AN EXPERIMENTAL ANALYSIS OF CRITICAL VISUAL DISPLAY PARAMETERS FOR COMPUTER-BASED TRAINING AND JOB PERFORMANCE AIDING

Rosemary Dunlop Garris
Daniel J. Dwyer
Patrick J. Moskal*
Christopher P. Mulligan*
Naval Training Systems Center
Orlando, Florida

ABSTRACT

The growing reliance on video display units to present graphic information in support of both military training and job aiding, is expected to continue. Empirical research has provided guidelines for display parameters associated with alphanumeric (textual) information, however research concerning graphics (particularly line drawings) is limited. This paper discusses the results of recent experiments which explored the effect of critical visual display parameters on task performance using line drawings as stimulus materials. The results suggested that in many cases, very low levels of graphics detail may be sufficient to produce adequate response times in locator task performance. Additionally, it is noted that, production of graphics with low levels of detail result in dramatic cost savings.

INTRODUCTION

The proliferation of training systems throughout the military services has contributed to a shifting trend, away from paper-based presentation media associated with traditional classroom lecture to computer-driven presentation formats. Training systems, as well as automated job performance aids (JPAs), rely heavily on visual presentation of information to support training and task performance. Consequently, the visual display characteristics of these systems have become a critical component in the user-system interface. Poor design, inaccurately specified visual display parameters, and/or omission of critical interface components associated with the delivery media can hinder legibility, resulting in systems which are not used or which may prove to be ineffective in providing technical information necessary for training and task performance.

Training systems often incorporate display images of training relevant equipment (e.g., via a videodisc system) with supplemental graphic line drawings, computer-generated graphic overlays, and textual information. Similarly, JPAs, which have traditionally existed in hardcopy (paper) formats and have now transitioned to computer-based modes, provide electronic delivery of technical information (schematics, illustrated parts breakdowns, etc.) through microprocessor control. This adaptation of hardcopy graphics to automated display media, makes it crucial to optimize graphics production efficiency and to determine the effects of critical visual display parameters on task performance and training effectiveness.

Past research has provided guidelines on display parameters associated with alphanumeric (textual) information. Meister (1984) provides a thorough review of this research, documenting efforts associated with display parameters such as symbol size, character fonts, symbol height-to-width ratios, and so on. However, empirically-based guidelines for graphics (especially line drawings) are lacking (Sweezey and Davis, 1983).

Recent research (Dwyer, 1985), which focused on locator task performance using a CRT display depicting printed circuit boards, revealed that a small (5" X 5") display screen resulted in acceptable locator task performance (measured by accuracy rates), but only when high discriminability existed among the components which made up the graphic. When low discriminability (a densely packed, cluttered display) existed in the graphic, a larger display screen (9" X 9" or 12" X 12") was warranted.

While this finding may not be a critical factor for school-based training systems, it is a critical issue in the development of portable, lightweight JPAs since screen size has a direct bearing on device size and weight characteristics. Thus, some of the preliminary research suggests that optimal levels of critical visual display parameters may vary as a function of intended application.

The Naval Training Systems Center (NAVTRASYSCEN) is currently engaged in a research program to identify optimum levels of critical visual display parameters within different domains (e.g, training, job aiding). The purpose of this paper is to present the results of two recent experiments in this area which explored the effects of certain display parameters on task performance.

EXPERIMENT 1

Despite the findings of Dwyer (1985) which advocated larger screen sizes for some tasks, portability remains a critical design issue in the development of automated JPAs. As a result, alternative techniques must be sought which enhance the legibility of graphics on small display screens. Experiment 1 was based upon the work of Regal and Knapp (1984), who found improvements in performance accuracy when unnecessary information was deleted from a visual display. One intent of this joint Navy-Air Force research was to remove non-critical areas of the graphics display in order to produce less clutter and facilitate task performance. The purpose of the experiment was to determine if reduced levels of graphics detail could compensate for performance decrements associated with small screen clutter. The

*This research was supported in part by Patrick J. Moskal's and Christopher P. Mulligan's appointment to the research participation program at the Naval Training Systems Center (NTSC), administered by Oak Ridge Associated Universities through an agreement between the U. S. Department of Energy and the NTSC.

experiment examined the effects of screen size (7" X 7" and 12" X 12"), image resolution (35 and 280 pixels per inch), and level of detail (high, medium, and low) on disassembly/assembly task performance. The levels of resolution were selected based on commercial availability and represented the extremes, such that any resulting performance differences would be exaggerated. High level of detail was operationally defined as all (100%) of the detail on the actual equipment used in the experiment; medium detail was defined as the component to be removed/installed, all immediate surrounding components, and the outline of the bomb ejector rack; and low detail was defined as the component to be removed/installed and the outline of the bomb ejector rack.

