HUMAN FACTORS, 1992, 34(6), 707–725

Time Stress and the Processing of Visual Displays

BRUCE G. COURY[1], *Johns Hopkins University, Laurel, Maryland, and* MARGERY D. BOULETTE, *Amherst Software, Houston, Texas*

Selecting the appropriate display format for time-constrained tasks is the focus of the research presented in this paper. The effect of time stress on operator performance was assessed by manipulating the time available to process the display. Twenty people were trained as operators and instructed to identify the state of a system using either a digital display or a polygon display. Participants were required to reach a prespecified criterion in training and were then tested under time-constrained conditions. Time constraints were set at 100%, 50%, and 25% of each person's mean unpaced response times obtained during training. Results showed that response to the time constrained conditions was significantly affected by uncertainty and the type of display format. Discussion focuses on the effects of time stress on performance and the selection of displays for time-constrained tasks.

INTRODUCTION

Time constraints are ubiquitous in virtually every job or task. This is especially true of tasks in which significant amounts of information from many different sources must be considered in a short period. Such situations are frequently encountered by air traffic controllers, military commanders, and operators of process and power generation plants.

In response to the needs of people in such situations, computer systems using powerful graphic display technology are being developed to aid decision making. Critically important to the design of those systems is selecting the appropriate display format of data and information. Selecting a display format that is least affected by the time constraints of a task has the potential to enhance performance and reduce the workload demands of a task (Hollnagel, Mancini, and Woods, 1986; Rasmussen, 1986). Selecting the appropriate display format is a major challenge. Previous research has shown that different representations (display formats) of the same system data can have a profound impact on performance (Carswell and Wickens, 1987, 1990; Sanderson, Flach, Buttigieg, and Casey, 1989), especially when the status of the decision problem or the state of a system is uncertain (Coury, Boulette, and Smith, 1989). In general, those studies have shown that an object display format is superior to a more separable display format when display attributes combine to produce useful perceptual cues. Conversely, when attention must be focused on a single element of the display

[1] Requests for reprints should be sent to Bruce G. Coury, Johns Hopkins University, Applied Physics Laboratory, BRE/Rm. 2-208, Laurel, MD 20723-6099.

or the display must be decomposed into its component parts, then a separable format produces the best performance (Carswell and Wickens, 1987, 1990; Coury et al., 1989; Sanderson et al., 1989).

Although many of the studies evaluating different display formats of system data have used paradigms that incorporate characteristics of time-constrained tasks (e.g., external pacing and deadlines), few have directly manipulated the time available for processing the display or have assessed the effect of that manipulation on performance with different display formats. The experiment reported in this paper explores the relationship between time stress and display formats by considering two important factors. First, the effect of time constraints on performance is assessed by systematically reducing the time available to process a display. Second, display formats allowing qualitatively different processing strategies are used. Although Coury et al. (1989) found strong evidence to support the argument that the polygon and digital formats used in their study were representative of integral and separable displays, there is sufficient controversy surrounding the definitions of integrality and separability to warrant a more detailed discussion of those concepts. Both issues are discussed in more detail in the following sections.

Time Stress and Performance

The effect of time constraints on performance is not a new issue. In the considerable research on the topic (e.g., see reviews by Hamilton and Warburton, 1979; Hockey, 1984; Salvendy and Smith, 1981), effects have been explained in terms of attention and theories of resource allocation (Hockey, 1984; Kahneman and Triesman, 1984; Moray, 1984; Wickens, 1980). Studies have shown that a number of factors mediate the effects of time constraints, including the structure of attentional resources (Wickens, 1980), the mapping of stimuli to responses (Hockey, 1984), and the pacing of a task (Coury and Drury, 1986). In general, one would expect that limiting the time available for processing a display would increase the workload demands of a task and result in degradation in performance when a person can no longer allocate sufficient processing resources.

Coury and Drury (1986) found that manipulating time constraints in a task does not always produce monotonically decreasing performance. By using a person's unpaced response time as a baseline for systematically reducing the time available for processing in a visual inspection task, Coury and Drury found that classification accuracy did not significantly decline in the most severely time-constrained condition. The adverse effects of stress were evident in other nonperformance measures of workload, such as heart rate variability and subjective measures of fatigue and discomfort. The authors concluded that the way in which resources are allocated and used is affected by the demands of the task and the nature of time constraints, an explanation consistent with the effects of arousal discussed by Hockey (1984) and Kahneman (1973).

Given that time stress imposes a constraint on the effective use of attentional and processing resources (and may, in fact, influence resource allocation policy), it is reasonable to expect that time stress will be mediated by the way in which system information and data are displayed. For instance, object displays (such as an integral or a configural display) can be processed faster than can separable displays in situations in which a person can exploit the perceptual cues and redundant information found in integral or configural formats (Carswell and Wickens, 1987, 1990; Pomerantz and Pristach, 1989; Sanderson et al., 1989) and relate a unique configuration of display elements to a specific response (Coury et al., 1989; Pomerantz and

Pristach, 1989). Thus the mapping of display elements to responses and the relationships among data can have a direct effect on the resource demands of a task. For instance, when the relationships among data produce a unique mapping of display elements to responses, a person can rely on the rapid processing of an object-like configuration or a set of salient features to make a quick and accurate response.

Separable displays, however, require a person to attend to individual display elements and to serially process displayed data. Because separable displays frequently require some mental manipulation of display elements, many researchers have concluded that those types of displays require more time to process (Coury et al., 1989; Goldsmith and Schvaneveldt, 1984; Pomerantz, 1986).

Integral and Separable Displays

Operationally defining a display format as integral or separable is, unfortunately, not as straightforward as one might expect. Garner's original (1974) definitions of the concepts of integrality and separability relied on a set of converging operations to define the way dimensional aspects of a stimulus are processed in a speeded classification task. Integral dimensions, as defined by Garner, are based on similarity rather than dimensional relations and are adversely affected by a filtering task that requires ignoring information on irrelevant dimensions (referred to as *filtering interference*). Integral dimensions also produce an improvement in classification performance in a correlated task involving redundant dimensions (a redundancy gain) and, in direct similarity scaling, produce a best fit of data with a euclidean distance metric. Conversely, separable dimensions are defined by classifications that are dependent on dimensional rather than similarity relations and are not adversely affected by irrelevant information in a filtering task. Separable dimensions produce no speed advantage with correlated dimensions, and they produce the best fit of data from direct similarity scaling using a city-block distance metric.

It is important to draw attention to the fact that Garner's criteria for defining integral and separable dimensions is dependent on a thorough understanding of the underlying statistical properties of a task. The correlational structure of a task, as explained by Garner (1974), determines its statistical properties by defining the mapping of task information or data to response categories. An observer's ability to make use of that mapping will be dependent, as previous research has shown, on the way in which system data are displayed. Integral displays combine correlated dimensions in such a way as to enhance the perception of dependencies among data; separable displays decouple the presentation of unrelated data, thus allowing attention to be focused on a single dimension of a task.

