HUMAN FACTORS, 1993, 35(1), 71–97

Emergent Features and Graphical Elements: Designing More Effective Configural Displays

KEVIN B. BENNETT[1] *and* MONA L. TOMS, *Wright State University, Dayton, Ohio, and* DAVID D. WOODS, *Ohio State University, Columbus, Ohio*

When performing tasks in complex, dynamic domains individuals must consider information regarding both high-level constraints (relationships among several variables, performance goals) and low-level data (the values of individual variables). Previous research has revealed mixed results concerning the effectiveness of configural displays in achieving these dual design goals. Two empirical studies were conducted to investigate these issues using a laboratory analogue of a complex, dynamic task modeled on a real-world domain. Performance with a configural display, which highlighted the low-level data, was compared with performance with a bar graph display. For the extraction of information about high-level constraints in a memory probe task, the configural display significantly increased accuracy with no cost in latency. For low-level data there were no differences in accuracy across the two display conditions, but there was a significant cost in latency with the configural display. However, this cost was dependent on both experience and system state. These results suggest that configural displays can be designed to support the extraction of both high-level constraints and low-level data in complex, dynamic domains. To support the extraction of information for high-level constraints, the emergent features produced by a configural display must reflect the critical data relationships that are present in the domain. To support the extraction of low-level data, the graphical elements of the display must be made more salient perceptually through a variety of techniques, including emphasis of scale, spatial separation, and color-coding.

INTRODUCTION

As advances in computer science and artificial intelligence provide new computational power with the potential to support human problem solving, it is easy to overlook other techniques for aiding human performance and cognition. One form of decision support is representation aiding (Woods 1991a; Woods and Roth, 1988b; Zachary, 1986), in which machine power is used to create and manipulate representations of the target world, rather than to create autonomous machine problem solvers. Direct manipulation and graphic techniques are used to help the human problem solver to find the relevant data in a dynamic environment, to visualize the semantics of the domain (that is, make concrete the abstract), and to reconceptualize the nature of the problem. Although technological developments have placed powerful computer graphic capabilities in the hands of human-computer interface designers, how

[1] Requests for reprints should be sent to Kevin B. Bennett, Psychology Department, 309 Oelman Hall, Wright State University, Dayton, OH 45435.

best to use these capabilities to support human cognition in a wide range of tasks may not be fully understood.

To complete tasks in complex dynamic domains, individuals must consider a continuum of information ranging from high-level constraints (e.g., status of processes, or properties defined by the relationship between variables) and low-level data (measured values of individual variables). Analog formats are more likely than digital formats to support performance in these domains. Providing digital values for high-level constraints and low-level data alone does not provide effective decision support; these values are meaningful only with respect to the overall context. For example, the significance of a particular value for an individual variable depends on the relative contributions of related low-level data: without consideration of other variables it may be impossible to determine whether that value will increase, decrease, or remain the same. Thus, in order to be useful in complex, dynamic domains, displays should allow the parallel extraction of both high-level constraints and low-level data in the context of performance boundaries.

One type of analog format that has the potential to achieve these dual design goals is the *configural display*. A configural display represents high-level constraints of the domain through the relationships among the low-level data that define the constraint. Instead of noting the value of the high-level constraint directly, it is represented as an emergent property of the structure and behavior of the low-level data. The emergent property is based on a relationship among low-level graphical elements or the configuration of these elements—in other words, there is an emergent pattern or configural property of the display. The pattern of configural relationships can be structural or dynamic—that is, the behavior or movement of graphical elements relative to others. Because the relationships that define the emergent property sometimes refer to how graphical elements fit together to create higher-order visual objects (e.g., two variables mapped in the x and y dimensions form the geometric pattern of a rectangle), they are sometimes referred to as *object displays*.

For example, a configural display currently used in some nuclear power plants is the polar graphic display: a polygon with eight spokes (Woods, Wise, and Hanes, 1981). The spokes of the polygon are dynamically scaled so that a regular polygon represents normal conditions and distortions in the polygon represent a developing abnormality. This display integrates information from more than 100 individual sensor values in order to represent process health.

The majority of laboratory researchers who have investigated the dual design goals for extraction of high-level constraints and low-level data have posed their questions in terms of object versus separable display formats. In a *separable display*, each individual state variable has its own unique representation (e.g., a bar graph). The primary principle that has been used to guide design and evaluation is that of proximity compatibility (Carswell and Wickens, 1987). The initial version of this principle made explicit and strong predictions regarding the dual design goals: namely, the extraction of high-level constraints will be facilitated with an object display format, and the extraction of low-level data will be facilitated with a separable display format. Although these predictions have been qualified in subsequent revisions (e.g., Wickens and Andre, 1990), the principle emphasizes the nature of the representation (the form of the graphic display) and the information-processing characteristics of the individual using that representation.

A number of alternative principles of display design suggest that consideration of these two factors is necessary—but not sufficient—for effective display design. From the perspective of the cognitive system triad (Woods and Roth, 1988b), the quality of performance in complex, dynamic domains is the result of three interactive and mutually constraining components: the cognitive demands produced by the domain of interest, the cognitive agent (or agents) that meet those demands, and the representation of the domain through which the agent experiences and interacts with the domain. The characteristics of each component, and the interactions among those characteristics, determine the ease or difficulty of problem solving. Similar conclusions have been expressed by Flach and Vicente (1989), Vicente and Rasmussen (1990), and Rasmussen (1986). These design perspectives stress that graphic form and information-processing characteristics also interact with, and are constrained by, the characteristics of the domain. Specific implications for the design of configural displays will be considered in the following section.

High-Level Constraints: Semantic Mapping

Viewed from these perspectives, the overall goal for display design is to map the domain semantics (low-level data, high-level constraints, and relevant performance goals) into the appearance and dynamic behavior of a graphic display so that this information is readily available (easily extracted or decoded by the user). Rather than objectness per se, the critical factors in the design of a configural display are the number and quality of *emergent features* (Pomerantz, 1986) that are produced by a graphic format and the mapping from these emergent features to the constraints in the domain (Bennett and Flach, 1992). Emergent features are the high-level, global perceptual features that are produced by the interactions among individual parts or graphical elements of a display (e.g., lines, contours, and shapes). For example, mapping two variables into a rectangle (one variable in the x axis and one variable in the y axis) produces the emergent feature of area. Emergent features are "dependent upon the identity and arrangement of parts but not identifiable with any single part" (Pomerantz, 1986, p. 8).

Different configural display formats will produce different emergent features, and the number and quality of these emergent features will contribute to its effectiveness. In a bar graph the interactions among the graphical elements are held to a minimum, and therefore the number and quality of emergent features are reduced. However, as Sanderson, Flach, Buttigieg, and Casey (1989) have illustrated, under certain conditions bar graphs can produce configural patterns that facilitate performance. Digital displays do not have a graphic form, do not produce configural patterns, and are therefore truly separable. Thus the configural-separable dichotomy is more appropriately characterized as a continuum with geometric pattern displays, bar graph displays, and digital displays in descending order with respect to the quality of the configural patterns they produce.

The second critical factor in designing configural displays for the extraction of high-level constraints concerns the mapping between the emergent features produced by the configural display and the cognitive demands produced by the domain. The emergent features must reflect the inherent data relationships that exist in the domain—that is, the highly salient emergent features must correspond to the information needed to complete domain tasks. In the previous example the rectangle configural display will be effective only if the user must make a decision based on the product of the two variables (which is directly reflected in the emergent feature of area). In fact, if the emergent features do not correspond to critical information, these highly salient perceptual features will have to be ignored and could actually degrade performance—hence the term *semantic mapping*. Thus it is not the presentation of domain information in a configural format per se that determines the effectiveness of a display (Sanderson et al., 1989). Rather, the crucial determinant of success is how well the inherent data relationships have been mapped into the appearance and dynamic behavior of the configural display (Bennett and Flach, 1992; Woods, 1991b).

Despite its intuitive appeal and practical success, research using laboratory tasks and naive subjects has produced mixed evidence regarding the effectiveness of configural displays in tasks that require the consideration of information from a number of variables (i.e., integration tasks). Experimental results usually indicate that the display of information in a configural format improves integration performance (Barnett and Wickens, 1988; Carswell and Wickens, 1987; Goldsmith and Schvaneveldt, 1984; Wickens and Andre, 1990), suggesting that there is a cost involved with the display of information in a more separable display. However, other studies have shown neither costs nor performance advantages of more separable display formats (Coury, Boulette, and Smith, 1989; Sanderson et al., 1989). A principle based on configurality appears to have a great deal of explanatory power, with respect to both explaining why a separable display might support performance better than an object display could and explaining differences between two object displays.

Low-Level Data: Salient Graphical Elements

It is often assumed that the presentation of information in a configural display format incurs a cost when performance depends on the consideration of an individual variable or when irrelevant variables need to be ignored; such tasks are usually referred to as *focused tasks*. Either implicitly or explicitly, this potential cost is assumed to result from the fact that perception of the individual elements is secondary to perception of the object itself. That is, as a function of being part of an object, information related to the individual parts is somehow less accessible. Several theories of attention and object perception cast doubt on this explanation. The *object file theory* (Kahneman and Treisman, 1984) suggests that once attention has been allocated to an object, information about low-level data should be readily available. Pomerantz (1986) argued that there is no evidence of "perceptual glue" that binds individual elements into a whole, therefore making the individual parts less accessible. He stated that "the groupings of parts into wholes does not make parts imperceptible or inaccessible; the evidence for a perceptual glue binding parts together is thin. Rather, when parts configure into wholes, new emergent features arise that are available alongside the parts but that are more salient perceptually and will therefore be attended to instead of parts when it is advantageous to do so" (Pomerantz, 1986, p. 28).

Both of these theoretical perspectives suggest that information about low-level data can be available in parallel with high-level emergent features. Thus it may be possible to design configural displays that allow an observer to attend to one or another aspect of the graphic form (global precedence effect) as a function of task demands. Pomerantz's theory suggests that low-level data can be made more available by making the graphical elements more salient perceptually. This might be accomplished through a variety of techniques: for example, maintaining or emphasizing the scale, separating the contributing elements spatially, color-coding the contributing elements, or even directly noting the values of low-level variables as an element of the configural graphic. When increasing the perceptual salience of the graphical elements representing low-level data, it is important not to destroy the emergent features that are relevant to high-level issues.