METHOD

SUBJECTS

Sixty Air Force maintenance training students from Lowry AFB, CO served as participants in the experiment. The students ranged in age from 17 to 31 years (mean = 20.0 years), and in tenure from 1.5 to 48 months (mean = 5.4 months).

PROCEDURE

All subjects were seated (individually) in front of a Megatek 7210 high resolution monitor which presented graphic displays of an Air Force bomb ejector rack (model MAU-12B/A). Twelve disassembly and 12 assembly frames were presented, each of which contained a graphic display of the bomb ejector rack and textual instructions explaining how to perform the task. An actual bomb ejector rack and the tools necessary to perform each disassembly/assembly task step were located on a workbench placed between the student and the monitor.

The experimental design was a 2 X 2 X 3 between subjects factorial with 5 (randomly assigned) subjects in each of the 12 cells. Performance measures were response accuracy (number of correct tasks performed), absorption time (time to read and interpret the frame), and manipulation time (time to perform the task following absorption time). Response accuracy was assessed by the experimenter and entered into the computer for storage and subsequent data analysis. Absorption and manipulation times were recorded by the computer, based upon experimenter input (i.e., a "stop" command was entered when the student had stopped reading (absorbing) and began manipulating).

RESULTS AND DISCUSSION

A multivariate analysis of variance (MANOVA) was used to analyze the results due to the multiple dependent variables assessed. The results of the analysis revealed no significant main effects nor any significant interactive effects, $p > .05$ in all cases. Although these results suggest that neither screen size, image resolution, nor level of detail significantly affected task performance, these results must be treated with extreme caution due to the small sample size (5 per cell) and the subsequent loss of statistical power to detect true differences.

EXPERIMENT 2

Independent variables such as screen size and image resolution are relatively easy to quantify because they can be defined and measured in discreet increments. Consequently, it is a straightforward process to select "levels" of these variables for examination. For example, virtually any level of screen size (5" X 5", 7" X 9", 12" X 15", etc.) can be selected for study. Similarly, image resolution can be defined and measured by the number of pixels per inch, hence we can select and study a particular level of resolution of interest (provided the display media can accommodate the desired level). However, measuring variables such as graphics detail is not so easy. It was therefore necessary to systematically develop a method for quantifying level of graphics detail, such that further experimentation could proceed using a standard metric with generalized applicability. Several techniques are available for operationally defining graphics detail. Four are presented below:

1. Method of cue presentation: Cues can be added to a graphic either (a) concentrically, (b) randomly, or (c) as a function of unique/outstanding landmarks.

2. Number of cues added: The number of cues can be increased some pre-determined number at a time, or by some percentage of the total number of cues on the actual equipment.

3. Amount of target detail: The amount of detail inherent in the target component can be varied (such as the appearance of pointers on dials, tick marks, etc.).

4. Amount of cue detail: The amount of detail inherent in the cue components can be varied (similarly to that of the target detail).

The scope of this experiment was limited to the first 2 techniques addressed above, method of cue presentation and number of cues added. (Examples of which can be found in Figures 1, 2, and 3.) In order to reduce the number of cue components which must be generated to a manageable number (such that graphics production is more efficient), a pilot study was run. Pilot data were gathered to determine the level of detail (number of cues added) at which locator task performance began to stabilize. These data identified the point at which added detail failed to produce measurable gains in locator task performance. Level of detail in paper-based line drawings of a printed circuit board and an Air Force bomb ejector rack (See Figure 4 for full detail illustrations of the two pieces of equipment.) was systematically increased by adding cue components, one-at-a-time, concentrically surrounding the target components, until subjects correctly identified the correct target component. Based upon the data collected, a range of detail levels was established for the subsequent locator tasks and was set at 1, 3, and 5 cue components for the bomb ejector rack, and 2, 6, and 10 cue components for the printed circuit board. The pilot data suggested that the overwhelming majority of students were successful in locating target components within these bounds.