The display formats used by Coury et al. (1989) fit Garner's criteria defining integrality and separability. Evidence to support such a conclusion is found in the interaction between display formats and uncertainty for response times in Experiment 1 of the Coury et al. study. In that experiment uncertainty defined the mapping of system data to system state categories. When the relationships among system data produced a direct mapping to a particular system state category, uncertainty was low; when the mapping to state categories was not direct, uncertainty was high. This definition of uncertainty is consistent with other views of uncertainty found in models of inductive reasoning (Holland, Holyoak, Nisbett, and Thagard, 1986) and behavioral decision theory (Hogarth, 1987; Morgan and Henrion, 1990).

Coury et al. (1989) found that the type of display format mediates the effects of uncertainty. In that study response times to the polygon and to the digital display formats were affected by uncertainty in different ways. When uncertainty was low and the values of the process variables uniquely defined a specific system state, response times to the polygon display were significantly faster than response times to the digital display. Once uncertainty reached a point at which specific values of the process variables were no longer diagnostic (and had to be ignored), response times to the polygon display increased and became greater than those for the digital display at the same level of uncertainty. Coury et al. concluded that the polygon display is difficult to decompose under conditions of high uncertainty, thereby increasing the amount of time required to detect subtle, but critical, changes in system variables. The fact that performance was adversely affected by irrelevant information and enhanced by correlated dimensions is indicative of the filtering interference and redundancy gains used by Garner (1974) to define integral dimensions. Response times to the digital display, however, were found to be relatively unaffected by uncertainty. The digital display was not adversely affected by irrelevant information, and there was no performance advantage when dimensions were correlated, which implies that the digital values were being processed as separable dimensions.

In a more recent study Coury, Zubritzky-Weiland, and CuQlock-Knopp (1992) showed that display formats can influence the attributes used by people to identify the state of the system. Using multidimensional scaling to evaluate the composition and structure of a person's mental model, the researchers found that people using the digital display format relied on attributes related to ranges of numeric values, whereas people using the polygon display format relied on perceptual cues of shape, size, and orientation. In addition, uncertainty was found to be positively correlated with perceived distance using a euclidean metric for the polygon display but not the digital display. Such results are consistent with Garner's (1974) definitions of similarity and dimensional structure and the criteria for defining integrality and separability. Judgments of similarity with the polygon display were based on similarity relations, whereas the same judgments with the digital display were based on dimensional relations.

Configurality Effects and Emergent Features

Not all researchers have been successful at applying Garner's criteria to different display formats. Carswell and Wickens (1990) systematically assessed the criteria defining configural and separable dimensions as well as the concept of emergent features proposed by Pomerantz (1986; Pomerantz and Pristach, 1989). Emergent features are the perceptual cues that arise from the interaction among component features in a display. The interaction of such features produces configural superiority effects "in which line segments are perceived or discriminated better when in the presence of other line segments that are nominally irrelevant to the task than when they are presented in isolation" (Pomerantz and Pristach, 1989, p. 636). Evidence for emergent features is found in speeded classification tasks when the following results occur: filtering interference (arising from a failure in selective attention) in tasks in which one of the dimensions of the stimulus varies randomly and must be ignored; condensation efficiency (attributable to divided attention) in tasks in which dimensions interact to produce configural cues; and no redundancy gain in tasks involving correlated dimensions.

In the study by Carswell and Wickens (1990), no specific evidence was found to support redundancy gains in a correlated task. The researchers did find evidence supporting the existence of emergent features, which led them to conclude that configurality effects may provide a better explanation than integrality for the superior performance found in some of the display formats. There is reason to believe, however, that the absence of integrality in the Carswell and Wickens experiments may have been attributable to three other factors. First, the task was relatively straightforward, requiring judgments along two dimensions that produced response times in the range of 300–400 ms. Those results are consistent with the fastest response times obtained by Pomerantz and Pristach (1989) using relatively simple stimuli composed of three line segments. Pomerantz and Pristach concluded that the fastest response times obtained in their experiments were most likely attributable to higher acuity for stimuli presented in the foveal region. Thus it is possible that the conditions used in the Carswell and Wickens experiments may not have allowed redundancy gains to occur.

Second, emergent features do not always emerge. The study by Pomerantz and Pristach assessed four different types of displays containing emergent features in a variety of conditions. Not all of the displays produced the expected configural superiority effects; the arrangement and saliency of features and the requirements of the task significantly influenced the utility of the emergent feature. The results led Pomerantz and Pristach to conclude that configural superiority effects will be found only when display elements interact to produce emergent features *and* when the person viewing the display can exploit those perceptual cues to make a response. This is consistent with the conclusion of Coury et al. (1989) that uncertainty (i.e., the mapping of system data to response categories) is a critical factor in determining the relative superiority of a display format in a multidimensional decision making task. Carswell and Wickens (1990) also acknowledged such a possibility; they pointed out that configurality effects may be an intermediate type of dimensional interaction lying somewhere between true integrality and true separability.

Finally, the similarity and dimensional structure of the stimuli used by Carswell and Wickens have not been determined. The use of scaling methods is critical to Garner's (1974) definitions of integral and separable dimensions and is a fundamental component of his set of converging operations. Without using similarity rating techniques (as in Coury et al., 1992), one cannot completely ascertain whether a specific display format possesses either integral or separable properties. In fact, Pomerantz and Pristach (1989) pointed out that the parts of forms (such as those found in configural display formats) have not yet been shown to have integral properties. Garner (1978) and Triesman (1986) also discussed the use of multidimensional scaling to identify integral and separable dimensions, but not configural dimensions. Thus stimuli that have clearly definable configural properties may not produce the patterns of results that can be attributed to integral dimensions.

Factors Affecting the Processing of Displays

Research has shown that two primary factors determine the integrality or separability of a display format: the correlational structure of the task and the integrality or separability of the dimensional aspects of the display. The original work by Garner and Pomerantz established the criteria for identifying integral dimensions, separable dimensions, and emergent features. Further research by Pomerantz and Pristach (1989), Coury et al. (1989), and Sanderson et al. (1989) has demonstrated the importance of task factors in determining the extent of integral processing and the utility of emergent features. Those studies have shown that the underlying statistical properties of the task—and the relationship between the elements of a display and the demands of a task—significantly affect performance and determine the relative merits of specific display formats.

To ensure that both factors were taken into consideration in this experiment, the display formats fulfilled two requirements. First, the displays were characterized in terms of their integral or separable properties. As previously discussed, the research by Coury et al. (1989, 1992) demonstrated the integral properties of the polygon display format and the separable properties of the digital display format.

Second, the cognitive requirements of the task allowed the full range of analyzability, from true integrality to true separability, to be considered. Coury et al. (1989) showed that manipulating uncertainty within a response category produced conditions that allow filtering interference, redundancy gains, and condensation efficiency to occur. In other studies evidence for integrality, separability, and configurality effects is usually obtained by comparing performance across qualitatively different tasks and classification schemes. The task used in this experiment minimizes the potential for negative transfer by using the same classification scheme and response categories in the same task. Uncertainty provides the only variation in the task by changing the mapping of system data to system state categories, thereby allowing the full range of processing strategies to occur for a single response and display format.

The task also allows consideration of the way in which irrelevant information in the display affects the interaction between uncertainty and display formats. Given that people who work with complex systems are typically faced with processing data that is not diagnostic, manipulating the relevancy of displayed information must be a critical component of any study attempting to generalize results to actual tasks. In addition, the task must simulate the processing demands associated with identifying the state of the system. This task meets that requirement by using a combination of values of system variables to identify a system state and ensures that the mapping of values of system variables to a state category is not direct.

