures would be displayed at different bit levels than some or all of the displayed information. If the displayed information was at a luminance sufficiently different than that of the failures, performance would be foreseeably increased by the increased ability to discriminate the information from the failures. To the authors' knowledge, this area of research has not previously been explored and may warrant consideration.

REFERENCES

Abramson, S. R., & Snyder, H. L. (1984). Operator performance on flat-panel displays with line and cell failures (Tech. Report HFL-83-3). Blacksburg, VA: Virginia Polytechnic Institute and State University.

Albert, D. E. (1975). Prediction of intelligibility of contextual dot matrix characters. Unpublished master's thesis, Virginia Polytechnic Institute and State University, Blacksburg, VA.

Bauer, D., & Cavonious, C. R. (1980). Improving the legibility of visual display units through contrast reversal. In E. Grandjean and E. Vigliani (Eds.), Ergonomic aspects of visual display terminals (pp. 137-142). London: Taylor and Francis.

Decker, J. J., Dye, C. J., Lloyd, C. J. C., & Snyder, H. L. (1991). The effects of display failures and symbol rotation on visual search and recognition performance (Technical Memorandum 4-91). Aberdeen Proving Ground, MD: U.S. Army Human Engineering Laboratory.

Decker, J. J., Kelly, P. L., Kurokawa, K., & Snyder, H. L. (1991). The effects of character size, modulation, polarity, and font on reading and search performance in matrix-addressable displays (Technical Memorandum 6-91). Aberdeen Proving Ground, MD: U.S. Army Human Engineering Laboratory.

Decker, J. J., Lloyd, C. J. C., Kurokawa, K., & Snyder, H. L. (1991) The effects of symbol rotation and matrix size on visual search performance (Technical Memorandum 8-91). Aberdeen Proving Ground, MD: U.S. Army Human Engineering Laboratory.

Erickson, R.A. (1978). Line criteria in target acquisition with television. Human Factors, 20, 573-588.

Florence, D., & Gieselman, R. E. (1986). Human performance evaluation of alternative graphic display symbologies. Perceptual and Motor Skills, 63, 399-406.

Gieselman, R. E., Landee, B. M., and Christen, F. G. (1982). Perceptual discriminability as a basis for selecting graphic symbols. Human Factors, 24, 398-407.

Greenhouse, S. W., & Geisser, S. (1959). On methods in the analysis of profile data. Psychometrika, 24, 95-112.

Human Factors Society (1988). American National Standard for human factors engineering of visual display terminal workstations (ANSI/HFS 100-1988). Santa Monica, CA.: Human Factors Society.

Knox, S. T., & Beaton, R. J. (1985). Adaptation to luminance in a complex visual field: peak versus spatial average (Technical Report 604-01). Portland, Oregon: Tektronics Imaging Research Laboratories.

Laycock, J. (1985, April). The effect of picture element failure on the legibility of a matrix display image. Displays, 70-77.

Lloyd, C. J. C., Decker, J. J., Kurokawa, K., & Snyder, H. L. (1988). The effects of line and cell failures on reading and search performance using matrix-addressable displays. Society for Information Display International Symposium Digest of Technical Papers, XIX, 344-347.

Lloyd, C. J. C., Decker, J. J., & Snyder, H. L. (1991). The effects of line and cell failures on reading and search performance using matrix-addressable displays (Technical Memorandum 7-91). Aberdeen Proving Ground, MD: U.S. Army Human Engineering Laboratory.

Pastor, J. R., & Uphaus, J. A. (1982). Significant reading errors in 7x9 dot-matrix ASCII numbers with two percent dot loss. S.I.D. International Symposium Digest of Technical Papers, XIII, 198-199.

Riley, T. M., & Barbato, G. J. (1978). Dot-matrix alphanumerics viewed under discrete element degradation. Human Factors, 20, 473-479.

Rupp, B. A. (1981). Visual display standards: A review of issues. In Proceedings of the Society for Information Display, 22, 63-72.

Semple, C. A., Heapy, R. J., Conway, E. J., & Burnett, K. T. (1971). Analysis of human factors data for electronic flight display systems (Technical Report AFFDL-TR-70-174). Wright-Patterson Air Force Base, Ohio: Flight Dynamics Laboratory.

Silbernagel, B. L. (1982). Using realistic sensor, target, and scene characteristics to develop a target acquisition model. Human Factors, 24, 321-328.

Snyder, H. L. (1980). Human visual performance and flat-panel display image quality (Technical Report HFL-80-1). Blacksburg, VA: Virginia Polytechnic Institute and State University.

Snyder, H. L. (1988). Image quality. In M. Helander (Ed.), Handbook of human-computer interaction (pp. 437-474). North Holland: Elsevier Science Publishers B.V.

Williges, R. C., & North, R. A. (1973). Prediction and cross-validation of video cartographic symbol location performance. Human Factors, 15, 321-336.

Winer, B. J. (1971). Statistical principles in experimental design. New York: McGraw-Hill.