Figure 1. Concentric Cue Addition

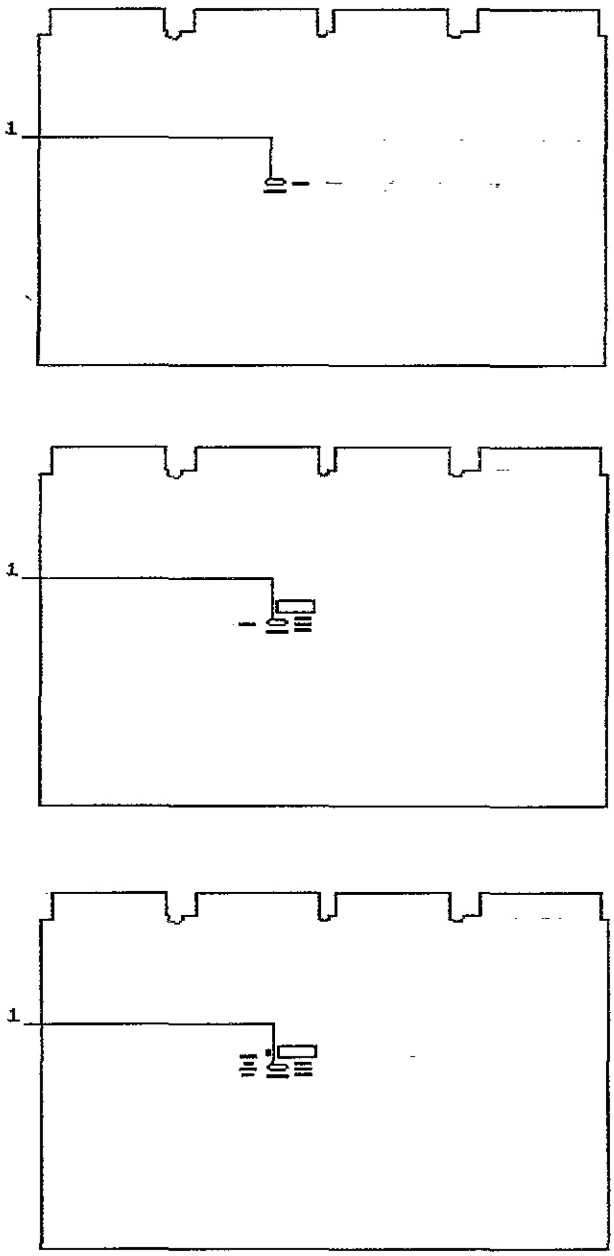


Figure 2. Random Cue Addition

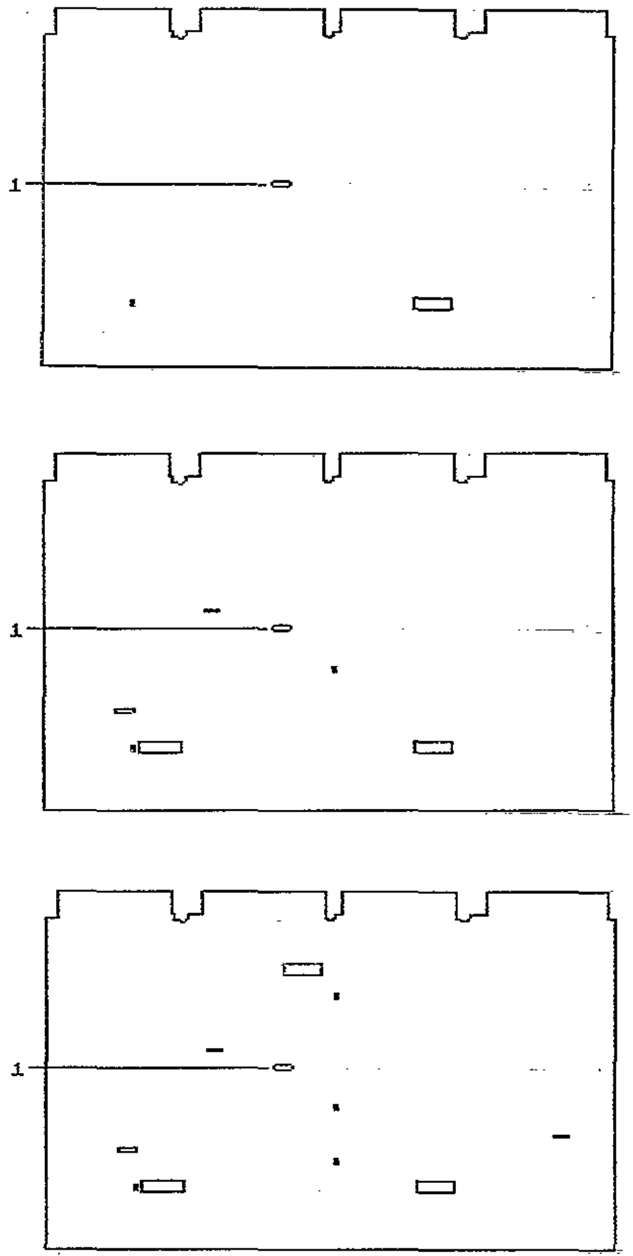


Figure 3. Landmark Cue Addition

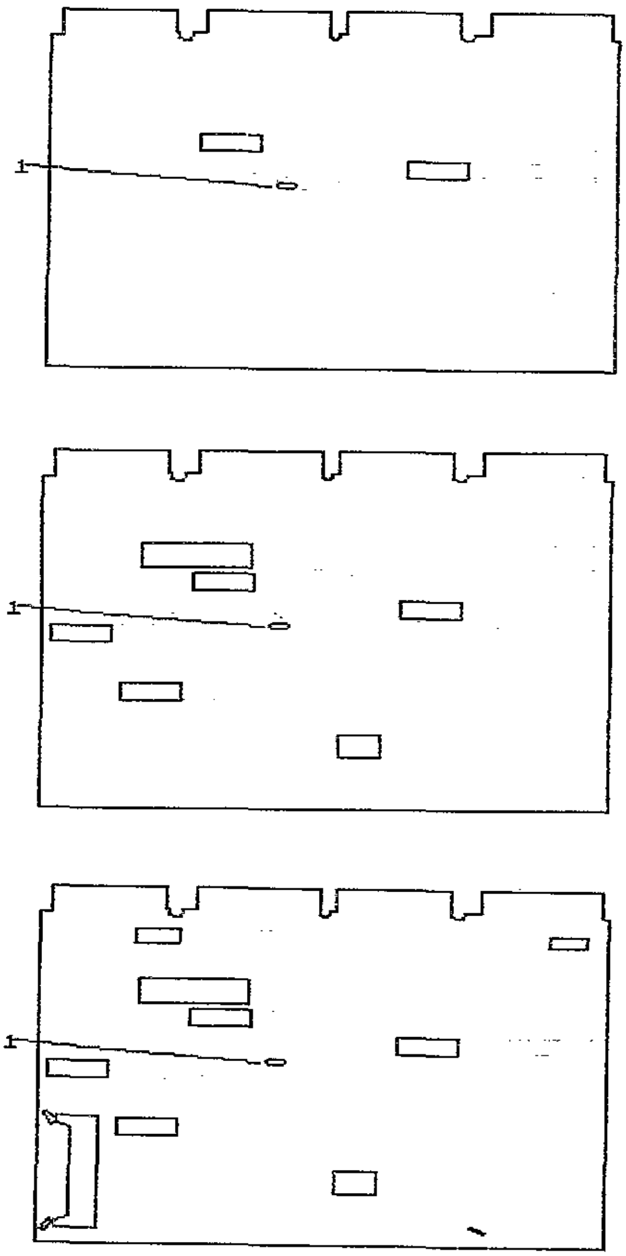
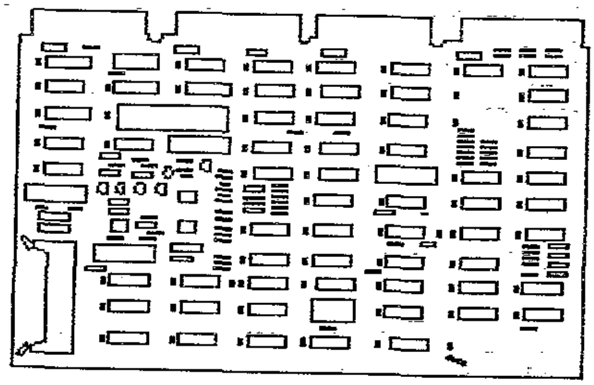
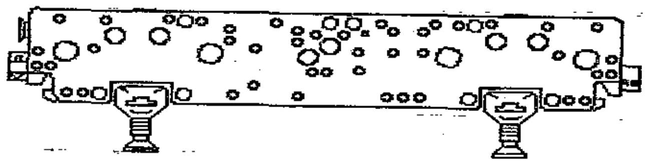


Figure 4. Circuit Board and Bomb Ejector Rack with All Components



Circuit Board



Bomb Ejector Rack

METHOD

SUBJECTS

Eighty-five male students from Torpedoman's "A" school and 25 male students from Nuclear Power "A" school participated in the experiment. All 110 students were from Service School Command, Naval Training Center, Orlando, Florida.

PROCEDURE

The three levels of detail addressed above and three methods of cue presentation (concentric, random, and landmark) were manipulated in a locator task procedure involving two equipment test beds: a printed circuit board, and an Air Force bomb ejector rack. Two additional conditions, a verbal description (of the target component and cues) and a target-only with the equipment outline, served as controls. Nine fifteen-page stimulus groups of line drawings were used for each test bed, one group corresponding to each of the experimental conditions. The same fifteen target components, albeit at different levels of the two independent variables, were depicted in the stimulus groups for the nine experimental graphics cells and in the two control conditions.