The two types of display formats used in this experiment are shown in Figure 1. The digital display (Figure 1a) presents system variables as independent numerical values

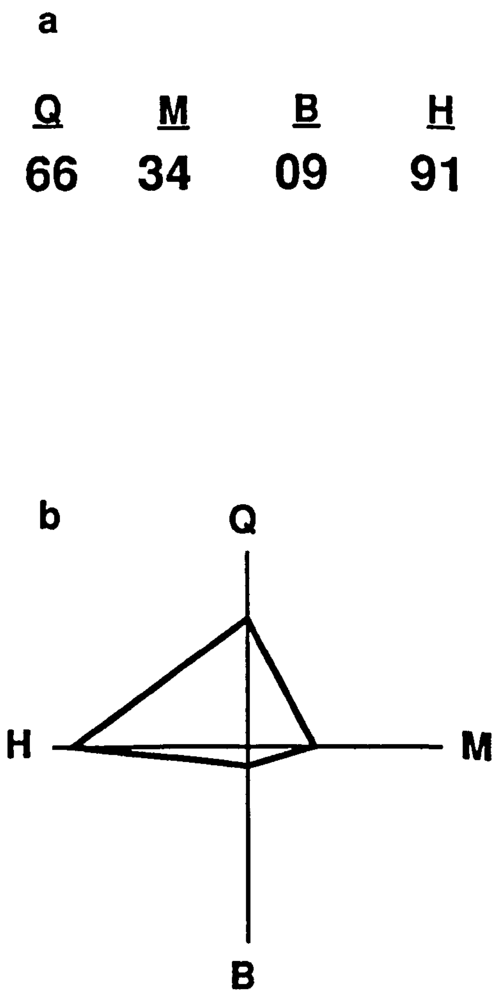

Figure 1. *The two types of displays used in the experiment: (a) an instance of system data presented as a digital display; (b) an instance of system data presented as a polygon display. In all cases the values of the system variables are the same.*

(i.e., separate digits). The digital format is a separable representation that maximizes the separation between process variables, places emphasis on verbal coding of the data, and minimizes the opportunity for integral or configural properties to emerge.

The polygon display (Figure 1b) presents the same system data in integral form. This display maps the values of system variables defining a system state onto a single object-like configuration (a polygon). The display allows a person to classify system state by attending to the overall shape or configuration of the display without attending to any specific value of a process variable.

The various ways in which the digital and polygon display formats are processed can be used to predict the interaction between time stress and display formats. In general one would expect overall performance to be adversely affected by increasing time stress. The adverse effects of time constraints can be ameliorated, however, by a display format that allows a person to process display elements rapidly and identify the state of a system accurately. Conversely, a display requiring more time to process will be more adversely affected by time stress. Thus a separable display requiring a serial processing strategy will require more time to process than will an integral display that allows a more global processing strategy or reliance on emergent features. In this experiment the relative merits of display formats will become evident when uncertainty is considered.

METHOD AND PROCEDURES

Subjects

Twenty subjects ranging in age from 18 to 28 years drawn from the University of Massachusetts student population participated in this experiment. Each was paid $5 per hour and had reached a prespecified criterion of classification performance during the first day of training.

Stimuli and System State Categories

The experimental stimulus set was composed of instances of system data uniformly distributed across the range of uncertainty within four system state categories. In technical terms, the construction of instances of system data relies on a discrete multivariate distribution that uses a vector map to relate values of the system variables to system state categories (Dougherty, 1990, pp. 200–210).

The process variables represent the four dimensions of the classification task, with the ranges of values along each dimension defining the value of that dimensional attribute for each of the four system state categories. The range of values defining each of the four system states is presented in Table 1 (Q, M, B, and H are the labels for the four system variables); these are the same definitions of system state used in Coury et al. (1989). Notice that all possible combinations of the ranges of values for the four dimensions establish the dimensional structure for this task as defined by Garner. A combination of values of Q, M, B, and H (e.g., $Q = 30, M = 65, B = 10, H = 85$) defines a system state category (in this example, System State 1), with the mapping of system data to system state categories defining the correlational structure (Garner, 1974) of this task. Thus a person's ability to classify an instance of system state accurately

TABLE 1

Ranges of Values for Process Variables

System State	System Variable			
	Q	*M*	*B*	*H*
1	25–51	49–75	0–26	74–100
2	25–51	49–75	74–100	0–26
3	49–75	25–51	0–26	74–100
4	49–75	25–51	74–100	0–26

is dependent on consideration of the joint probability of the values of the four system variables and their relation to a system state category.

Notice that Q and M play a role that is distinctly different from that of B and H in the categorization scheme. When both Q and M are in the range 49–51, an overlap between state categories is produced that simultaneously defines all four system states; only by considering the values of B and H can the number of system state alternatives be reduced to two (e.g., if $Q = 49$ and $M = 51$ in the previous example, then that set of values defines both System State 1 and 3). B and H, however, do not have a similar overlap region. The overlap region represents a unique condition in this experiment, and we refer to it as the *borderline.* When a combination of values falls in the borderline, the best a person can do is reduce the number of system state categories from four alternatives to two. Thus either state of a category pair (States 1 and 3 or States 2 and 4) is a correct response, and a person could be correct 50% of the time even if he or she were guessing.

The type of multidimensional classification task used in this experiment is most characteristic of a markovian decision process with stationary transition probabilities. The process of systematically reducing the number of system states is equivalent to eliminating certain alternative states from consideration in a markov process and terminating the process with a stationary stopping rule (DeGroot, 1970; Winston, 1987). Note too that the system states act as absorbing states, which terminate the process once the correct system state has been identified (DeGroot, 1970). This type of problem is also similar in character to the typical probabilistic dynamic programming problem found in operations research (Winston, 1987).

By systematically varying the extent to which values of the variables define membership in a state category, a full range of uncertainty was produced, and the diagnosticity of specific system variables was manipulated. Combinations of values of system variables that uniquely define a state category have high diagnostic value and are low in uncertainty; conversely, data that do not uniquely define a state category have low diagnostic value and are high in uncertainty. Garner (1974) described this type of correlational structure as a four-term interaction uncertainty.

Experimental Task and Procedure

The task and procedures used in this experiment were based on the methodology employed in Coury et al. (1989) and Coury and Drury (1986). In this experiment participants acted as operators whose task was to identify the state of a system. They were required to identify system state by classifying an instance of system data into one of the four state categories described in the previous section.

The experiment was divided into five sessions across two days: two training sessions on Day 1 and three time-constrained conditions (100%, 50%, and 25% of mean unpaced time) on Day 2. Participants were randomly assigned to either the digital display group or the polygon display group. They were seated before two 33-cm monochrome monitors, each driven by a DEC Pro380 microcomputer. The monitor positioned directly in front of the person was used to present the instances of system data for classification. The second monitor, positioned to the right of the presentation monitor, provided feedback to the participant during training sessions. Without describing the underlying classification scheme, the experimenter reviewed the procedures, demonstrated the task, and explained the information provided by the feedback monitor with each person. The feedback monitor provided information regarding the accuracy of a response, the correct system state, and variable-specific information about state category membership. At no time, however, did the participant receive specific information (such as in Table 1) about the possible range of values for each state category.