Most of the laboratory research that has been conducted to investigate potential costs with configural formats has not used the aforementioned techniques to make low-level data more salient (however, see Wickens and Andre, 1990, for a counterexample) and has therefore been biased toward finding a cost. Despite this, the results do not support the existence of an inherent cost to the display of low-level data in a configural format. The most common finding is a lack of significant performance differences between the two types of display formats (Bennett and Flach, 1992). When significant differences have been observed, they usually have favored more separable display formats (Carswell and Wickens, 1987; Casey and Wickens, 1986; Wickens and Andre, 1988, 1990), though advantages for configural displays have been reported (Wickens et al., 1985, Experiment 1).

Discovering the Semantics of a Complex Domain

The previous discussion suggests that configural displays can be designed to support the extraction of both high-level constraints and low-level data. We investigated these issues using a simulated version of a complex, dynamic control task: the manual control of water level in a boiler during the start-up of a power plant. This is a *nonminimum phase* dynamic system that is characterized by right half plane zeros in the transfer function relating output to input. One example of nonminimum phase behavior is a case in which the initial response of the output to change in an input variable is to move in the opposite direction of the change in input before going to its positive asymptote. Another example is pure time delay, in which a lengthy gap occurs between the time an event occurs or an action is taken and the time its effect on other variables can be observed.

It must be emphasized that the goal of these studies was not to make specific claims regarding the advantages of a particular configural display for a particular task. Rather, the goal was to investigate central issues in configural displays to discover design principles that will generalize to a variety of domains. Because of the importance of domain semantics for both design and evaluation, these issues must be investigated in domains that approximate the complexity found in the real world (Woods and Roth, 1988b). The manual control of feedwater is representative of the multiple, interleaved task situations that are found in the real world: it has highly intermingled subtasks that require the consideration of both individual variables and high-level constraints. Thus it affords an excellent opportunity to investigate critical issues in configural display design and evaluation in a task that approximates the complexity found in the real world.

The laboratory task is a second-order simulation of the system dynamics and was designed to capture the critical demands that the actual task places on human performance (see Roth and Woods, 1988, for a detailed investigation of this task from a cognitive perspective). A similar laboratory process system with simpler dynamics has been used in studies of process control skill (see Crossman and Cooke, 1974; Moray, 1987; Moray, Lootsteen, and Pajak, 1986). In the real-world version of the task, energy from a source (in this particular case, a nuclear reactor) is used to convert water to steam in multiple boilers or steam generators (SG). When the supply of steam is sufficient, it is used to load a turbine to generate electricity. During the start-up of this process, the task of the feedwater operator (one of several human operators who monitor and control the process in conjunction with automatic control loops) is to control feedwater flow in order to maintain the water level in the steam generators between high and low boundaries. Exceeding these limits results in a plant shutdown, which has high economic penalties, and the start-up must begin anew.

Previous research has established that several types of task demands make performance difficult (Roth and Woods, 1988). First, two goals interact in feedwater/level control: (a) to generate electricity (that is, to generate sufficient steam to meet the electricity goal) and (b) to maintain indicated steam generator level (ISGL) within limits. The processes that effect these goals are the energy inflow/outflow through the steam generator (energy inflow vs. the energy leaving in the form of steam to the turbine or to some other location) and the water (mass) inflow/outflow through the steam generator (mass inflow in the form of feedwater flow [FF] vs. mass outflow in the form of steam flow [SF]).

Goal competition can arise because changes in each process affect both goals. Thus control input that helps to satisfy one of the goals can, at the same time, degrade the status of the other goal. Changes in steam and feed flow affect level by changing the mass balance. Changes in these processes also affect the energy balance (the temperature of the feedwater relative to the water in the steam generator and the rate of steam flow is proportional to the rate of energy outflow). Although changes in the energy balance do not affect the amount of water mass in the steam generator, they do affect the energy content of the water and therefore the measured level. The level control goal is specified in level units, not mass; one can think of the measured level as a measure of the volume occupied by the mass of water present. As a result changes in energy affect the level goal, but at a time constant that is different from that of the mass effects. Strong pressure to complete the start-up rapidly intensifies competition between goals. Rapid maneuvers produce large disturbances in SG level that must be counteracted; they also introduce multiple forces that simultaneously act on the level, complicating situation assessment, especially the assessment of where the level is going to be in the future, based on past actions and influences.

The demand characteristics of the task can also be described in terms of the dynamic system properties. Most feedwater/boiler systems, and many other systems, are characterized by nonminimum phase dynamics, which are characterized by right half plane zeroes in the transfer function relating output to input. One example of nonminimum phase dynamics is pure time delay, in which there is a lengthy gap between the time an event occurs or an action is taken and the time its effect on other variables can be observed. Another example of nonminimum phase behavior is when the initial response of the output to change in an input variable is in the opposite direction of the change in input before going to its positive asymptote. In the manual control of feedwater (MCF) task operators referred to this as *shrink/swell effects.* In shrink/swell, the initial effect of a change in a state variable on SG level is the opposite of its long-term effect. For example, when the flow of relatively cold feedwater is increased, the fact that the energy of the feedwater is less than the energy level in the SG causes the net energy in the SG to decrease. The energy drop decreases the indicated level at one time constant while the net increase in mass inflow increases the indicated level, but at a longer time constant. The result is that an increase in feedwater flow results in indicated level first decreasing (shrink) because of the energy effect and then increasing because of the mass effect. Because of time delays and shrink/swell effects, only weak evidence about critical-state variables is available to controllers. In addition, accurate measures of steam flow and feed flow do not exist at low power.

The complex dynamics place a premium on a controller's ability to anticipate the effects of changes in plant state or to control actions on SG level and make compensatory responses before the ultimate effect of the event on SG level is seen. In order to do this the controller needs to know what energy and mass factors have been introduced into the system. However, the dynamics and the poor state information make this assessment difficult (see Roth and Woods, 1988). For example, determining whether a change in ISGL represents a longer-term change in mass or a transitory energy effect can be quite difficult.

Representation Aiding in the Feedwater/Boiler Control Task

The cognitive analysis of this task (Roth and Woods, 1988) showed that performance could be enhanced if information were provided that helped the controller to better anticipate level behavior. Effective assistance should help the problem solver to build a better situation assessment of where the level is, what factors are influencing the level (shrink/swell vs. changes in the mass balance), and where the level will go given these influences and possible interventions. Performance at the feedwater/level task fundamentally revolves around separating the relative contributions of two independent functional processes on indicated level (water mass effects from the effects of energy changes or shrink/swell) or providing cues that would assist operational personnel to do this.

A new form of predictive information was developed for nonminimum phase dynamic systems (see Haley and Woods, 1988; Woods and Roth, 1988a) to assist the operators in separating the contributions. The compensated steam generator level (CSGL) is a measure of the mass contribution to indicated SG level, calibrated in terms of indicated level units because the shutdown limits are expressed in terms of indicated level. It is an estimate of current SG level that is not confounded by any shrink/swell effects or, alternatively, it is an estimate of SG mass transformed into level units. CSGL provides information about the future course of indicated level. It can be thought of as an indication of where measured SG level will be in the future once shrink and swell effects dissipate. Note that CSGL is calculated for each time sample of data; it is not a literal projection of level behavior into the future, though it can be used to derive such a projection. It computes what indicated level would be if the energy state of the SG were nominal. Initial studies have shown that compensated state variables aid human control performance in nonminimum phase dynamic systems (Bennett, Woods, Roth, and Haley, 1986).

To support human control performance, then, a visual representation or display (Woods, 1991b) must be designed which portrays state values in a form so that the operator can better anticipate and control the feedwater/boiler process during start-up and thereby avoid unnecessary plant shutdowns. The information to be communicated in this representation revolves around showing the operator the relationship between state variables related to the goal of the task—that is, to keep the level within limits. Data on state variables associated with energy and mass balance processes that influence level behavior—feedwater temperature, feedwater flow, steam flow, and energy inflow—must be integrated and interrelated with the goal data. The operator must integrate this diverse set of data in order to assess where the level is, what factors are influencing the level (changes attributable to shrink/swell vs. changes in mass balance), and where the level will go given these influences and possible interventions.

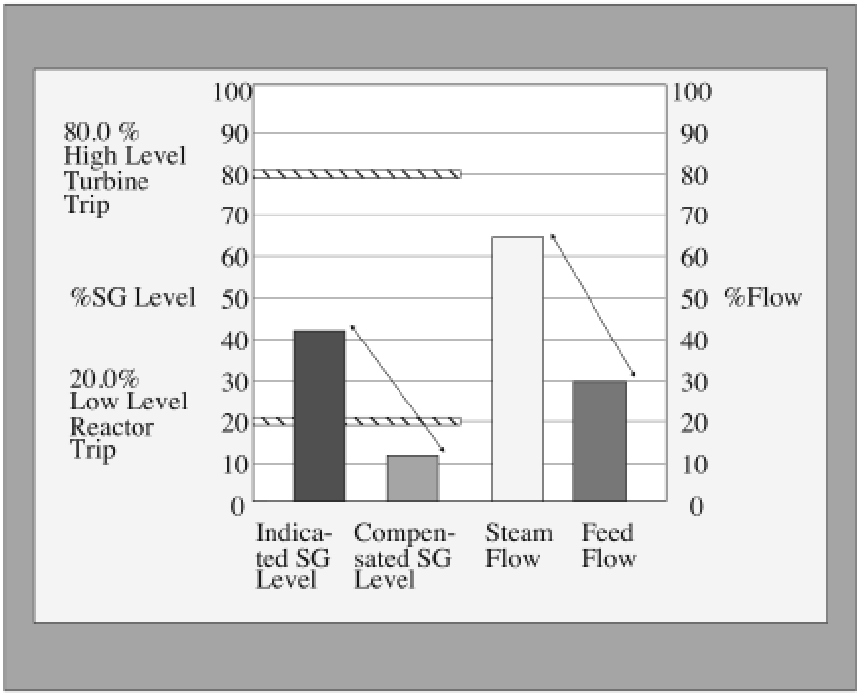
We used a scaled-world simulation of this dynamic control task to investigate issues about how separable and configural displays of data support human performance. This task is appropriate to this end on several grounds. First, task performance requires consideration of several interacting state variables. Second, it is a tractable laboratory version that possesses the same critical demand characteristics as the real-world analogue. Third, past studies have examined separable and configural display issues in laboratory tasks that have no relation to actual settings or actual displays, whereas the data obtained in this study directly relate to a real-world situation, and the displays investigated relate more closely to displays that practitioners might actually use.