Subjects were randomly assigned to one of the 11 conditions. The order in which each subject performed the locator tasks was counterbalanced across test beds in order to control for practice effects. Additionally, the fifteen pages, each containing a different target component, were randomly presented to the student. All students performed the locator task for both pieces of equipment.

Subjects were seated at a table which held either the printed circuit board or the bomb ejector rack and the corresponding set of stimulus materials. The experimenter presented one drawing at a time and asked the student to locate and identify the target component on the actual printed circuit board/bomb ejector rack that was identified by a callout (line which pointed to the target) in the line drawing. The experimenter recorded accuracy (correct/incorrect) and task time using a stop watch. Time was measured from the point when the drawing (or verbal description) was first presented until the student correctly identified the target component or when he indicated that he could not locate it. Prior to the actual data collection, students were given two practice trials (with a target not used for actual data collection).

RESULTS

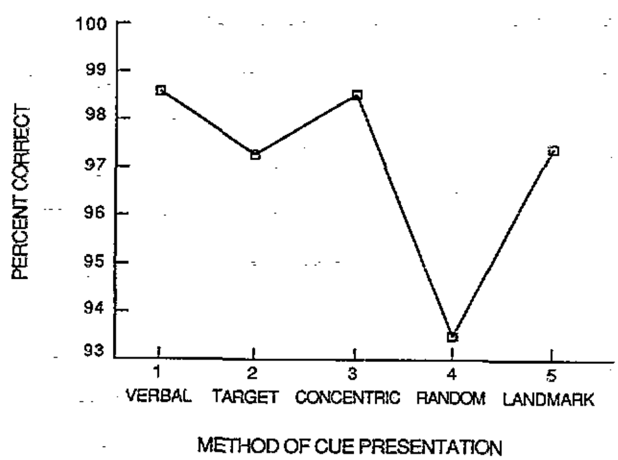
For each subject, fifteen time scores (in seconds) were recorded, as well as an accuracy score reflecting the percentage of target components which were identified correctly. A mean for each subject was then computed for the fifteen times scores. This score and the percent correct represent the two dependent variables, time and accuracy respectively. The findings of the data analysis are presented below.

ACCURACY

Printed Circuit Board. A one-way ANOVA was computed on the printed circuit board accuracy scores for method of cue presentation in order to assess the verbal only and target with outline methods with the other methods of cue presentation. This analysis yielded a significant main effect, $F=3.51(4,105)$, $p < .01$, (See Figure 5 for the graphed means). A Scheffe post hoc test identified a significantly greater level of accuracy for concentric presentation of cues over that of the random presentation of cues, ($p < .05$). None of the other pairwise comparisons were statistically significant. Results of a one-way ANOVA on level of detail indicated no significant effects in accuracy rates across the detail levels.

Next, a two-way ANOVA using accuracy scores was computed, with method of cue presentation and level of detail as the independent variables (excluding verbal and target only conditions). A main effect for method of cue presentation was found which supports the results of the one-way ANOVA which examined this variable, ($F(2,81)=5.96$, $p < .01$). Again, the findings suggest level of detail does not significantly affect task performance. The method of cue presentation was an important factor, however, with the concentric method found to produce greater accuracy than the random method.

Figure 5. Circuit Board - Accuracy All Conditions

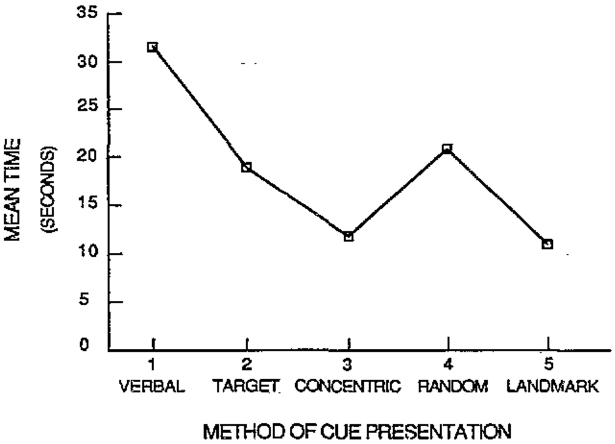


Bomb Ejector Rack. Because a significant heterogeneity of variance was present, a Kruskal-Wallis one-way ANOVA was computed on the method of cue presentation. This analysis yielded no significant effects. This was also true for a one-way ANOVA computed for level of detail. Next, a two-way ANOVA using accuracy scores was computed and this too, yielded no significant effects. Apparently, neither method of cue presentation nor level of graphic detail affects location performance accuracy on the bomb ejector rack.