Training. During the two training sessions on Day 1, the operators classified 384 instances of system data. Presentation of each instance constituted a training trial. Each training trial followed the same pattern: presentation of an instance of system data, the person's response, and feedback on the accuracy of the response. The classification task in the training sessions was self-paced. Only those people who reached a 90% classification accuracy criterion in the last 100 trials of training were allowed to continue to Day 2. This criterion ensured that the results obtained in the time-constrained conditions would not be confounded by inadequate learning of the concepts defining state category membership.

Time constrained conditions. On Day 2 operators classified the instances of system data under conditions of time constraints with no feedback. The three time-constrained conditions were based on each person's mean correct response times during the last 100 trials of training. Presentation intervals for the time-constrained conditions were based on individual performance to ensure that task demands were based on each person's ability. The three display duration times (representing the three time-constrained conditions) were 100%, 50%, and 25% of mean unpaced response time. Selection of the duration times was based on the results from the Coury and Drury (1986) experiment. In that study the 100% and 50% time-constrained conditions did not adversely affect performance; consequently, this experiment included the 25% condition in order to create a condition with severe time constraints.

Each person was tested under each time-constrained condition for 256 trials. These trials were composed of the same stimulus set used in the first session of training but presented in a different random order. Each trial within one of the time-constrained conditions followed the same procedure: the participant initiated the trial by pressing the spacebar of the keyboard; an instance of system data was presented for the predetermined display duration time for that block of 256 trials; and a state category was selected by pressing the appropriate response key during the time the stimulus was presented or in a 500 ms buffer period following the offset of the stimulus.

Data Measurement

Throughout the experiment response times and accuracy data for each person were recorded. Response times were measured as the interval between the onset of an instance of system data and the person's response. Only times for correct responses were used in the analysis. Accuracy was defined as the proportion of instances of system data correctly classified. During the training session both response time and percentage correct were summarized by averaging consecutive sequences of 32 trials. This produced 12 sequences of 32 trials on Day 1.

To evaluate the effect of uncertainty, an analysis was performed on accuracy and response times for instances of system data occupying predetermined incremental distances from the borderline condition. Instances were selected for analyses that were at the borderline and six steps away from the borderline, with the six steps representing equal step sizes across the range of category membership. The borderline and six steps operationally define *distance from borderline* and the levels of uncertainty in the experiment. Specifically, in the borderline condition (as described in a previous section) both variables Q and M had values that were in the overlap region of 49–51, and a response to either state in the pair would be a correct response. In Steps 1 and 2 at least one of the variables Q or M had a value in the overlap region. In Steps 3 through 6 all four variables took on values that were more distant in the range of possibilities from the overlap region (e.g., an instance defined by $Q = 25$, $M = 75$, $B = 0$, and $H = 100$ would be the most distant from the overlap region and uniquely defines System State 1). Because guessing could result in a higher level of accuracy at the borderline relative to Steps 1–6, the accuracy data used in the analysis of uncertainty were corrected for chance performance (Green and Swets, 1966).

Experimental Design and Analysis

The experiment used a multifactorial repeated-measures design with participants grouped under display type. Display type (polygon vs. digital) was the between-subjects variable, with block of trials the within-subject variable for the training session and time constraints and distance from the borderline the within-subject variables for the time-constrained sessions. Each of the dependent measures for each session was evaluated using a multifactorial ANOVA to assess the effects of display type, time constraints, and distance from borderline. All factors except participants were treated as fixed. In the time-constrained sessions, the order of presentation of the three time-constrained conditions was counterbalanced across subjects.

RESULTS

The results will be presented separately for the training sessions and for the time-constrained conditions.

Training Sessions

Accuracy. Both the digital display group and the polygon display group exceeded the classification accuracy criterion by the end of training. Classification accuracy for the digital display group in the second session was equal to 98%; accuracy for the polygon display group in the same session was equal to 93%. The ANOVA of those data showed no significant difference in accuracy between the two groups. In addition, there was no Display Type × Training Trials interaction, implying that the learning curves for the two groups were statistically equivalent. Mean classification accuracy for both displays is presented in Figure 2.

Response times. The greatest change in performance during training occurred in response times. Response times significantly decreased for both groups during the two sessions of training on Day 1, $F(1,18) = 80.70$, $p < 0.001$. The polygon group's response times decreased from a mean of 3534 ms in the first session of training to 1801 ms in the second session. The digital group's response times decreased from 6141 ms to 3628 ms. Because there was no significant Display Type × Training Trials interaction, the rate of decrease across trials for the two groups was approximately the same. Mean response

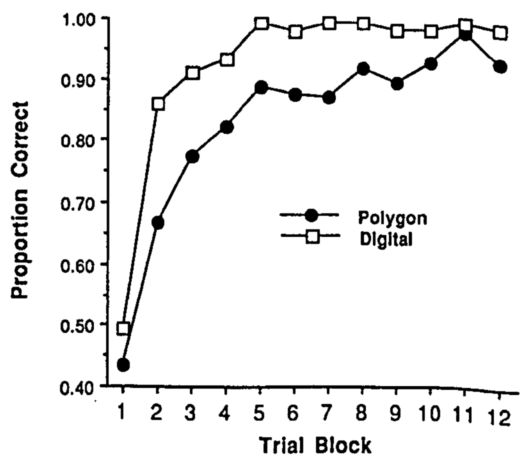

Figure 2. *Mean proportion correct responses for the digital and polygon displays as a function of blocks of trials during training sessions.*

times for the two display groups are presented in Figure 3.

Time-Constrained Conditions

Accuracy. As expected, overall classification accuracy decreased as time stress increased and was significantly affected by the time-constrained conditions, $F(2,36) = 24.42$, $p < 0.001$. Mean classification accuracy for each display type for the three time-constrained conditions is presented in Table 2. The mean unpaced time used to establish the presentation interval in the time-constrained conditions is also given for comparison.

Accuracy was highest during the 100% and 50% time-constrained conditions for both display types; the greatest decline in accuracy occurred during the 25% time-constrained condition. Response to time constraints was not the same for both types of displays. The polygon display group was able to maintain higher classification accuracy during the 25% condition than was the digital display group. In addition, the decrement in performance from the 50% condition to the 25% condition was only 7% for the polygon display but 30%

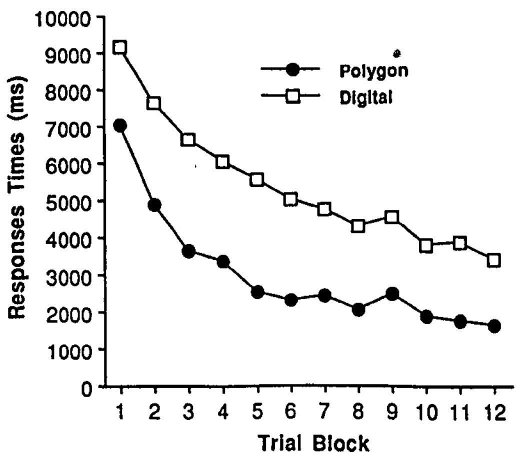

Figure 3. *Mean correct response times (in ms) for the digital and polygon displays as a function of blocks of trials during training sessions.*

for the digital display. This difference in response to time stress produced a significant Display Type × Time Constraints interaction, $F(2,36) = 3.56$, $p < 0.04$. It is interesting to note that the decrement in accuracy from the 100% to 50% condition for both display types (2% for the digital display and 1% for the polygon display) was small despite a significant reduction in stimulus duration.