Mapping the Domain Semantics into Computer-Based Configural Displays

Two displays, a configural and a bar graph display, were developed for the scaled-world simulation of the MCF task (see Figures 1 and 2). As previously mentioned, bar graph displays can produce emergent features; we chose the label, *separable*, for this display so as to be consistent with previous research. However, we maintain that configural properties should be defined as a relationship between domain semantics and the properties of its visual representation, not as a property of the visual appearance of a graphic (Bennett and Flach, 1992; Woods, 1991a, 1991b). The separable display mapped the four task variables (ISGL, CSGL, SF, and FF) into separate, color-coded bars with a common baseline (see Figure 1). The upper and lower set point boundaries are represented by the horizontal lines at 20% and 80% on the left scale; the set point boundaries are relevant only to indicated and compensated SG level and thus do not extend across the display. Figure 1 illustrates the separable display in two different process states. As discussed previously, the relationships between SF/FF and ISGL/CSGL are highly informative about the state of the process and about appropriate control actions. For example, the value of CSGL functions as a prediction of the future state of ISGL. The emergent features in the separable display that correspond to these relationships are the inferred linear relationships between the tops of the bars. These relationships are represented by the lines with arrows between bars in Figures 1a and 1b—these lines are included for illustrative purposes and were not present in the actual display. Thus for the variables ISGL and CSGL the inferred line indicates that the future trend of indicated level will be decreasing in the situation captured in Figure 1a. Similarly, the difference between SF and FF indicates the state of the mass balance component of the process. In Figure 1a steam flow greatly exceeds feed flow, indicating that ISGL will eventually decrease; in Figure 1b the two flows are relatively balanced, indicating that ISGL will remain the same in the near future. Note that there are partially redundant sources of information in the variables shown. Mass balance state can be extracted from a comparison of the steam and feed flow indications or from the trend in compensated level: if inflow and outflows are equal, the steam and feed flow bars are equal in height and compensated level is constant over time. (Mass balance cannot be determined from indicated level alone because of delays and shrink/swell effects.)

In Figure 1a ISGL is well within the shutdown limits, but CSGL indicates that in the future indicated level is in danger of crossing the lower set point value. A comparison of the steam-feed flow bars indicates that mass outflow is much higher than mass inflow. The controller should increase feed flow to eliminate the imbalance and generate net inflow, but the complex dynamics place the controller in a dilemma. Increasing feed flow will create an energy effect—a shrink effect—on ISGL. In other words, the correct input for resolving the problem in the long run will, in

(A)



(B)

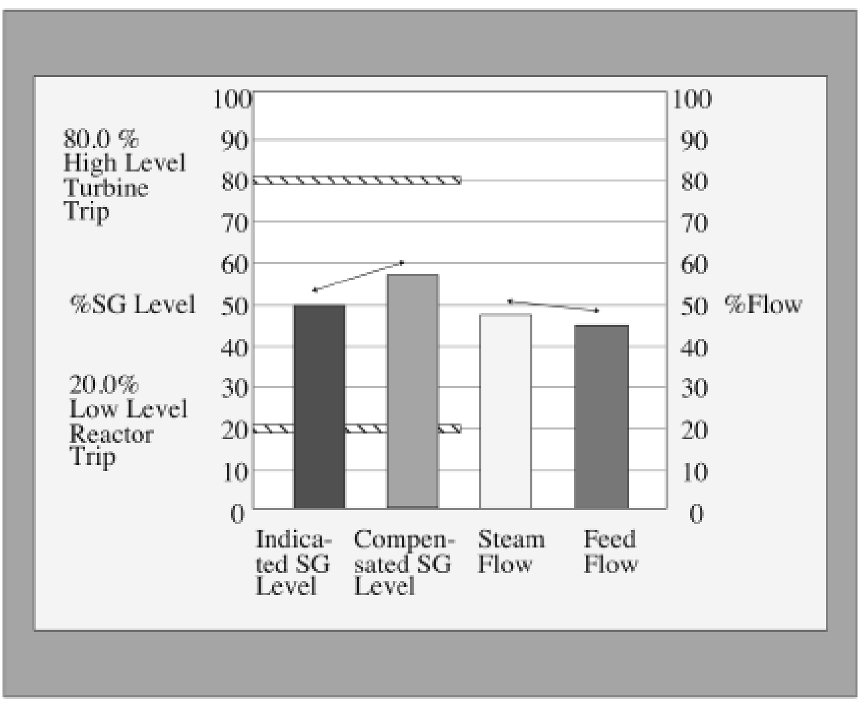


Figure 1. *The separable bar graph display used in Experiments 1 and 2. Each of the four variables are represented by color-coded bar graphs with a common baseline. The upper and lower set point boundaries are represented by the horizontal lines at 20% and 80% on the scale (relevant to ISGL and CSGL only). The four short lines with arrows represent the emergent feature of inferred linear relationship (these lines were not present in the actual displays). See text for details of (A) and (B).*

the short run, exacerbate the trouble and fail to resolve this threatening situation. This example illustrates several aspects of the task from an operational point of view. One aspect is situation assessment—that is, identifying the state of the process. Another aspect is response formulation: developing/selecting response strategies to avoid or recover from trouble. For example, subjects may learn to recognize that the state illustrated in Figure 1a is very undesirable before they learn response strategies to avoid or recover from this state. In addition, this example illustrates the critical role of anticipating developments in process state to avoid undesirable states.

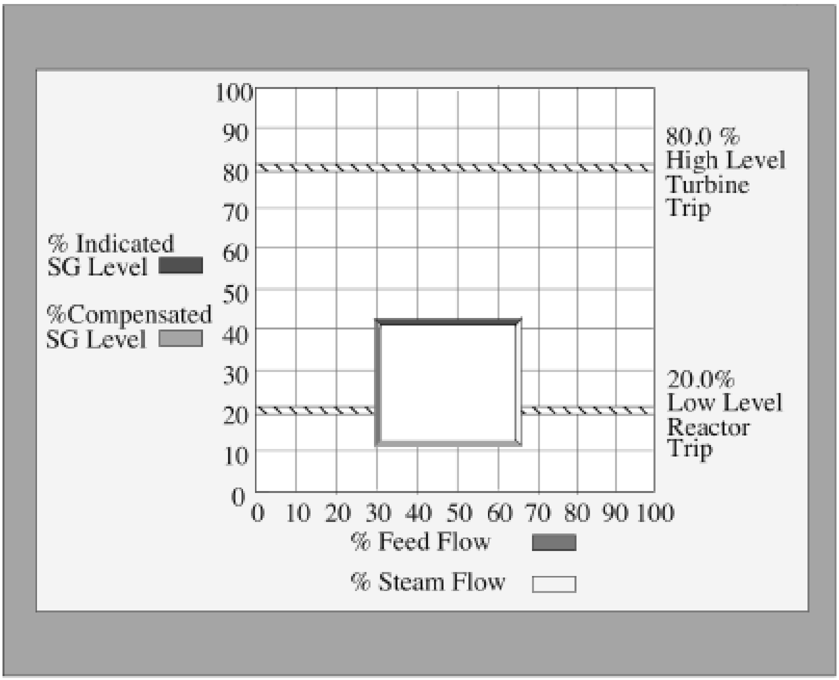
In comparison, Figure 1b illustrates the separable display when the process is well-balanced; ISGL is within the target region and is likely to remain there in the near future. Overall, the mapping of data onto bar charts produces emergent cues about the state of the process. Wide variations in bar heights, especially between relevant pairs of bars, are a cue that the process is in an undesirable state (Figure 1a). A relatively linear arrangement of bar heights is a cue that the system is in balance or needs only minor tuning inputs (Figure 1b).

The configural display provides additional emergent features, or cues, that testify about the state of the process. These cues highlight the highly coupled and interconnected nature of the relationships of the variables. Figures 2a and 2b illustrate the same process states as in Figures 1a and 1b (i.e., the values of the four variables are exactly the same). This format mapped the four variables onto a single geometric object: a rectangle (see Figure 2). The values for indicated and compensated SGL are plotted on the vertical axis (imagine two horizontal lines extending from the left to the right of the display grid); the values for steam flow and feedwater flow are plotted on the horizontal axis (imagine two vertical lines extending from the top to the bottom of the display grid). Only the rectangular shape resulting from the intersection of the four variables is actually shown in the display. The sides of the rectangle are color-coded to reflect the contribution of an individual variable, and the set point boundaries are represented by the horizontal lines at 20% and 80% on the vertical scale.

Mapping these four variables onto a single geometric object results in a number of emergent perceptual features. For example, the critical relationships between SF/FF and ISGL/CSGL are directly mapped onto the height and width of the rectangle, respectively. Additional emergent features include the area and shape of the rectangle, the location of the rectangle in the display grid, and the direction and rate of movement within this grid. For example, in Figure 2a the undesirable process state can be seen from the size of the rectangle (i.e., large differences between ISGL and CSGL and between inflow and outflow) and its position over the marker of the lower limit value. In Figure 2b the rectangle is small, indicating that the critical parameters are approximately in balance and the rectangle is positioned well within the target region. Thus the critical relationships are directly mapped onto the graphic form, and strong perceptual cues that signal the state of the process are produced. Note that the dynamic cue of seeing the rectangle move relative to the target area and limits and seeing it change in size—very compelling parts of this graphic form—are not captured in the static illustrations in Figure 2.

