Evaluating Two Aspects of Direct Manipulation in Advanced Cockpits

James A. Ballas, Constance L. Heitmeyer and Manuel A. Pérez

Human-Computer Interaction Lab
Naval Research Laboratory
Washington, D. C. 20375-5000
(202) 767-2774, ballas@itd.nrl.navy.mil

ABSTRACT

Increasing use of automation in computer systems, such as advanced cockpits, presents special challenges in the design of user interfaces. The challenge is particularly difficult when automation is intermittent because the interface must support smooth transitions from automated to manual mode. A theory of direct manipulation predicts that this interface style will smooth the transition. Interfaces were designed to test the prediction and to evaluate two aspects of direct manipulation, semantic distance and engagement. Empirical results supported the theoretical prediction and also showed that direct engagement can have some adverse effects on another concurrent manual task. Generalizations of our results to other complex systems are presented.

KEYWORDS: Direct manipulation, interface styles, interface design, adaptive automation, intermittent automation, aircraft interfaces, intelligent cockpit.

INTRODUCTION

The rapid evolution of user interfaces is occurring not only in office systems but also in modern cockpits, which are computer-based and include advanced graphical displays [14]. However, modern cockpits differ from traditional office systems in several fundamental ways. First, unlike office systems, they often include sophisticated automation, such as the ability to fly on automatic pilot. Moreover, unlike office applications, the cockpit application is dynamic and complex. The pilot must not only handle large quantities of real-time, often continuous, input data; he must also perform several demanding tasks concurrently, usually under severe timing constraints. Finally, unlike users of office systems who typically communicate via electronic mail, the pilot of a modern cockpit communicates in real-time via networked voice and data links. Given these differences, the cockpit interface presents many design challenges that the developers of office systems seldom encounter.

1992 ACM 0-89791-513-5/92/0005-0127

An important question in designing the user interface of modern cockpits is how to handle automation. Our research is part of a larger research program in adaptive automation whose role is to allocate tasks between the pilot and the computer system in an optimal manner [10]. In adaptive (i.e., intermittent) automation, the pilot performs a task only intermittently. Given a dual task situation, a rise in the level of difficulty of one task causes automation of the second task. Having the computer system take over the second task allows the pilot to focus his efforts on the increased difficulty task. Once the difficulty level of the first task returns to normal, the pilot resumes control of both tasks. Such an approach to automation is expected to result in better overall pilot/system performance [10]. Because the pilot only performs the first task intermittently, a challenging problem, and the problem that this paper addresses, is how to design an interface that supports a smooth transition from automated to manual mode.

This paper presents the results of our empirical research on interface styles for adaptive automation. Our research is designed to test predictions from a theory of direct manipulation. A fundamental goal of the research is to determine whether a direct manipulation interface has performance benefits in adaptive automation; i.e., does direct manipulation lead to improved performance when a pilot must quickly resume a task that has been previously automated? A related goal is to separate and evaluate two aspects of direct manipulation identified by the theory, namely, distance and engagement. In this paper, we introduce the direct manipulation theory, present our hypothesis about the effect of interface style in adaptive automation, describe the interfaces developed to test our hypothesis, and summarize the empirical results. We conclude with a discussion of the implications of our results.

BACKGROUND

Designing an interface for an adaptive system involves many issues and decisions, but little theoretical guidance or empirical information is available. There is general agreement on what the interface should accomplish. As a first priority, the interface should enable the pilot to maintain both situational awareness and system control

[10]. We define *situational awareness* as the extent to which the pilot has the knowledge needed to perform a specified task or tasks. Clearly, this knowledge depends upon the specific state of the aircraft and selected aspects of the aircraft environment. In adaptive (i.e., intermittent) automation, the pilot shifts from manually performing a task to monitoring its automated performance and then back to manual operation. In this situation, the key to assessing situational awareness is how well the pilot can resume a task that has been previously automated. We claim that a critical factor in achieving a smooth transition from automated to manual performance of a task is interface style.

Hutchins, Hollan, and Norman (HHN) have developed a theory of direct manipulation [6]. They characterize direct manipulation interfaces according to a model world metaphor; the user interacts with an interface that represents the task domain itself, the domain objects and the effect of user operations on those objects. Command language interfaces behave according to a conversational metaphor; the user and the interface have a conversation about the application domain. The interface acts as an intermediary between the user and the domain. Although typically associated with office computer systems, direct manipulation is also being considered for large safety-critical systems, such as nuclear power plants [2]. HHN concluded that two aspects of direct manipulation account for its performance advantages, low distance and direct engagement. The first aspect is the "information processing distance between the user's intentions and the facilities provided by the machine". Performance advantages come with less distance, because there is less cognitive effort needed to understand and manipulate the domain objects. HHN call such an interface *semantically direct* and claim that it can be achieved by "matching the level of description required by the interface language to the level at which the person thinks of the task".