When uncertainty (as defined by distance from borderline) is taken into consideration, the source of the difference in accuracy for the two types of display formats is revealed. Mean classification accuracy as a function of distance from borderline for each time-constrained condition is presented in Figure 4a for the digital display and in Figure 4b for the polygon display. Classification accuracy was significantly affected by uncertainty, $F(6,108) = 32.07$, $p < 0.001$; in general, people were least accurate when uncertainty was high; classification accuracy increased as uncertainty decreased.

The effect of uncertainty was not the same for the two types of displays. Accuracy with the digital display remained high across all levels of uncertainty until the most severely time-constrained condition occurred, at which point accuracy suffered a significant and almost constant decline across all distances from the borderline. The polygon display group, however, exhibited a different pattern of results: accuracy declined from the unpaced levels for all time-constrained conditions. Although the decline was evident at all levels of uncertainty, the decrease in accuracy was most pronounced under conditions of high uncertainty. The ANOVA of these data revealed a significant Display Type × Distance from Borderline interaction, $F(12,216) = 2.42$, $p = 0.006$. Analysis of the simple main effects of distance from borderline for each of the display types found the effect of uncertainty to be highly significant for the polygon display for each

TABLE 2

Effects of Time Constraints on Classification Accuracy and Response Times

Measure/Type of Display	*Unpaced*	*Time-Constrained Condition* 100%	50%	25%
Classification accuracy				
Digital	0.98	0.90	0.88	0.58
Polygon	0.93	0.81	0.80	0.73
Response time (ms)				
Digital	3658	2178	1499	862
Polygon	1765	1076	744	638
Presentation interval				
Digital	—	3658	1829	915
Polygon	—	1765	883	599

time-constrained condition: 100%, $F(6,54) = 8.48$, $p < 0.001$; 50%, $F(6,54) = 19.64$, $p < 0.001$; and 25%, $F(6,54) = 11.57$, $p < 0.001$. However, this effect was significant only in the 25% condition, $F(6,54) = 10.45$, $p < 0.001$, for the digital display.

Errors. The decrease in accuracy under conditions of uncertainty and increasing time stress can be attributed to two types of error: selecting the incorrect state category, or failing to respond within the specified response window. Both types of errors are indicative of time-constrained tasks and represent different failures in a person's ability to process system data.

Selecting the incorrect state category (a wrong response) would be an error attributable to incomplete or inaccurate processing of system data. Failing to respond within a given interval (a missed response) would be indicative of insufficient processing time. The proportion of missed responses for each type of display at each level of uncertainty in the time-constrained conditions are presented in Figure 5, and the proportion of wrong responses for each type of display at each level of uncertainty in the time-constrained conditions is presented in Figure 6. ANOVAs of the proportion of missed and wrong responses across levels of uncertainty for each time-constrained condition were conducted separately for the digital and polygon display.

The results show that the proportion of errors associated with wrong and missed responses increased as time constraints became more severe, with the greatest increase occurring for the missed responses. Time constraints significantly affected the proportion of missed responses to both the polygon display, $F(2,18) = 13.85$, $p < 0.001$, and the digital display, $F(2,18) = 6.10$, $p = 0.009$. Although uncertainty was found to be a significant factor for missed responses, the effect was not the same for each type of display. Analysis of the simple main effect of distance from borderline at each of the time-constrained conditions for the polygon display confirmed that the effect of uncertainty was significant at all levels of time stress: 100%, $F(6,54) = 4.01$, $p = 0.002$; 50%, $F(6,54) = 5.43$, $p < 0.001$; and 25%, $F(6,54) = 2.46$, $p < 0.04$. The same analysis for the digital display found distance from borderline to significantly affect the proportion of missed responses only at the 25% time-constrained condition, $F(6,54) = 3.10$, $p = 0.011$. Analyses of the simple main effects of uncertainty at the 100% condition, $F(6,54) = 0.825$, and the 50% condition, $F(6,54) = 0.756$, showed no significance.

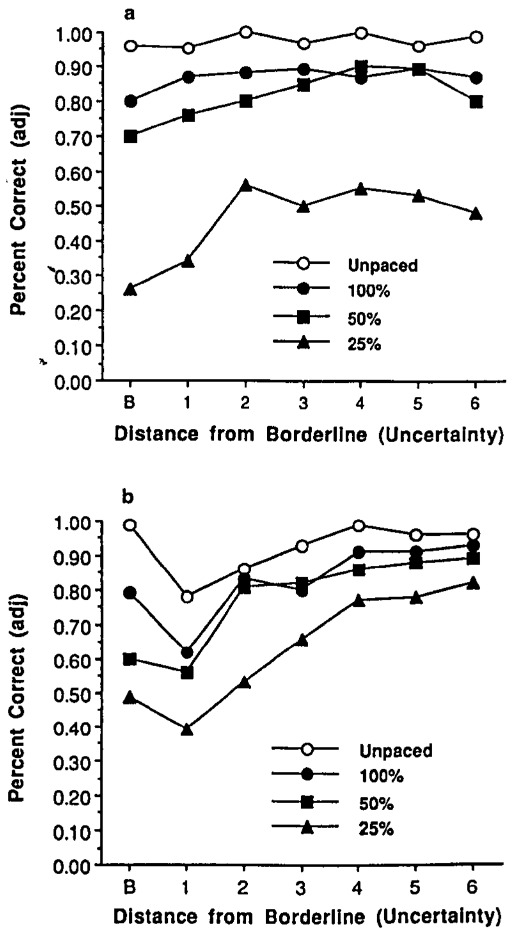

Figure 4. *Mean proportion correct responses for the digital (a) and the polygon (b) display groups as a function of distance from borderline for the unpaced and three time-constrained conditions. B is the borderline condition at which uncertainty is highest; numbers 1 through 6 indicate increasing distance from the borderline, with Step Size 6 representing an instance with the lowest uncertainty. Data have been corrected for chance performance.*

The increase in the proportion of wrong responses as the time available for processing decreased was not the same for both types of displays. The primary source of the difference between the two displays was uncertainty. Wrong responses to the polygon display were affected by distance from borderline, $F(6,54) = 28.62$, $p < 0.001$, but a significant Time Constraints × Distance from Borderline interaction, $F(12,108) = 1.93$, $p = 0.04$, indicates that the effect was different for each level of time stress. Analysis of the simple main effects of uncertainty for each time-constrained condition with the polygon display was found to be significant for wrong responses: 100%, $F(6,54) = 16.19$, $p < 0.001$; 50%, $F(6,54) = 13.77$, $p < 0.001$; and 25%, $F(6,54) = 10.67$, $p < 0.001$.