To successfully complete the MCF task, the user must consider several high-level constraints (e.g., mass balance, energy balance) so that the current system state can be assessed and the correct control input determined. Part of the inherent difficulty of the task arises because of the tight intercoupling between subsystems. Changes in a primary

(A)



(B)

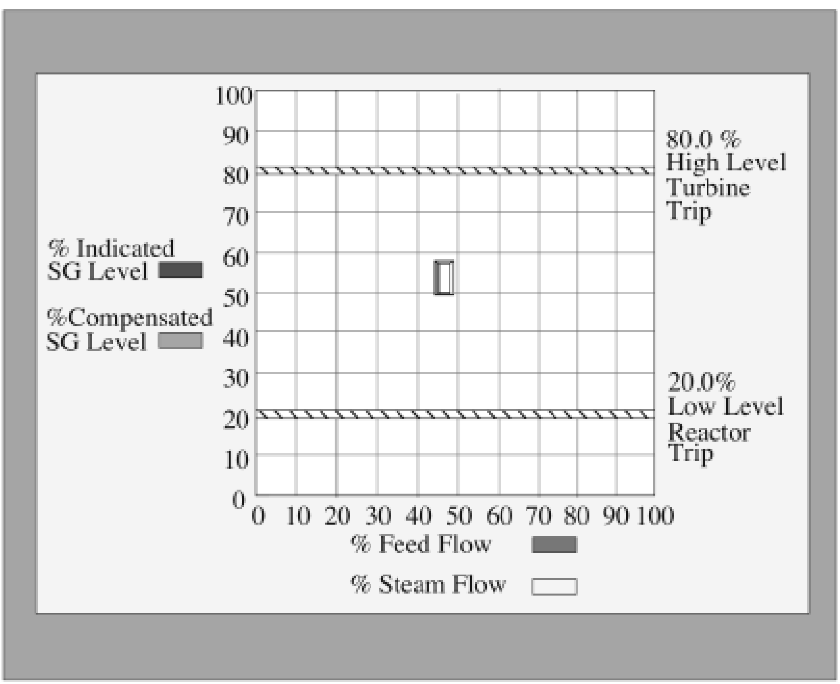


Figure 2. *The configural display used in Experiments 1 and 2. The ISGL and CSGL variables are plotted in the vertical axis (represented by horizontal lines); the steam and feed flow variables are plotted in the horizontal axis (represented by vertical lines). Where the lines representing the four variables intersect a rectangle is formed, and the sides of this rectangle are color-coded to reflect the individual contributions. The upper and lower set point boundaries are represented by the horizontal lines at 20% and 80%. See text for details of (A) and (B).*

variables (e.g., feedwater flow or steam flow) will produce changes in both the mass balance and the energy balance, which in turn, affect the critical performance variable, ISGL. The configural display provides a representation that highlights this tight coupling by providing a number of emergent features, including the height, width, and area of the rectangle, and its location, direction of movement, and rate of movement within the display grid. In contrast, the separable display provides only one primary emergent feature: the inferred linear relationship between bar heights. Experiment 1 compares performance at the MCF task with the configural and separable displays.

EXPERIMENT 1

Method

Subjects. Twenty students (10 men and 10 women) participated in the experiment and were paid $5.00 an hour. The subjects' ages ranged from 19 to 35 years of age, and all had normal or normal-corrected vision with no color-blindness deficiencies. The subjects were assigned to one of the two display conditions randomly.

Apparatus. All experimental events were controlled by a general-purpose laboratory computer (Sun Microsystem 4-110 Workstation). Subjects were seated in an enclosed experimental room. A color video monitor (40.64 cm) with a resolution of 1152 × 900 pixels was used to present experimental prompts, and a standard keyboard was used to record user responses.

Simulation model. The subjects' task was to control a second-order nonminimum phase dynamic system (see Figure 3). The simulated system possesses the same basic dynamic characteristics as a single nuclear power plant steam generator (e.g., time delays, shrink/swell behavior). It incorporates the influence of a number of factors on the indicated steam generator level, including steam flow, feedwater flow, excess heat, and temperature of the feedwater. The simulation also computes the value of CSGL. The appendix contains a description of the simulation.

System dynamics. The excess heat and the feedwater temperature parameters remained fixed throughout the experiment. Programmed changes to the steam flow parameter produced the primary challenges that the subjects had to meet. There were two types of changes: (1) continuous variation and (2) disturbances (see Figure 3). The continuous variations were constant changes to steam flow resulting from the combination of three sine waves and a ramp. The net result was to produce a trial in which the steam flow rate was either oscillating, oscillating with a gradual rise, or oscillating with a gradual fall (a rising ramp is illustrated in Figure 3). The second type of change was the result of asynchronous changes, or disturbances, to the steam flow parameter. The number and size of these disturbances varied as a function of the elapsed time of an experimental trial. During the first 30 s of an experimental trial one randomly timed disturbance was introduced. For each 30-s increment in time that followed, the number of disturbances was increased by two. The direction of each disturbance was either positive or negative and was also randomly determined. The size of the disturbance was randomly selected within a time-dependent range: disturbances within the first 30-s time interval ranged between 0% and 1%, and for each additional 30 s the range was increased by 0.5%. The effect of the programmed changes to steam flow was to produce an experimental trial that changed constantly in a relatively unpredictable fashion and that became progressively more difficult as time on task increased.

Stimuli. Two different graphic displays were used to present ISGL, CSGL, SF, and FF

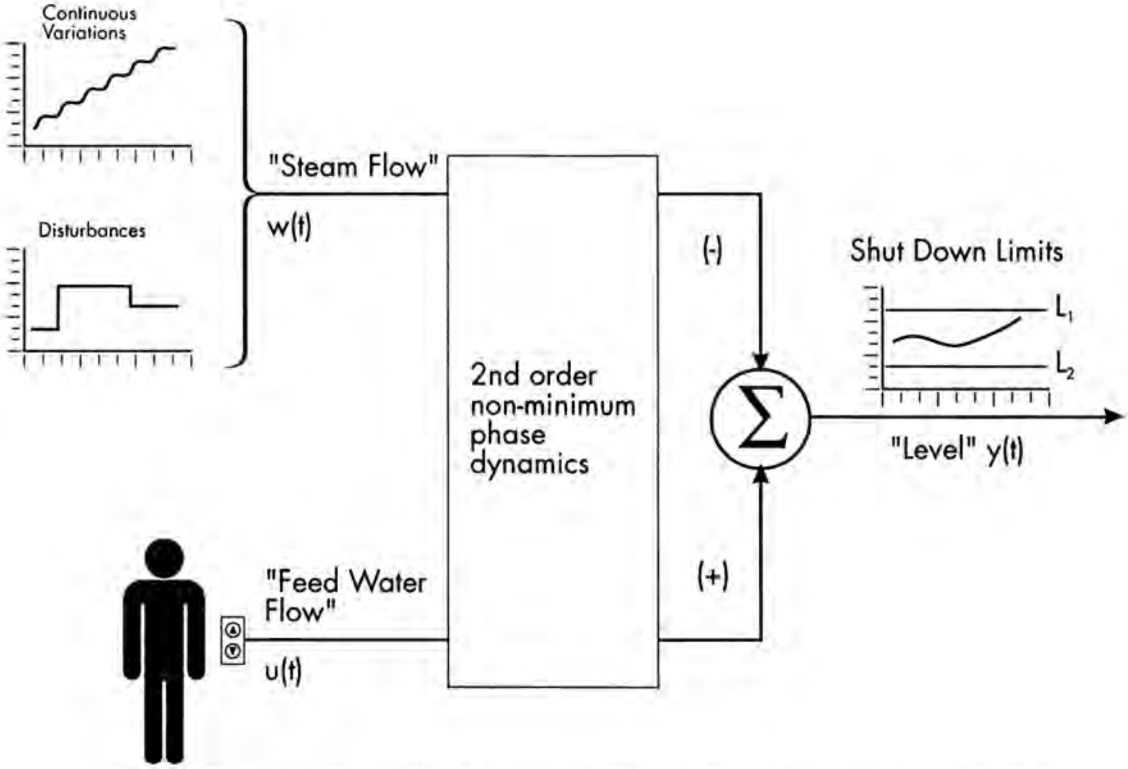


Figure 3. *The manual control of feedwater simulation and the experimental control task.*

(see Figures 1 and 2). The background mat for each graph was 11.5 cm high × 14 cm wide; it was colored a light gray, and the lines and letters were colored black. Both displays were updated with information from the simulation model every 2 s, and random noise ranging from −2% to +2% was added to the value of each variable displayed (these changes were not permanently added to the mathematical model, only displayed to the subject). For the separable display each variable was presented as an individual bar with a common baseline in the *x* axis. Each bar was color-coded: blue for ISGL, white for SF, green for FF, and yellow for CSGL. Each bar was 0.6 cm wide with a maximum height of 5.5 cm and was equally spaced in the *x* axis. Assuming a subject seating distance of 50 cm, the bar chart display subtended a visual angle of 4.9 deg horizontally and a maximum of 6.3 deg vertically. The *y* axis was labeled using a scale of 0% to 100%, and black, horizontal grid lines extended the length of the *x* axis and were placed at 10% intervals. The trip set points were placed at 20% and 80% as horizontal red lines that began on the left side of the display and extended half the length of the *x* axis.

Whenever possible, the configural display used the same sizing, scaling, and coloring conventions as did the separable display. In the configural display the *x* axis was also scaled and labeled, and the trip set points were extended to cover the length of the *x* axis. The configural display mapped the four critical variables onto a single geometric object: a rectangle. The difference between ISGL and CSGL was mapped onto the *y* axis, and the difference between steam flow and feedwater flow was mapped onto the x axis. A black rectangular shape was drawn at the points where the four variables intersected in the display grid, and the sides of the rectangle were color-coded (using the same colors as the separable display) to reflect the contribution of an individual variable. The maximum dimensions of the rectangular object were 4.5 cm high by 5.5 cm wide. Assuming a subject seating distance of 50 cm, the configural display subtended a maximum visual angle of 6.3 deg horizontally and a maximum of 5.1 deg vertically.

Procedure. The experiment was conducted during a one-week period; one experimental session lasting one hour was performed each day, for a total of five sessions. The subjects were individually tested in an enclosed room. During the first experimental session the subjects were provided with a written and a verbal explanation of the task and a verbal description of their respective display. The experimenter remained in the room during this session to answer any general questions about the task but did not provide any information regarding control strategies. During each experimental session the subject completed at least 12 trials, a factorial combination of three steam flow ramp types (rising, null, falling) and four ISGL starting positions (35%, 45%, 55%, and 65%). The order of these trials was randomly determined. To ensure equal training time, each subject was required to complete additional trials (a randomly determined subset of the original 12 trials) until one hour had expired. Because the dependent measure was time on task, this manipulation provided an additional control for experience. The additional trials were not considered in the data analyses.

During each experimental trial the rate of feedwater flow was under the control of the subject, whose task was to adjust this rate by pressing an up arrow to increase feedwater flow and a down arrow to decrease feedwater flow in order to maintain the ISGL between the upper and lower trip set points (see Figure 3). The subject initiated a trial by pressing a designated key. Each trial could last up to 5 min and ended either when 5 min had elapsed or when the ISGL surpassed one of the two trip set points. The subject was provided with feedback concerning time on task after each trial.

In summary, the experimental design contained four independent variables: display (separable vs. configural, between-subjects), day (1–5, within-subjects), starting position of ISGL (35%, 45%, 55%, and 65%, within-subjects) and steam flow ramp type (rising, null, falling, within-subjects). The primary dependent variable was time on task, which was recorded at 1/100-s accuracy.