TIME

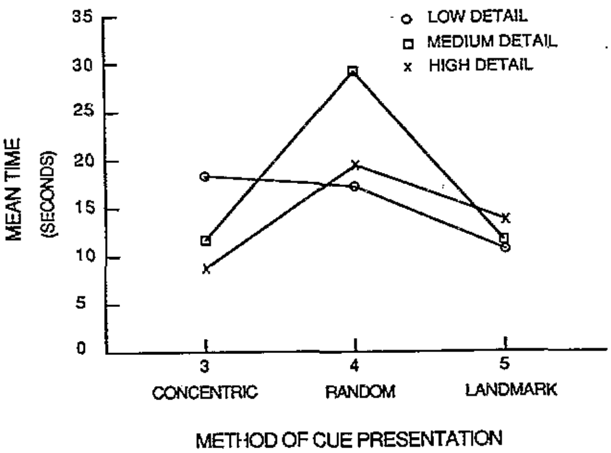
Printed Circuit Board. In order to determine if method of cue presentation affected the task time, a Kruskal-Wallis one-way ANOVA was computed for methods of cue presentation, including verbal description only and target with outline only. This analysis, which was used because there was significant heterogeneity of variance, resulted in a significant effect, $H=25.1$, $p < .001$ (see Figure 6 for the graphed means). A Scheffe post hoc test revealed that presentation of landmark cues resulted in significantly faster location times than both the random method of presenting cues and the verbal description only method ($p < .05$).

Figure 6. Circuit Board - Time
All Conditions



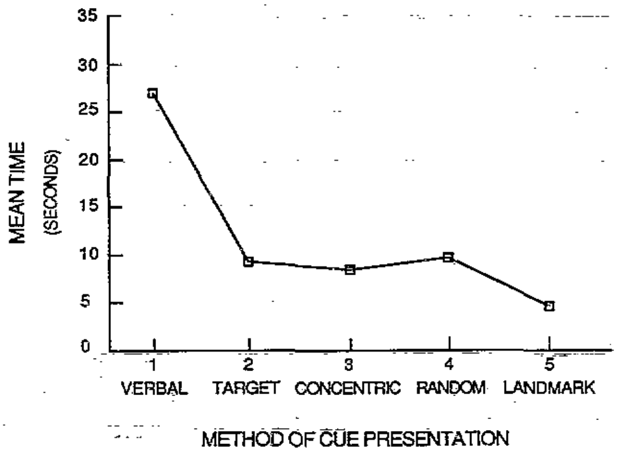
A two-way ANOVA using mean time scores for the printed circuit board was computed with method of cue presentation and level of detail as independent variables. Scores for the verbal description only and target with outline only were not included in this analysis since they were addressed previously. An interaction effect between the two independent variables was found, $F(4,81)=2.71$, $p < .05$ (see Figure 7 for the graphed means). Post hoc follow-up tests identified significant time differences between the concentric - high detail condition (mean=8.65 seconds) and the concentric - low detail condition (mean=17.91 seconds), $p < .05$, and also between the concentric - high detail condition (mean=8.65 seconds), and the random - high detail condition (mean=19.05 seconds), $p < .05$. The analysis also yielded a main effect for the method of cue presentation, $F(2,81)=8.44$, $p < .001$, confirming the results of the Kruskal-Wallis ANOVA. Finally, a one-way ANOVA performed on four levels of graphics detail (target only, two cues added, six cues added, and ten cues added) revealed no significant effects ($p > .05$) suggesting statistically equivalent location times across all levels of detail. These analyses suggest that location cues presented using the landmark and concentric methods produce faster identification of target components than either the verbal or random methods. There were no clear trends regarding the effect of level of graphics detail on task times.