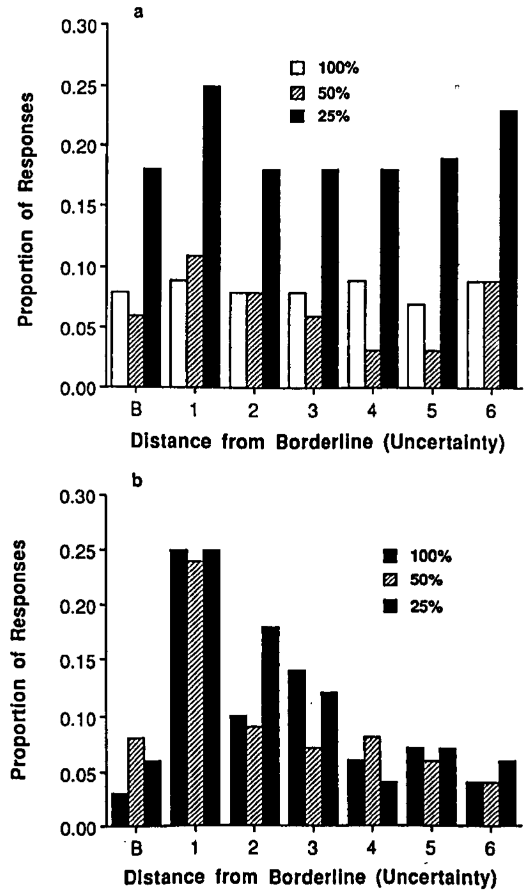

Figure 5. *Proportion of missed (a) and wrong (b) responses for the digital display in the 100%, 50%, and 25% time-constrained conditions. Distance from borderline represents the range of uncertainty from the highest level at B to the lowest level at Step Size 6.*

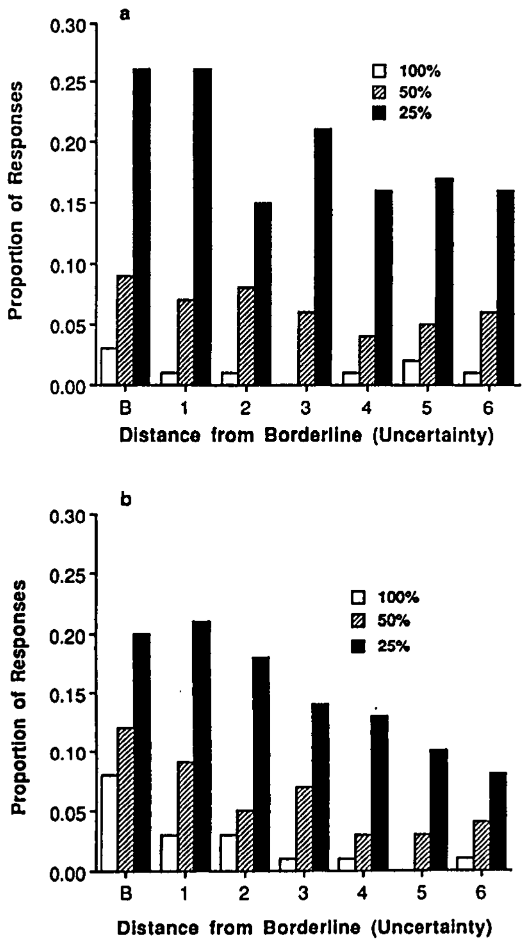

Figure 6. *Proportion of missed (a) and wrong (b) responses for the polygon display in the 100%, 50%, and 25% time-constrained conditions. Distance from borderline represents the range of uncertainty from the highest level at B to the lowest level at Step Size 6.*

Although wrong responses to the digital display were also found to be significantly affected by time constraints, $F(2,18) = 5.57$, $p = 0.04$, and distance from borderline, $F(6,54) = 2.54$, $p = 0.03$, the analysis of the simple main effects revealed that uncertainty was not a factor at each level of the time-constrained conditions: 100%, $F(6,54) = 0.46$; 50%, $F(6,54) = 1.86$, $p = 0.11$; and 25%, $F(6,54) = 1.09$, $p = 0.38$. These results indicate that the effect of uncertainty was a tenuous one for wrong responses to the digital display.

Response Times. The effect of uncertainty on response times for each of the time constraints is presented for the digital display in Figure 7a and for the polygon display in Figure 7b. In general response times for the polygon display group were faster than for the digital display group (1076 ms vs. 2128 ms),

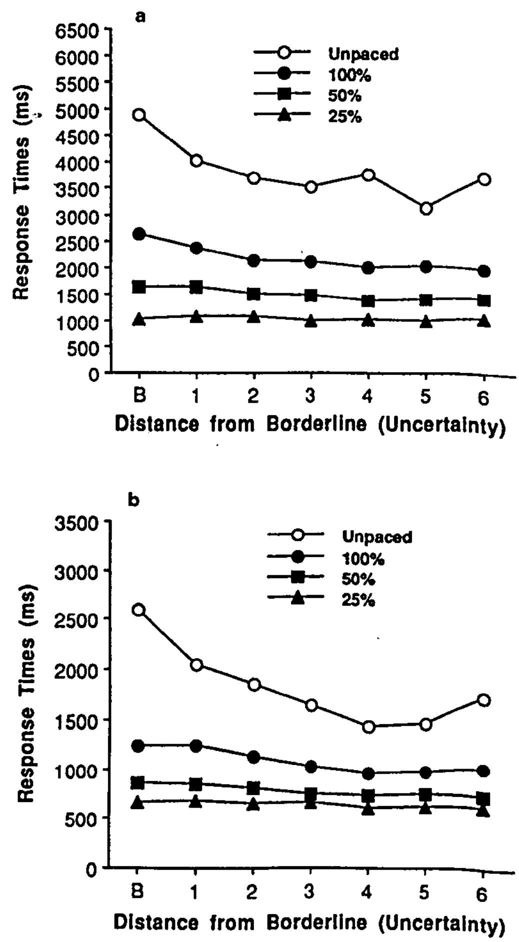

Figure 7. *Mean correct response times for the digital (a) and the polygon (b) display groups as a function of distance from borderline for the unpaced and three time-constrained conditions. Distance from borderline represents the range of uncertainty from the highest level at B to the lowest level at Step Size 6.*

and that difference was evident at all levels of time-constrained conditions. The ANOVA of these data revealed a significant main effect of display type, $F(1,18) = 25.80$, $p < 0.001$, and a significant Display Type × Time Constraints interaction, $F(3,54) = 15.43$, $p < 0.001$.

Uncertainty also significantly affected response times, $F(6,108) = 9.43$, $p < 0.001$, but not in the same way for each of the time-constrained conditions. The effect of uncertainty found in previous research (slowest response times when uncertainty was high and fastest response times when uncertainty was low) was found in this experiment only in the unpaced condition. Once time constraints were introduced, uncertainty was no longer a factor, and response times remained constant across distance from borderline for all three time-constrained conditions. The differential effect of the time-constrained conditions produced a significant Time Constraints × Distance from Borderline interaction, $F(18,324) = 4.72$, $p < 0.001$. No other main effects or interactions were significant.

Response times did not directly correspond to the time the stimulus was actually available for processing (i.e., the stimulus presentation interval). When the mean unpaced response times obtained during training were compared with the response times obtained during the time-constrained conditions (see Table 2), an interesting result appeared: in general, people did not use all of the time available to them and responded before the offset of the stimulus. Only the people using the polygon display in the 25% condition required more time than that provided by the presentation interval.