Results

A $2 \times 5 \times 4 \times 3$ mixed analysis of variance (ANOVA) was performed on the regular time-on-task scores. The main effects of day, $F(4,72) = 2.64$, $p < 0.04$, starting position, $F(3,54) = 12.74$, $p < 0.0001$, and ramp, $F(2,36) = 6.60$, $p < 0.004$, were significant, as was the interaction between ramp and starting position, $F(6,108) = 19.70$, $p < 0.0001$. The main effects can be summarized by stating that performance improved with experience at the task, the rising and falling ramps were more difficult than the null ramp, and performance was degraded when the ISGL starting position was closer to the trip set points. The interaction effect indicates that performance was particularly poor when (1) the ISGL starting position was low and the steam generator ramp was rising or (2) the ISGL starting position was high and the ramp was falling. All other effects, including the predicted main effect of display, $F(1,18) = 0.03$, $p < 0.85$, were not significant. For comparison purposes the means for the Display × Task × Day interaction effect are illustrated in Figure 4.

Discussion

Experiment 1 compared performance in a complex, dynamic task when information was presented on a separable graphic display format with that obtained when using a configural display format. Contrary to expectations, the configural display did not facilitate performance for the MCF task significantly, as measured by the time-on-task variable. One possible explanation is that the emergent features produced by the separable display (the inferred linearity between the tops of bars) are equally effective in facilitating extraction of task-relevant information as the emergent features produced by the configural display (height, width, area, shape, location, and movement of the rectangle). A second possibility is that the five hours of experience were not sufficient for the observers to become proficient at this rather difficult task, regardless of the graphic form. A third possibility is that the expected performance differences do exist but that the time-on-task variable was not sensitive enough to capture them.

Two of these three possibilities were investigated in Experiment 2. To assess the possibility that additional experience at the task would reveal more distinct time-on-task performance differences, the same observers participated in the follow-up study. In addition, a second task was included to directly assess the quality of the emergent features that the two displays produced. Observers were given *retrospective memory probes:* the screen was blanked and observers were asked to recall information about the state of the system. Wickens and Andre (1988, 1990), Wickens

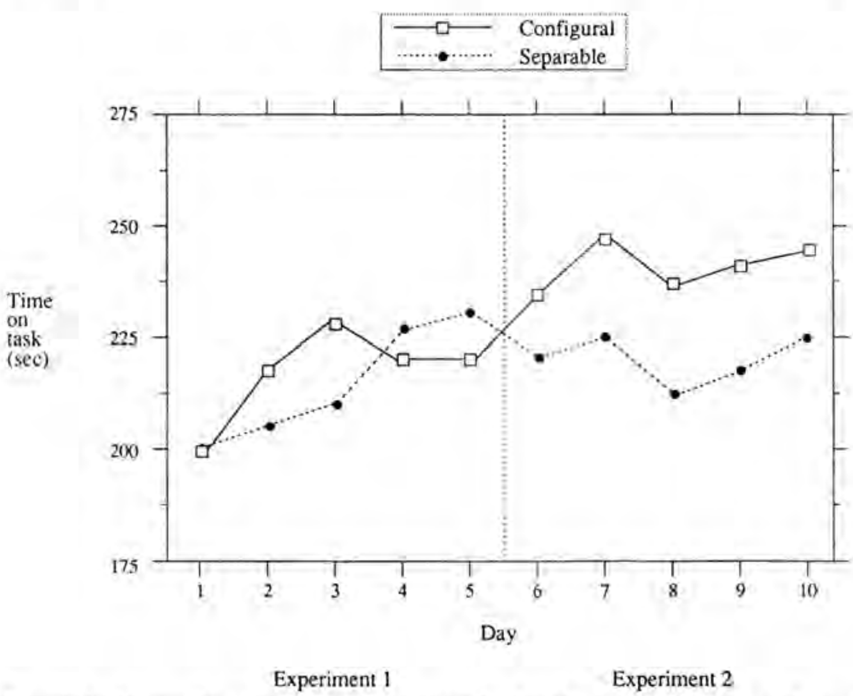


Figure 4. *Means for time on task (in seconds) for the separable and configural displays in Experiments 1 and 2.*

et al. (1985, Experiments 1 and 4), and Barnett and Wickens (1988) have used this technique to assess performance in tasks that require selective or focused attention to individual variables.

The rationale behind the use of this methodology is as follows. In order to complete retrospective memory probes, the subject must extract information from the displays, represent information internally, recall information from memory, and generate a response. Thus performance on a memory-probe task has the potential to reveal the availability of information in a graphic display. One interpretation of differences in performance between two displays is that one format has presented the information in a manner that is more compatible with observers' perceptual and cognitive capabilities than has the other. In the present context, differences in performance could be interpreted as an indication that there are differences in the salience of the emergent features produced by alternative graphic displays. If the memory probes correspond to critical information in the domain, then differences in performance also indicate more-effective or less-effective mappings between domain and displays. Thus improved performance at memory-probe tasks might also be interpreted as an indication that the fundamental requirements for more complex tasks (system control, detection, or diagnosis) have been fulfilled. Whether or not improved performance is actually evident in these high-level tasks depends on many other factors.

In Experiment 2 we used a modified retrospective memory-probe technique that assessed memory for both high-level constraints and low-level data. There were two types of memory probe tasks: focused and integrated. In the focused memory-probe task subjects were asked to recall the value of one of the four variables relevant to the MCF task (SF, FF, ISGL, or CSGL). In the integration memory-probe task the subjects were probed on critical differences between these variables: the difference between SF and FF (mass balance) and the difference between ISGL/CSGL (energy balance). It is also possible that the effectiveness of the emergent features produced by the two displays might vary as the relationships between the critical variables change. To check for this possibility, an additional factor, *plant state,* was included. All memory probes were administered in one of two system states: when the difference between SF and FF or ISGL and CSGL was small (less than 5%) or when the difference was large (greater than 15%).

As previously mentioned, comparisons of configural displays with more separable displays have yielded mixed results when observers were required to focus on individual variables. Pomerantz's theory (1986) suggests that one way to improve the availability of low-level data is to make the graphical elements more salient perceptually. However, the vast majority of studies investigating the issue have not done so (Bennett and Flach, 1992). In the design of both displays two techniques were used to increase the salience of low-level graphical elements: color-coding and maintaining and emphasizing scale. Color-coding the low-level graphical elements makes them more salient perceptually and therefore should make the underlying information more accessible. In addition, it has been shown that color and shape do not configure to produce emergent features (Carswell, 1988; Treisman and Gelade, 1980). Thus color-coding the graphical elements is unlikely to destroy the emergent features, and therefore it is unlikely to disrupt performance on integration tasks.

When maintaining and emphasizing scale, the design goal is to provide a context that increases the salience of the data itself (relative to nondata elements of the graph), the relationship between a datum and other data, or the relation between a datum and the potential values that it might assume. Both Cleveland (1985) and Tufte (1983) provided recommendations for the design of frames, labels, coordinate grids, and other graph elements that can be used to achieve these goals. In both the separable and the configural displays explicit coordinate grids were provided. Although they were not manipulated as an independent variable, we were interested in determining whether these two design techniques might facilitate the extraction of low-level data from configural formats.

EXPERIMENT 2

Method

The subjects, apparatus, simulation model, system dynamics, and stimuli were identical to those used in Experiment 1.

Procedure. The procedure was essentially the same as in Experiment 1, with the only changes being those necessary to implement the memory probes. All subjects were tested on the five consecutive days immediately following Experiment 1. On the first day the subjects were informed of the memory probes and were instructed to respond as quickly and as accurately as possible. The subjects performed the MCF task, as in Experiment 1, and were interrupted to complete memory probes. During a memory probe the simulation was stopped, the screen was blanked, a probe was presented, and the subjects responded by entering a numeric value via the keyboard. The simulation was then restarted in the previously existing system state, including the appropriate adjustments for simulation time constants. Response time was measured from the time that the screen was blanked until the first digit of the subject's response was entered, and it was measured with 1/100-s accuracy. Accuracy (error magnitude) was measured by computing the absolute value of the difference between the subject's estimate and the actual value (as it appeared on the subject's screen—that is, the value of a simulation variable plus the random noise for that particular screen update). Feedback on accuracy was provided before the simulation was restarted.

There were two different types of memory-probe tasks: a focused task and an integration task. In the focused task subjects were asked to estimate either (1) the rate of steam flow, (2) the rate of feedwater flow, (3) the indicated steam generator level, or (4) the compensated steam generator level. In the integration task the subjects were asked to estimate either (1) the difference between steam and feedwater flow or (2) the difference between indicated and compensated steam generator level. Memory probes could also occur during two different system states: (1) when the differences between SF/FF or ISGL/CSGL were large (greater than 15%) and (2) when these differences were small (less than 5%). Approximately 128 probes were obtained in an experimental session. An algorithm was developed to ensure that approximately equal numbers and distributions of probes were obtained in an experimental session. The 64 probes for the focused task consisted of eight probes for each of four categories (SF, FF, ISGL, and CSGL) in both states (large and small differences). For the integration task the 64 probes consisted of 16 probes for each of two categories (differences between SF and FF and differences between ISGL and CSGL) in both states (large and small differences). In summary, the experimental design contained four independent variables: display (separable vs. configural, between-subjects), day (1–5, within-subjects), task (focused vs. integration, within-subjects) and state (large vs. small differences, within-subjects).

Results

Accuracy. All memory probe scores in an experimental session were averaged across probe categories (Task × State) for a total of four scores for each individual. A 2 × 5 × 2 × 2 mixed ANOVA was performed on these data. The main effect of display was significant, $F(1,18) = 4.40, p < 0.05$, indicating that performance with the configural display (average error magnitude = 4.13%) was more accurate than performance with the separable display (5.21%). The main effect for task was significant, $F(1,18) = 18.75, p < 0.0007$, indicating that performance in the integration task (4.12%) was significantly better than performance in the focused task (5.22%). A significant Display × Task interaction, $F(1,18) = 16.41, p < 0.002$, indicated that the configural display facilitated performance in the integration task (see Figure 5a). *F* tests for simple effects indicated that the differences between displays were not significant for the focused task, $F(1,18) = 0.01$, but highly significant for the integration task, $F(1,18) = 34.33, p < 0.00007$. As Figure 5a illustrates, the configural display facilitated the accuracy of performance for the integration memory task. There was also a significant main effect for state and a significant State × Task interaction, but all other effects were not significant.

Latency. All memory probe scores for an experimental session were averaged across probe categories for a total of four scores (Task × State) for each individual. A 2 × 5 × 2 × 2 mixed ANOVA was performed on these data. The main effects of day, $F(4,72) = 42.55, p < 0.000001$, state, $F(1,18) = 60.45, p < 0.000009$, and task, $F(1,18) = 69.31, p < 0.000006$, and the interaction effects of State × Task, $F(1,18) = 50.72, p < 0.00002$, Day × Task, $F(4,72) = 3.89, p < 0.007$, Display × Day, $F(4,72) = 3.32, p < 0.02$, Display × Task, $F(1,18) = 13.25, p < 0.003$, Display × Task × State, $F(1,18) = 4.51, p < 0.05$, and Display × Task × State × Day, $F(4,72) = 2.54, p < 0.05$, were significant. All other effects were not significant. Because of the large number of significant effects, only those directly related to the primary effect of interest, the Display × Task interaction, will be discussed.