Figure 7. Circuit Board - Time
Experimental Conditions Only



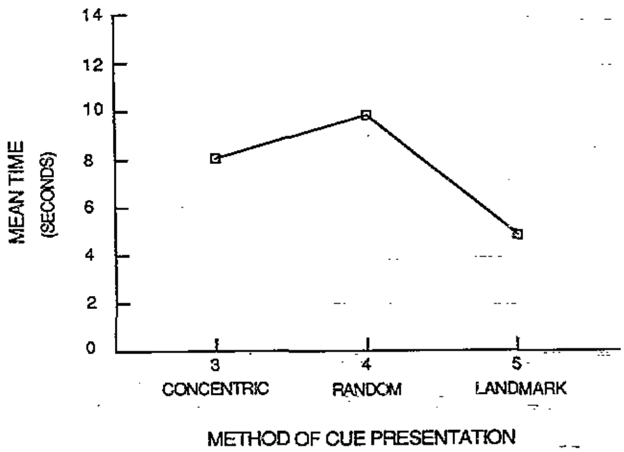
<u>Bomb Ejector Rack</u>. Because heterogeneity of variance was evident, a Kruskal-Wallis one-way ANOVA was computed for method of cue presentation. This analysis resulted in a significant main effect, $H=29.47$, $p < .001$, (see Figure 8 for the graphed means). A Scheffe post hoc test revealed that the verbal description only method resulted in significantly slower location times that that of all of the other methods of cue presentation. A one-way ANOVA for level of detail resulted in non-significant effects for location times on the bomb ejector rack test bed.

Figure 8. Bomb Ejector Rack - Time
All Conditions



A two-way ANOVA was then computed using the bomb ejector rack mean time scores, resulting in only a main effect for method of cue presentation, $F(2,81)=4.24$, $p < .05$, (see Figure 9 for the graphed means) thereby substantiating the results of the Kruskal-Wallis ANOVA. These results, for the bomb ejector rack, support the findings obtained for performance times on the printed circuit board, such that verbal description only was found to be a poorer method of supporting a locator task than graphic methods, however, varying the level of graphic detail did not significantly affect task times.

Figure 9. Bomb Ejector Rack - Time
Experimental Conditions Only



CONCLUSIONS

ACCURACY

The findings with respect to response accuracy are equivocal. The only significant effect emanating from the results of the accuracy data analysis was that the concentric method of cue presentation resulted in higher performance accuracy rates than did the random method of presentation. This finding only held for the printed circuit board test bed. The lack of any meaningful pattern in the accuracy data might be explained by examining the method used to gather performance data.

Students attempted to locate target components on the actual printed circuit board/bomb ejector rack based upon information in the line drawing. The student continued with a task until he accurately identified the correct target. Consequently, he may have incorrectly identified several components before accurately identifying the target (correct) component. However, his response was recorded as correct, regardless of his number of "misses", as long as he ultimately identified the correct target. A task was "graded" as incorrect only if he gave up.

TIME

The results of the response time data provide only one clear conclusion - the method of verbal description consistently resulted in slower identification times. This was true for both test bed applications. These results are intuitively logical - trying to identify a physical component with only verbal instructions is a difficult task.

One interesting observation was that very low levels of detail (i.e., the target with test bed outline) resulted in response times statistically equivalent to the higher levels of detail graphics. Apparently, simply providing the outline of a piece of equipment provides a sufficient amount of information to locate targets in a timely fashion. If this finding is born out in subsequent research, it could represent significant cost savings to the military in the preparation of instructional and performance aiding graphic illustrations. For example, recently prepared job performance aids for the fire control system of the M1-Tank contain 500 graphic illustrations. If the highest possible level of graphics detail (100% of the detail appearing on the actual equipment) was required to support the maintenance of this system, it would take 37.5 weeks to reproduce these graphics on a computer display, and the approximate cost would be $67,500. However, if the lowest level of detail (target with outline of equipment only) was sufficient, it would take only 2.1 weeks to produce the necessary graphics and would cost only $3,750. When this is generalized over the massive number of equipment systems operated by the military, the magnitude of possible savings becomes apparent. (See Figures 10 and 11 to see the time and cost savings, respectively.)

Clearly, additional research in this area is warranted. Performance measures associated with location accuracy should include an assessment of the number of errors made during task performance. Also, the locator tasks were performed using paper media; it is not clear if the findings observed in this experiment will generalize to a computer-based video display unit. However, the experiments which will follow the research reported here will also be examining the dimension of computer display screen size as it relates to the level of graphics detail.