DISCUSSION AND CONCLUSION

The interaction between time stress and uncertainty determined performance in this experiment. Time stress significantly altered the way in which uncertainty affected performance, and the interaction between time stress and uncertainty determined the extent of integral and separable processing of display elements.

In previous research Coury et al. (1989) used the interaction between display formats and uncertainty for response times to characterize the differences between integral and separable displays. In this experiment that interaction was found only in the unpaced condition. There is a straightforward explanation for the difference in results. In previous studies in which time constraints were not a factor, a person could contend with uncertainty by adjusting the time to process the display; as uncertainty increased, the time taken to process the display also increased. In this experiment time constraints prohibited people from spending more time contemplating the status of the system and forced them to process the display elements rapidly before the response deadline occurred. Consequently response times were no longer sensitive to the effects of uncertainty.

The expected Display Format × Uncertainty interaction found in previous research emerged in this experiment in the accuracy data. The pattern is consistent: people using the digital display were unaffected by uncertainty until the most severe time-constrained condition. People using the polygon display were affected by uncertainty, with the best classification accuracy occurring for instances low in uncertainty and increasing errors with increasing uncertainty. Note that this interaction between display types and uncertainty in the accuracy data is the complement to the one found for response times by Coury et al. (1989). When the time available for processing becomes fixed and a person no longer has sufficient time to adequately process the elements in the display, accuracy declines. In the polygon display used in this experiment, the decline occurred under conditions of high uncertainty; in the digital display the decline was constant across all levels of uncertainty. Both displays produced the greatest number of errors when time stress was the most severe.

The decline in accuracy at the borderline condition illustrates clearly the adverse effects of time stress. Even in a situation in which guessing could produce a correct response, accuracy declined with increasing time constraints. The effect was especially evident with the polygon display; adjusting for guessing, performance at the borderline was worse than at any other level of uncertainty.

In general, however, the polygon display produced the best performance, allowing people to identify the state of the system accurately even under the most severe time stress. The superior performance of the polygon display resulted primarily from a high level of accuracy for instances low in uncertainty even under the most severely time-constrained condition. In this experiment the unique mapping characteristic of low levels of uncertainty was exploited by those using the polygon display in order to rapidly process that display, even under the most severely time-constrained conditions. Such a strategy allowed those people to correctly classify more than 75% of the instances low in uncertainty, even when instances were presented at a fraction of the original mean unpaced time.

Types of Perceptual Cues

An important goal of this research was to determine whether performance with the polygon display was attributable to integral processing or configurality effects (i.e., emergent features). For the results of this experiment to support integral processing rather than configurality effects, at least two conditions must be met. First, evidence of both filtering interference and redundancy gains with correlated dimensions must be present. Second, condensation efficiency—the primary criterion for defining configurality effects—should not be evident in these results.

The first condition is met by the interaction between display formats and uncertainty. The polygon display suffers from filtering interference under conditions of high uncertainty but benefits from redundancy gains under conditions of low uncertainty. Performance with the digital display provides an excellent contrast: there is no adverse effect of filtering with that display, no advantage for correlated dimensions, and no evidence of condensation efficiency. The digital display is clearly being processed in a separable fashion.

For the second condition to be met, individual features in the polygon display must interact to produce useful emergent features. In addition, condensation efficiency requires that at least one of the interacting features be irrelevant to the task. Thus emergent features can occur in this experiment when at least one of the system variables is nondiagnostic but the values of the other variables provide the information necessary for categorization to occur. Such a situation exists when uncertainty is in the range from moderate (Step 3) to high (Step 1). Although accuracy was found to improve significantly over that range of uncertainty, the results are more indicative of redundancy gains associated with integral processing than of configurality effects. As the correlation among dimensions increased, performance improved.

Perhaps one reason for the apparent reliance on integral processing in this experiment is that the polygon display allowed the complex relationships between system data and system states to be mapped onto a single, unique, and useful representation of that category. People could therefore use that representation to rapidly evaluate instances of system data that were clearly members of a state category. Such a strategy would be consistent with theories of concept learning that rely on a schema or prototypical instance to define category membership (Holland et al., 1986; Smith and Medin, 1981) and would produce results consistent with Garner's (1974) criteria for integral processing. In those situations in which the elements of the polygon display combine to produce useful perceptual cues, the processing demands of a multidimensional decision-making task can be reduced to the level of a more straightforward unidimensional discrimination task. The digital display, however, provides no such advantage, and people using that display format must separably process each of the relevant task variables. It is worth noting that the results of this experiment do not rule out the possibility of emergent features in this type of task but serve only to illustrate the integral processing of a polygon display when the cognitive demands of a task allow it.

Analysis of the two different types of errors (missed and wrong responses) provides further evidence of the integral processing of the polygon display and the separable processing of the digital display. The results presented in Figure 6 show that wrong responses to the polygon display were always greatest in the range of highest uncertainty, whereas wrong responses to the digital display were always independent of uncertainty, except in the most severely time-constrained condition. The major source of errors in the time-constrained conditions were, however, missed responses; the number of missed responses significantly increased with time stress and accounted for more than half of all errors for both display formats when time stress was most severe. Although there is some evidence in Figure 4a that uncertainty was a factor for the digital display group at the highest levels of uncertainty, the effect is significant only in the missed responses when the time constraints were most limited. One could speculate that at high levels of time stress, when people are forced to respond in a given period, they will come to rely on whatever perceptual cues are in the display irrespective of the type of format.

One can conclude from this study that time stress, uncertainty, and display formats interact in two fundamental ways. First, uncertainty significantly affects the way in which system data are processed, with the type of display format determining the conditions under which errors will occur. Second, the type and magnitude of error will be determined by both time constraints and the type of display format. Both results add to the evidence indicating that the interaction between uncertainty and display formats provides a reliable criterion for distinguishing among integrality, separability, and configurality effects.

Selecting the Best Display Format

One clear conclusion from this research is that the effects of time stress complicate the selection of display formats for system state identification tasks. Although designers of operator interfaces may be tempted to select the polygon display because of its apparent superiority under the most severe time constraints, such a choice would ignore a number of important trade-offs in performance.

Consider changes in processing strategy. Although overall performance with the polygon display was good, the digital display was superior in many situations. For instance, the digital display was unaffected by uncertainty, and overall accuracy remained high and unaffected by time constraints as long as time stress remained at moderate levels. Given that critical situations (such as severe accidents) are typically fraught with uncertainty, one could argue that the stable and predictable performance obtained with the digital display can be a distinct advantage in critical situations that involve moderate time constraints. Unfortunately, critical situations are often characterized by severe time stress, and the polygon display provides the best overall performance when the time available to process a display is severely limited.

The results of this research suggest, then, a number of potentially useful guidelines for selecting a display format.

First, performance will decline under time stress irrespective of the type of display format. Selecting a display format will mediate the adverse effects of time stress, but only in certain types of situations.

Second, polygon formats produce the best performance when there is a unique mapping of system data to system state categories. In other words, when uncertainty is low and the state of the system can be readily determined, an object-like display format that allows integral processing or the use of emergent features will produce the quickest and most accurate identification of system state.