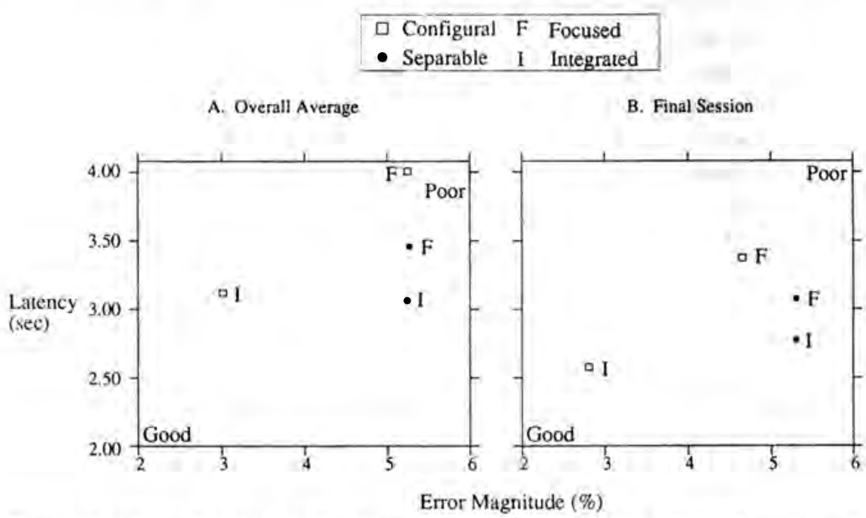


Figure 5. *Means for the Display (configural and separable) × Task (focused and integration) interactions for both latency (in seconds) and accuracy (error magnitude) in Experiment 2. (A) illustrates the averages for all five experimental sessions; (B) illustrates the averages for the final session.*

Because the significant four-way interaction effect constrains the interpretation of lower-level interaction effects, it will be described in detail.

The Display × Task × State × Day interaction indicated that the nature of the interaction between display and task was dependent on both the size of the differences between variables and experience at the task (see Figure 6). In the interest of brevity, only the results for the Days 1 and 5 (i.e., the first and last panels in both graphs) will be discussed. For large differences (bottom graph of Figure 6), F tests for the simple interaction effects

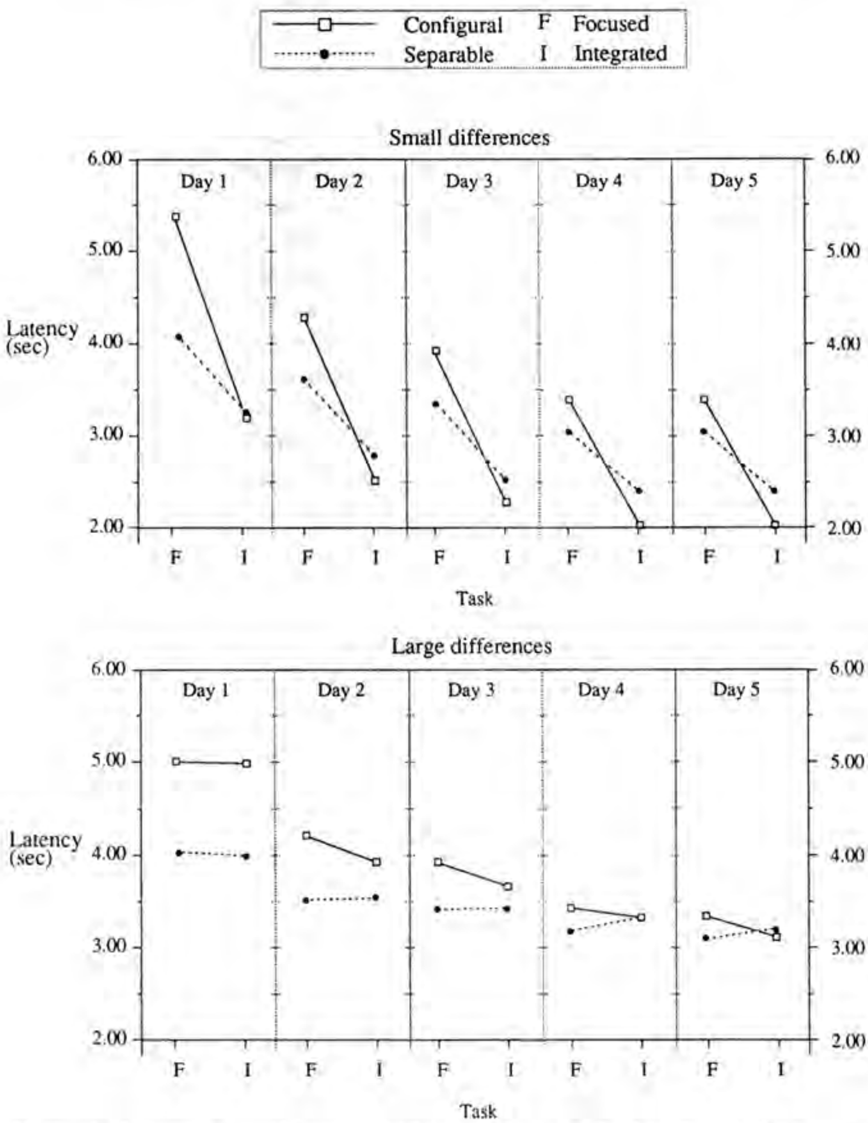


Figure 6. *Latency means (in seconds) for the Display × Task × State × Day interaction effect in Experiment 2. (The interaction between display and task is represented across each state and day.)*

between display and task were not significant for Day 1 or Day 5. Additional F tests directly comparing performance between the two display formats revealed that on Day 1 the latency of responses was significantly lower for the separable display for both the integrated task, $F(1,72) = 51.69$, $p < 0.000001$, and the focused task, $F(1,72) = 40.84$, $p < 0.000001$. However, by Day 5 the differences between the two displays were not significant for either task.

For small differences (top graph in Figure 6), F tests for the simple interaction effects between display and task were significant for both Day 1, $F(1,72) = 47.35$, $p < 0.000002$, and Day 5, $F(1,72) = 10.80$, $p < 0.002$. Additional F tests to directly compare performance indicated that on Day 1 the latency of responses was significantly slower for the configural display with the focused task, $F(1,72) = 89.14$, $p < 0.000001$, but that performance differences for the integration task were not significant. On Day 5 the latency of responses was significantly slower for the configural display with the focused task, $F(1,72) = 7.22$, $p < 0.009$, but was marginally faster for the configural display with the integrated task $F(1,72) = 3.84$, $p < 0.051$.

Time on task. As in Experiment 1, time on task was recorded for each experimental trial. The experimental design contained four independent variables: display (separable vs. configural, a between-subjects variable), day (1–5, a within-subjects variable), starting position of ISGL (35%, 45%, 55%, and 65%, a within-subjects variable) and steam flow ramp type (rising, null, and falling, a within-subjects variable). A $2 \times 5 \times 4 \times 3$ mixed ANOVA was performed on the regular time-on-task scores. The main effects of starting position and ramp were significant, as were the interactions between ramp and starting position and between day by starting position. All other effects, including the predicted main effect of display, $F(1,18) = 1.33$, $p < 0.26$, were not significant.

Discussion

The critical factor in designing configural displays to support the extraction of high-level constraints is the mapping between the inherent data relationships in the domain and the emergent features produced by a display. For the MCF task two critical, high-level constraints are mass balance and energy balance. In the separable display this information was mapped onto the emergent features of the inferred linear relationship between the tops of bars. In the configural display this information was mapped into the emergent features of the height and the width of the rectangle. The integration memory probe tested how effective these alternative emergent features were in conveying this critical domain information.

The results indicate that the emergent features produced by the configural display were more effective than those produced by the separable display. The accuracy of memory for mass and energy balances was significantly better with the configural display than with the separable display. Figure 5a represents the Display × Task interaction (for both accuracy and latency) averaged across the five experimental sessions. As this figure indicates, the significant increase in accuracy performance for the configural display was obtained with no cost in latency. Figure 5b illustrates these interactions for the final experimental session. (Please note that the Display × Task × Day interaction effect for accuracy was not significant—the graph is provided for illustrative purposes.) On Day 5 observers with the configural display were responding to integrated memory probes faster and with significantly more accuracy than were observers with the separable display.

The second design goal for configural

displays is to allow the effective extraction of low-level data. Previous research has revealed a potential cost when low-level data must be extracted from a configural format. However, the possibility that this cost might be offset by raising the perceptual salience of the graphical elements has received little attention (Bennett and Flach, 1992). In both displays the salience of the graphical elements were increased through color-coding and the inclusion of scales.

In the focused probe task, in which individuals were tested for their memory of individual variables, there was no significant difference in the accuracy of responses between the two graphic displays (see Figure 5a). This level of accuracy performance appears to have been maintained at a cost in latency for the configural display. However, as Figure 5b indicates, the pattern of performance changed across experimental sessions. On Day 5 observers with the configural display were performing more accurately (though not significantly so) while the cost in latency became smaller.

The significant four-way interaction for latency confirms that this cost was dependent on both experience at the task and system state. In both system states the cost for focused memory probes was quite large initially but became smaller with additional experience at the task (see Figure 6). With large mass or energy imbalances, this cost was not significant by Day 5; when mass and energy were balanced, the cost was much smaller on Day 5 than on Day 1 but still significant. Thus the results with the retrospective memory-probe technique provide some evidence that color-coding graphical elements and maintaining and emphasizing scale can be used to alleviate the potential costs associated with the configural display of low-level data.

The results of the integrated memory probe indicate that the configural display produced more effective emergent features than did the separable display. In addition, the integrated memory probes measured functional constraints that are critical to successful performance of the MCF task (Roth and Woods, 1988). Considered together, these factors suggest that the configural display provided the basis for improved performance at the control task. However, as in Experiment 1, the time-on-task measure did not reveal statistically significant differences.