Figure 10. Production Time Differences by Level of Detail

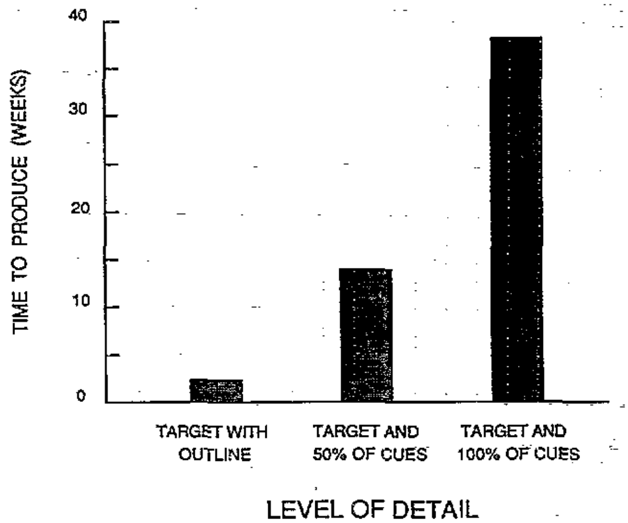
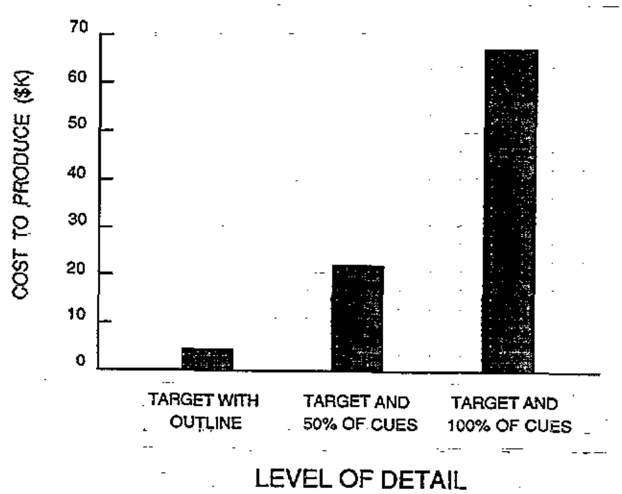


Figure 11. Production Cost Differences by Level of Detail



REFERENCES

Dwyer, D. J. (1985). An Exploratory Study of the Effect of Screen Size and Resolution on the Legibility of Graphics in Automated Job Performance Aids. NAVTRAEQUIPCEN IH-355, Naval Training Equipment Center, Orlando, FL.

Meister, D. (1984). Human Engineering Data Base for Design and Selection of Cathode Ray Tube and Other Display Systems. NAVPERSRANDCEN TR-84-51, Navy Personnel Research and Development Center, San Diego, CA.

Regal, D. M. & Knapp, B. G. (1984). An aid for improved information processing of high density computer generated visual displays, Proceedings of the Human Factors Society 28th Annual Meeting

Sweezey, R. W. & Davis, E. G. (1983). A case study of human factors guidelines in computer graphics, IEEE, Computer Graphics and Applications, pp. 21-30.

ABOUT THE AUTHORS

MS. ROSEMARY DUNLOP GARRIS is a Research Psychologist in the Human Factors Division of the Naval Training Systems Center. She has five years experience in military training research and development. Specific work experience has been in the following areas: computer-based electronic troubleshooting training; computer-based maintenance job performance aiding and training; radio instruments navigation training; chemical, biological, and radiological defense training; and visual display parameters for training and aiding applications. Ms. Garris received her B.A. from Stetson University in 1979 and her M.S. degree in Industrial/Organizational Psychology from the University of Central Florida in 1986.

MR. DANIEL J. DWYER is a Research Psychologist in the Human Factors Division of the Naval Training Systems Center. He has over eight years experience in military training R&D. Specific work experience has been in the areas of human factors front-end analysis for maintenance training, development of design guidelines for maintenance simulators, evaluation of visual display parameters for training and aiding applications, computer-assisted instruction, and embedded training. Mr. Dwyer received his B.S. degree from SUNY Cortland in 1977 and his M.S. degree in Industrial/Organizational Psychology from the University of Central Florida in 1979. Currently, he is working on his Ph.D. in Industrial/Organizational Psychology at the University of South Florida.

DR. PATRICK J. MOSKAL is a Post-doctoral Fellow appointed to the research participation program in the Human Factors Division of the Naval Training Systems Center (NTSC) administered by Oak Ridge Associated Universities through an agreement between the U.S. Department of Energy and the NTSC. His appointment is in its second year. Dr. Moskal's research has focused on performance under stress, development of training aids for chemical, biological, and radiological defense, and the development of micro-computer based part-task training. He received his M.A. and Ph.D. in Experimental Psychology from the University of Notre Dame, specializing in Sensation and Perception and Psychophysical Scaling.

MR. CHRISTOPHER P. MULLIGAN is a research assistant in the Human Factors Division of the Naval Training Systems Center. He has been with the Human Factors Division for over a year and is currently on an educational fellowship from the Department of Energy. Specific work experience has been in the following areas: evaluation of visual display parameters for computer-based training and aiding applications, computer-assisted instruction, and developing training aids for performance in stressful environments. Mr. Mulligan received his B.S. degree from Florida State University in 1986, and is currently working on his M.S. degree in Industrial/Organizational Psychology at the University of Central Florida.