Third, polygon display formats are not well suited for highly uncertain situations. Although previous research has shown that an object-like display format may allow people to quickly determine deviation from a normal state (Sanderson et al., 1989), this research suggests that such formats are less effective than separable display formats for identifying the specific state of the system under conditions of high uncertainty.

Finally, the precise representation of system data provided by a separable format produces the most accurate and stable overall performance and is the most useful format in highly uncertain situations when time stress is moderate. Unfortunately, that advantage quickly disappears when time available for processing is less than the minimum amount of time required to read the values of the process variables.

One potentially useful design guideline that emerges from this study is that both integral and separable displays should be used to ensure accurate and timely recognition of system state under all conditions of time stress and uncertainty. Such a guideline would be consistent with the principle of display design proposed by Coury et al. (1989) and discussed in the introduction, and it is supported by previous research that has demonstrated the superiority of a combined alphanumeric and graphical display format in the monitoring and control of a dynamic system (Coury and Pietras, 1989; Hooper, Coury, and Terranova, 1991).

All of these conclusions must be considered in light of one important factor. In the time-constrained conditions, the number of trials was fixed at 256. Consequently, the total amount of time spent in the 25% time-constrained condition was significantly less than that in the 100% time-constrained condition. One might argue that the relatively high level of performance in the 25% condition could not have been sustained if people had been required to perform the task for a longer period. This would assume that the magnitude of effort/duration or effort trade-off for cognitive task discussed by Coury and Drury (1986) could operate in this task. If such a trade-off exists, then more trials in the 50% and 25% condition of this experiment would have produced a greater decrement in performance.

In summary, then, the research presented in this paper has provided a number of significant insights into the effect of time stress on the processing of visual displays. The research has further demonstrated the need to consider the choice of display format within the context of time constraints and uncertainty.

ACKNOWLEDGMENTS

This research was supported by the U.S. Army Human Engineering Laboratory, Aberdeen Proving Grounds, Maryland, through Contract DAAA15-85-K-0010 to Bruce G. Coury. The authors would like to thank contract monitor V. Grayson CuQlock-Knopp for her support, encouragement, and input during the conduct of this research. The authors would also like to thank Nadeem Bezar for aiding in the collection of data and two anonymous reviewers for providing insightful and valuable comments that led to a much-improved version of this paper.

REFERENCES

Carswell, C. M., and Wickens, C. D. (1987). Information integration and the object display: An interaction of task demands and display superiority. *Ergonomics, 30,* 511–527.

Carswell, C. M., and Wickens, C. D. (1990). The perceptual interaction of graphical attributes: Configurality, stimulus homogeneity, and object integration. *Perception & Psychophysics, 47,* 157–168.

Coury, B. G., Boulette, M. D., and Smith, R. A. (1989). The impact of uncertainty on classification of multidimensional data with integral and separable displays. *Human Factors, 31,* 551–569.

Coury, B. G., and Drury, C. G. (1986). The effects of pacing on complex decision-making inspection performance. *Ergonomics, 29,* 489–508.

Coury, B. G., and Pietras, C. M. (1989). Alphanumeric and graphic displays for dynamic process monitoring and control. *Ergonomics, 32,* 1373–1389.

Coury, B. G., Zubritzky-Weiland, M., and CuQlock-Knopp, V. G. (1992). Probing the mental models of system state categories with multidimensional scaling. *International Journal of Man-Machine Studies, 36,* 673–696.

DeGroot, M. H. (1970). *Optimal statistical decisions.* New York: McGraw-Hill.

Dougherty, E. R. (1990). *Probability and statistics.* Englewood Cliffs, NJ: Prentice-Hall.

Garner, W. R. (1974). *The processing of information and structure.* New York: Wiley.

Garner, W. R. (1978). Selective attention to attributes and stimuli. *Journal of Experimental Psychology: General, 107,* 287–308.

Goldsmith, T. E., and Schvaneveldt, R. W. (1984). Facilitating multiple-cue judgments with integral information displays. In J. C. Thomas and M. L. Schneider (Eds.), *Human factors in computer systems* (pp. 243–270). Norwood, NJ: Ablex.

Green, D. M., and Swets, J. A. (1966). *Signal detection theory and psychophysics.* Huntington, NY: Robert E. Krieger.

Hamilton, V., and Warburton, D. M. (Eds.). (1979). *Human stress and cognition.* New York: Wiley.

Hockey, R. (1984). Varieties of attentional state: The effects of environment. In R. Parasuraman and D. R. Davies (Eds.), *Varieties of attention* (pp. 449–483). New York: Academic.

Hogarth, R. M. (1987). *Judgment and choice* (2nd ed.). New York: Wiley.

Holland, J. H., Holyoak, K. J., Nisbett, R. E., and Thagard, P. R. (1986). *Induction: Processes of inference, learning, and discovery.* Cambridge, MA: MIT Press.

Hollnagel, E., Mancini, G., and Woods, D. D. (1986). *Intelligent decision support in process environments.* New York: Springer-Verlag.

Hooper, K. N., Coury, B. G., and Terranova, M. (1991). Team performance in dynamic systems. In *Proceedings of the Human Factors Society 35th Annual Meeting* (pp. 949–953). Santa Monica, CA: Human Factors Society.

Kahneman, D. (1973). *Attention and effort.* Englewood Cliffs, NJ: Prentice-Hall.

Kahneman, D., and Triesman, A. M. (1984). Changing views of attention and automaticity. In R. Parasuraman and D. R. Davies (Eds.), *Varieties of attention* (pp. 29–61). New York: Academic.

Moray, N. (1984). Attention to dynamic visual displays in man-machine systems. In R. Parasuraman and D. R. Davies (Eds.), *Varieties of attention* (pp. 485–513). New York: Academic.

Morgan, M. G., and Henrion, M. (1990). *Uncertainty.* New York: Cambridge University Press.

Pomerantz, J. R. (1986). Overview of visual form perception. In E. C. Schwab and H. C. Nusbaum (Eds.), *Perception of speech and visual form: Theoretical issues, models, and research* (vol. 2, pp. 1–30). New York: Academic.

Pomerantz, J. R., and Pristach, E. A. (1989). Emergent features, attention, and perceptual glue in visual form perception. *Journal of Experimental Psychology: Human Perception and Performance, 15,* 635–649.

Rasmussen, J. (1986). *Information processing and human-machine interaction.* New York: North-Holland.

Salvendy, G., and Smith, M. (Eds.). (1981). *Machine pacing and occupational stress.* London: Taylor & Francis.

Sanderson, P. M., Flach, J. M., Buttigieg, M. A., and Casey, E. J. (1989). Object displays do not always support integrated task performance. *Human Factors, 31,* 183–198.

Smith, E. E., and Medin, D. L. (1981). *Categories and concepts.* Cambridge, MA: Harvard University Press.

Triesman, A. M. (1986). Properties, parts, and objects. In K. Boff, L. Kaufmann, and J. Thomas (Eds.), *Handbook of perception and human performance* (pp. 35/1–35/70). New York: Wiley.

Wickens, C. D. (1980). The structure of attentional resources. In R. S. Nickerson (Ed.), *Attention and performance VIII* (pp. 239–257). Hillsdale, NJ: Erlbaum.

Winston, W. L. (1987). *Operations research: Applications and algorithms.* Boston: Duxbury.