It is possible that performance differences did exist but that the time-on-task measure was not sufficiently fine-grained to measure them. The results of Experiment 2 revealed a consistent advantage in time-on-task performance for the configural display (an increase of approximately 22 s averaged across days; see Figure 4). Thus the lack of statistical significance would appear to be the result of large subject variability in performance, and it is possible that the variability is inherent to the time-on-task measure. For example, an observer who changes an upward trend in ISGL and narrowly avoids the upper trip set point could add, regardless of input, an additional 1–2 min of time on task because of the system dynamics. The time-on-task measure was adopted because it is the most ecologically valid measure of performance (i.e., it is the only performance criterion that counts in the real world), but other measures of control performance may prove to be more effective in revealing performance differences (e.g., time to target, control reversals).

An alternative (perhaps complementary) explanation of why time on task did not reveal significant performance differences is that the observers had not fully developed the skills and knowledge that are required to successfully complete the task. The manual control of feedwater is a difficult control task, and the scaled-world simulation used in the present experiments captured this complexity.

Developing an appropriate internal model of the system dynamics, effective response strategies, and contextualized knowledge of the appropriate time to apply them is likely to require more than 10 hours of experience at the task. As Flach and Vicente have noted,

> Just because the interface is designed in such a way that control via direct perception and direct manipulation is possible does not necessarily guarantee that operators will indeed exhibit such skilled behaviors. For instance, an operator who is just learning to use a direct perception interface may initially rely on higher levels of cognitive control. As Gibson [1979] points out, the fact that the affordances exist and are perceivable (through the interface, in this case) does not ensure that the actor will pick up that information. However, it does make it possible to pick up that information. (1989, p. 33)

GENERAL DISCUSSION

The graphic display of information is a form of decision support that provides the very real potential to improve the overall performance of human-machine systems in complex, dynamic task domains. This vast potential arises, in part, because of continuing advances in hardware and software technology. Nevertheless a better understanding of how these technological capabilities can be used effectively is needed. Previous research has discounted the role that the semantics of the domain must assume in the design of effective decision support for complex, dynamic domains. A perspective is needed that considers the cognitive triad: the cognitive demands of the domain, the cognitive resources of the agents meeting these demands, and the interface (graphic displays) considered as a representation of the domain that supports demand/resource mismatches (Flach and Vicente, 1989; Vicente and Rasmussen, 1990; Woods and Roth, 1988b).

To accomplish tasks in complex domains, the operator must consider the system from different levels of abstraction and alternate between these levels (Rasmussen, 1986). Thus information about high-level constraints that are defined by the relationships between variables needs to be available, as does the low-level information about individual variables that contribute to them. For configural displays the critical design issue is to provide a representation that allows the parallel extraction of both types of information in the context of critical performance boundaries.

Designing Configural Displays for the Extraction of High-Level Constraints

The results of the present experiment add to accumulating evidence that the capability of a display to improve performance at integration tasks depends on how well the semantics of a domain have been mapped onto the form and dynamic behavior of a display (Bennett and Flach, 1992). In particular, the results indicate that if the low-level graphical elements configure to produce emergent features that correspond to the critical relationships between variables in the domain, then performance at integration tasks will be improved. Two separate activities must be completed to achieve this goal, and failure at either will reduce the effectiveness of the resulting display. First, the semantics of the domain must be determined. Second, a display must be designed that produces emergent features that directly reflect the domain semantics. Each of these activities will be discussed in greater detail.

Discovering the domain semantics. In complex, dynamic domains discovering the domain semantics is not an easy task. Researchers in cognitive engineering (Hollnagel and Woods, 1983; Norman, 1986; Rasmussen, 1986; Woods, 1991a; Woods and Hollnagel, 1987) have developed disciplined, top-down approaches to identify the semantics of a domain. The "goal/means" hierarchy (Rasmussen, 1986) is a description of domain semantics organized in five separate levels of abstraction, ranging from the physical form of a system (e.g., What are the system components? What do they look like? Where are they located?) to the high-level purposes it serves (e.g., What is the system's purpose? What constraints does the system operate under to fulfill this purpose?). It describes the high-level constraints of interest, the low-level data that are relevant to those constraints, the relationships between these low-level data, and the relevant goals and constraints. The semantics of the domain must then be mapped onto the static form and dynamic behavior of a display.

Mapping domain semantics into emergent features. Designing a configural display to facilitate performance with respect to high-level issues involves more than integrating variables into an object. The key to designing a successful configural display is to integrate the variables into a graphic form that produces emergent features that highlight the critical data relationships in the domain. This is not an easy task. Even with simple visual stimuli, the relationship between task and emergent features is difficult to define (Pomerantz, 1986); in complex, dynamic domains the relationships will be more complex (e.g., compare the results of Sanderson et al., 1989, to those of Carswell and Wickens, 1987).

As an example of some of the problems that might be encountered consider the graphic form of the human face (Chernoff, 1973). Because of the nature of the face, mapping from data to facial features is complex and arbitrary. Kleiner and Hartigan (1981, and the commentary by Howard Wainer that follows that paper) provide an example in which the same set of data was mapped into three different versions of a facial display. Each of the three versions substantially changed the resulting perceptual cues. A second problem is that the individual features of the face are not equal in perceptual salience (Brown, 1985; MacGregor and Slovic, 1986; Naveh-Benjamin and Pachella, 1982), which causes the information presented on some features to dominate the information presented on other, less salient features.

The problem is not eliminated by choosing a geometric object format. Buttigieg (1989) illustrated that the same graphical format can produce qualitatively different emergent features. Some relevant questions are: Which variables need to be included in the graphic form? How should the individual elements be assigned to the dimensions of the object—that is, does the relationship between two variables suggest that they should be juxtaposed or that they should be adjacent? Should all variables be converted to a common scale? How can the decision-making context (relevant goals and constraints) be represented? The answers to these questions must be based on the results of a cognitive task analysis and an analysis of the domain itself. The rationale provided for the development of the configural display is one example of how this information can be used to develop effective configural displays.

Designing Configural Displays for the Extraction of Low-Level Data

It is often assumed that there is an inherent cost to low-level data when this information is presented in a configural format. A review of the relevant literature has indicated that there is a potential cost but that it appears to be much less than anticipated (Bennett and Flach, 1992). The pattern of experimental results is consistent with predictions derived from Pomerantz's theory of form perception (1986). This theory maintains that low-level data are not "lost in" or "glued to" the high-level form but are simply less salient perceptually. Thus this potential cost may be alleviated by increasing the perceptual salience of these graphical elements. It should be emphasized that the vast majority of experimental studies that have investigated this cost have highlighted high-level constraints (through inclusion in configural formats) but have not emphasized the contributions of low-level data. Four possible techniques will be described.

Color-coding. As previously mentioned, color-coding the graphical elements will increase the salience of low-level data and is not likely to interfere with the high-level emergent features that are produced by their interaction. Wickens and Andre (1990) investigated these issues, comparing performance with a monochrome and color-coded configural display. They found that color-coding produced a speed-accuracy trade-off: color-coding the graphical elements facilitated accuracy at the focused task and did not disrupt accuracy at the integration task. However, the improvement in accuracy performance was associated with a decrement in latency performance: the chromatic version of the configural display increased the latencies associated with both types of tasks. Wickens and Andre (1990) also included a monochrome separable display in their study, but performance comparisons between this display and the chromatic configural display were not made.

Maintaining and emphasizing scale. The consequences of failing to maintain scale are most evident in the face display (e.g., MacGregor and Slovic, 1986). When low-level data are mapped onto the graphical elements of a face, all traces of their contributions disappear: data that were originally ordinal is mapped into nominal facial features. This results primarily from the fact that no scale is available to provide a context in which to evaluate individual values. It should be noted that with rare exceptions (e.g., the present study), scales and coordinate grids have not been included in experimental comparisons of graphic formats. Dolan, Elvers, and Schmidt (1991) included two aspects of scale (the presence or absence of tick marks and zero reference points) as independent variables in their study. They found that reaction time performance for a focused task was facilitated by the presence of scale but that performance at an integration task suffered.

Spatial separation. Another method to increase the salience of graphical elements is spatial separation, which can be thought of as a design-space continuum. The "contribution graphic" described by Woods and Roth (1988b) is at the high end of the spectrum. In this type of display each graphical element is mapped onto a spatially separated bar graph, whereas the contributions are mapped into an integrated (overlapped or tiled) bar graph. Spatially separating the elements both suppresses irrelevant emergent features and emphasizes the contributions of low-level data. An intermediate level of spatial separation is used in the configural display described by Wickens and Andre (1988, 1990). In that display the contributions of low-level data were independent of the high-level object that they defined. The configural display in the present experiment is located at the low end of the spatial separation continuum: the graphical elements directly define the object. The research conducted on this issue has revealed mixed results and indicates that the role of spatial separation for graphic displays in general may be complicated (Cleveland, 1985; Wickens and Andre, 1988, 1990).

Digital values for low-level data. One final technique to be considered is the inclusion of digital values for low-level data. If the value of an individual variable is critical, then a digital display would provide greater precision than would an analog representation. The dual design goals may be achieved by combining (1) a configural format that produces emergent features corresponding to high-level constraints and (2) digital values that provide precise information about low-level data. Although digital values were not provided in the present experiments, we have developed versions of the configural display that incorporate digital values in the placeholders outside the coordinate grids (see Figure 2). One criticism of this arrangement is that placing the values outside the display grid might also increase visual search times. Including the digital value inside the display grid would reduce search times and might increase configurality. For example, in the configural display used in the present experiment the digital values could be placed outside the rectangular form on the corresponding side. The digital value would be fixed in this location with respect to the rectangle but free to move in the coordinate grid along with the rectangle. The results of Hansen (in press) support the potential of this design technique. He found that incorporating digital values in this fashion facilitated the latency of detection for trends across variables without a cost in accuracy.

The theoretical perspective of configurality and the results of the experiments that have been conducted to date both indicate that these four techniques can be used to offset the potential costs associated with configural formats. The empirical results, however, also indicate that the trade-offs associated with their use need to be examined in greater detail.

SUMMARY

To support performance in complex, dynamic domains graphic displays must be designed to allow the extraction of both high-level constraints and low-level data. Designing graphic displays that allow the efficient extraction of information at both levels would be an economical and elegant solution, and it appears that configural displays have the potential to achieve both of these goals. The results of the present experiment, and the literature in general, suggest that configural displays are likely to support the extraction of high-level constraints, when properly designed. The critical design consideration is to provide a direct mapping between the inherent data relationships in a domain and the emergent features that are produced by a graphic format. For the extraction of low-level data, the critical design consideration is to raise the perceptual salience of the graphical elements in a display relative to its emergent features. Basic theories of object perception and attention suggest that this information is available, and the pattern of experimental results does not strongly support the existence of an inherent cost (even though these experiments were biased toward finding one). Perhaps the most important observation is that display design and evaluation do not occur in a vacuum. Research on graphic forms and processing characteristics is incomplete without consideration of how these factors interact with, and are mutually constrained by, the characteristics of the domain.

ACKNOWLEDGMENTS

The authors thank John Flach for discussions and comments on earlier drafts and two anonymous reviewers for their helpful comments. Funding was provided by Wright State University (Research Challenge and Research Incentive Grants) and by the U.S. Air Force Office of Scientific Research (Summer Faculty Research Program and Research Initiation Program). The part-task simulation was originally developed at Westinghouse Electric Corporation, primarily in the department of Human Sciences Research, with the assistance of numerous individuals and organizations.

REFERENCES

Barnett, B. J., and Wickens, C. D. (1988). Display proximity in multicue information integration: The benefits of boxes. *Human Factors, 30,* 15–24.

Bennett, K. B., and Flach, J. M. (1992). Graphical displays: Implications for divided attention, focused attention, and problem solving. *Human Factors, 34,* 513–533.

Bennett, K. B., Woods, D. D., Roth, E. M., and Haley, P. H. (1986). Predictor displays for complex, dynamic tasks: A preliminary investigation. In *Proceedings of the Human Factors Society 30th Annual Meeting* (pp. 684–688). Santa Monica, CA: Human Factors Society.

Brown, R. L. (1985). Methods for graphic representation of systems of simulated data. *Ergonomics, 28,* 1439–1454.

Buttigieg, M. A. (1989). *Emergent features in visual display design for two types of failure detection tasks* (Tech. Report EPRL-89-04). Urbana-Champaign: University of

Illinois, Engineering Psychology Research Laboratory and Aviation Research Laboratory.

Carswell, C. M. (1988). Integral, configural, and unitary graphs. In *Proceedings of the Human Factors Society 32nd Annual Meeting* (pp. 1345–1349). Santa Monica, CA: Human Factors Society.

Carswell, C. M., and Wickens, C. D. (1987). Information integration and the object display. *Ergonomics, 30*, 511–527.

Casey, E. J., and Wickens, C. D. (1986). *Visual display representation of multidimensional systems: The effect of information correlation and display integrality* (Tech. Report CPL-86-2). Urbana-Champaign: University of Illinois, Cognitive Psychophysiology Laboratory.

Chernoff, H. (1973). The use of faces to represent points in *k*-dimensional space graphically. *Journal of the American Statistical Association, 68*, 361–368.

Cleveland, W. S. (1985). *The elements of graphing data*. Belmont, CA: Wadsworth.

Coury, B. G., Boulette, M. D., and Smith, R. A. (1989). Effect of uncertainty and diagnosticity on classification of multidimensional data with integral and separable displays of system status. *Human Factors, 31*(5), 551–570.

Crossman, E. R. F. W., and Cooke, J. E. (1974). Manual control of slow-response systems. In E. Edwards and F. P. Lees (Eds.), *The human operator in process control* (pp. 51–66). London: Taylor & Francis.

Dolan, N. J., Elvers, G. C., and Schmidt, J. (1991). Measurement scale and zero point indicators affect object display separate and integral task performance. In *Proceedings of the Human Factors Society 35th Annual Meeting* (pp. 1589–1592). Santa Monica, CA: Human Factors Society.

Flach, J. M., and Vicente, K. J. (1989). *Complexity, difficulty, direct manipulation and direct perception* (Tech. Report EPRL-89-03). Urbana-Champaign: University of Illinois, Engineering Psychology Research Laboratory.

Gibson, J. J. (1979). The ecological approach to visual perception. Boston: Houghton-Mifflin.

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: Erlbaum.

Haley, P. H., and Woods, D. D. (1988). Methods and apparatus for dynamic systems control. Patent Number 4,770,841, Patent Date September 13, 1988.

Hansen, J. P. (in press). Representation of system invariants by optical invariants in configural displays in process control. In J. M. Flach, P. A. Hancock, J. K. Caird, and K. J. Vicente (Eds.). *An ecological approach to human-machine systems: II. Local applications*. Hillsdale, NJ: Erlbaum.

Hollnagel, E., and Woods, D. D. (1983). Cognitive systems engineering: New wine in new bottles. *International Journal of Man-Machine Studies, 18*, 583–600.

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

Kleiner, B., and Hartigan, J. A. (1981). Representing points in many dimensions by trees and castles. *Journal of the American Statistical Association, 76*, 260–269.

MacGregor, D., and Slovic, P. (1986). Graphic representation of judgmental information. *Human-Computer Interaction, 2*, 179–200.

Moray, N. (1987). Intelligent aids, mental models, and the theory of machines. *International Journal of Man-Machines Studies, 27*, 619–629.

Moray, N., Lootsteen, P., and Pajak, J. (1986). Acquisition of process control skills. *IEEE Transactions on Systems, Man, and Cybernetics, SMC-16*, 497–504.

Naveh-Benjamin, M., and Pachella, R. G. (1982). The effect of complexity on interpreting "Chernoff" faces. *Human Factors, 24*, 11–18.

Norman, D. A. (1986). Cognitive engineering. In D. A. Norman and S. W. Draper (Eds.), *User-centered system design* (pp. 31–61). Hillsdale, NJ: Erlbaum.

Pomerantz, J. R. (1986). Visual form perception: An overview. In H. C. Nusbaum and E. C. Schwab (Eds.), *Pattern recognition by humans and machines: Vol. 2. Visual perception* (pp. 1–30). Orlando, FL: Academic.

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

Roth, E. M., and Woods, D. D. (1988). Aiding human performance: I. Cognitive analysis. *Le Travail Humain, 51*(1), 39–64.

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

Treisman, A. M., and Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology, 12*, 97–136.

Tufte, E. R. (1983). *The visual display of quantitative information*. Cheshire, CT: Graphics.

Vicente, K. J., and Rasmussen, J. (1990). The ecology of human-machine systems: II. Mediating "direct perception" in complex work domains. *Ecological Psychology, 2*(3), 207–249.

Wickens, C. D., and Andre, A. D. (1988). Proximity compatibility and the object display. In *Proceedings of the Human Factors Society 32nd Annual Meeting* (pp. 1335–1339). Santa Monica, CA: Human Factors Society.

Wickens, C. D., and Andre, A. D. (1990). Proximity compatibility and information display: Effects of color, space, and objectness on information integration. *Human Factors, 32*, 61–78.

Wickens, C. D., Kramer, A., Barnett, B., Carswell, M., Fracker, L., Goettl, B., and Harwood, K. (1985). *Display-cognitive interface: The effect of information integration requirements on display formatting for C3 displays* (Tech. Report EPL-85-3). Urbana-Champaign: University of Illinois, Engineering Psychology Research Laboratory and Aviation Research Laboratory.

Woods, D. D. (1991a). The cognitive engineering of problem representations. In G. R. S. Weir and J. L. Alty (Eds.), *Human-computer interaction and complex systems* (pp. 169–188). London: Academic.

Woods, D. D. (1991b). Representation aiding: A ten-year retrospective. In *Proceedings of the 1991 IEEE International Conference on Systems, Man, and Cybernetics* (pp. 1173–1176). Charlottesville, VA: IEEE.

Woods, D. D., and Hollnagel, E. (1987). Mapping cognitive demands in complex problem-solving worlds. *International Journal of Man-Machine Studies, 26*, 257–275.

Woods, D. D., and Roth, E. M. (1988a). Aiding human performance: II. From cognitive analysis to support systems. *Le Travail Humain, 51*(2), 139–171.

Woods, D. D., and Roth, E. M. (1988b). Cognitive systems engineering. In M. Helander (Ed.), *Handbook of human-computer interaction* (pp. 1–41). Amsterdam: Elsevier.
Woods, D. D., Wise, J. A., and Hanes, L. F. (1981). An evaluation of nuclear power plant safety parameter display systems. In *Proceedings of the Human Factors Society 25th Annual Meeting* (pp. 110–114). Santa Monica, CA: Human Factors Society.
Zachary, W. (1986). A cognitively based functional taxonomy of decision support techniques. *Human-Computer Interaction, 2*, 25–63.

APPENDIX

This appendix describes the process simulation programmed for the experiments.

Parameter	:	Description
$y_p(t)$	:	indicated level
$y_c(t)$	:	compensated level
$u(t)$	:	feedwater flow (subject input)
$w(t)$	:	steam flow
$w_d(t)$	:	disturbance in steam flow (input)
$w_0(t)$	:	continuous variation in steam flow (input)
$q(t)$	:	excess heat input rate (constant)
$m(t)$	:	modeling error parameter (set to 0)
$n(t)$	:	noise parameter (set to 0)
T_0	=	40
T_1	=	24*
T_2	=	36*
T_3	=	24*
T_4	=	15
T_5	=	20
Δt	:	simulation time step

* The values of T_1, T_2, and T_3 model the effect of variations in feedwater temperature.

Coefficients computed between time steps:

$$E = e^{-\frac{\Delta t}{T_1}} F = e^{-\frac{\Delta t}{T_4}} G = e^{-\frac{\Delta t}{T_5}}$$

Given initial input $y_p(0)$, $w(0)$, then:

$$u(0) \leftarrow w(0)$$

$$x_0(0) \leftarrow y_p(0) - x_1(0)$$

$$x_1(0) \leftarrow \frac{T_3 - T_2}{T_0} \cdot w(0)$$

$$x_2(0) \leftarrow 0$$

$$x_3(0) \leftarrow 0$$

Time step equations:

$$w(k) \leftarrow x_3(k) + w_0(k) + w_d(k)$$

$$x_0(k+1) \leftarrow x_0(k) + \frac{\Delta t}{T_0}(u(k) - w(k))$$

$$x_1(k+1) \leftarrow E \cdot x_1(k) + (1 - E) - \left(\frac{T_1 + T_3}{T_0} \cdot w(k) - \frac{T_1 + T_2}{T_0} \cdot u(k)\right)$$

$$x_2(k+1) \leftarrow F \cdot x_2(k) + (1 - F) \cdot m(k) - \left(\frac{T_4 + T_3}{T_0} \cdot n(k)\right)$$

$$x_3(k+1) \leftarrow G \cdot x_3(k) + (1 - G) \cdot q(k)$$

$$y_p(k+1) \leftarrow x_0(k+1) + x_1(k+1)$$

$$y_c(k+1) \leftarrow x_0(k+1) + \frac{T_1}{T_1 - T_4} \cdot x_1(k+1) + x_2(k+1) + \frac{T_3 - T_2}{T_0} \cdot w(k)$$