Distance is of two types, semantic and articulatory. *Semantic* distance is the difference between the user's intentions and the meaning of the expressions available in the interface, both expressions that communicate the user's intentions to the computer and expressions whereby the computer system provides user feedback. For example, if the user wishes to delete all files whose names end in *text* and the computer system (e.g., the Macintosh) has no single expression for this purpose, then significant semantic distance exists between the user's intentions and the expressions available in the interface. *Articulatory* distance is the difference between the physical form of the expressions in the interface and the user's intentions. For example, when a Unix user wants to display a file and to do so he must invoke a command named "cat", significant articulatory distance exists between the name of the Unix command and the intended user operation. Our studies have focused on semantic distance. We have proposed followup studies to investigate issues concerned with articulatory distance.

The second aspect of direct manipulation is *engagement*, i.e., the involvement that comes when the user is able to interact directly with the application domain and the objects within it rather than interacting through an intermediary. The key to direct engagement is interreferential I/O, which permits "an input expression to incorporate or make use of a previous output expression". For example, if a listing of file names are displayed on the screen, one of these names can be selected and operated on without entering the name again. In Draper's view [5], the important aspect of interreferential I/O is that the user and the computer system share a common communications medium. This takes the notion of interreferential I/O beyond the Unix concepts of channels and pipes. In direct manipulation, the shared medium is usually a visual display which presents an explicit, often graphical, view of the task domain.

Related Research on Direct Manipulation

An early study comparing several interfaces concluded that usability depends more on specific interface design than interface style [12]. Contrary to expectations, iconic interfaces were inferior to menu systems and command language interfaces for new and transfer users. More recent studies have generally shown advantages for direct manipulation over command language interfaces [15]. For example, Karat [7] found consistently faster times for several file management tasks in a direct manipulation interface that used pointing and dragging operations on iconic representations of files. However, Karat did find an advantage for the command language interface on one particular type of file management task. Thus, evaluations of interface styles need to be sensitive to task-specific effects. Along this line, Elkerton and Palmiter (cited in Kieras [8]) suggest that the basic principle of direct manipulation lies in the replacement of complex cognitive operations with perceptual and motor activities. Thus the advantage of direct manipulation may lie in tasks with complex cognitive operations that can be transformed into motor and perceptual operations.

Research on direct manipulation has been mostly on conventional applications, such as word processing and file management. A notable exception is a study by Benson and her colleagues [3] which compared a conventional interface to a direct manipulation interface for a parts manufacturing system. The conventional interface used menus, function keys, typed commands, displayed textual information, and paged displays. The direct manipulation interface used a mouse as the only input device and provided a continuous display of important information. The evaluation of these interfaces used performance measures relevant to manufacturing, such as cost, inventory levels and status, and late deliveries. Performance with direct manipulation was superior on three of five dependent measures.

All previous research on direct manipulation has not attempted to tease apart semantic distance and direct engagement, and determine which is important in user performance. Furthermore, previous research has evaluated applications designed for purposes other than evaluation of

an interface style. The interfaces in our study were designed specifically to study direct manipulation by separating and evaluating the two aspects identified in the HHN theory.

Direct Manipulation in the Cockpit

The effectiveness in the cockpit of a direct manipulation interface and its two aspects remains an open question. Some studies suggest that navigation displays should present a model world to the pilot. For example, in 1987, Marshak, Kuperman, Ramsey, and Wilson [9] found that moving-map displays in which the viewpoint is similar to what would actually be seen by looking outside the plane led to improved performance. However, in other kinds of displays, a graphical representation of the model world does not provide an advantage. For example, Reising and Hartsock [11] found that in warning/caution/advisory displays, a schematic of the cockpit showing the controls that were needed to handle an emergency did not improve performance. The important factor in improved performance was a checklist of the required procedures (which is closer to what a command language interface would offer).

Ironically, in modern flight control systems, some trends have been away from direct manipulation. For example, fly-by-wire systems remove the pilot from direct control of wing surfaces. Bernotat [4] argues against this trend, suggesting that, in such systems, the pilot needs direct sensory feedback about the aircraft's performance. Such feedback is consistent with the notion of direct manipulation. Other trends in cockpit controls suggest a move toward direct manipulation, e.g., the incorporation of touchscreen displays. However, the incorporation of pointing devices into the flight deck needs to be carefully evaluated; e.g., what is the effect of the pilot's use of two pointing devices concurrently (a touchscreen and a joystick)?

Direct Manipulation in Intermittent Automation

An issue in interface design for intermittent automation is *automation deficit*, the initial decrease in pilot performance that occurs when a task that has been previously automated is resumed. This deficit may reveal itself in several ways: slower human response, less accurate human response, subjective feelings of not being in control, subjective feelings of stress, etc. Some previous studies have shown an automation deficit for manual control tasks, while others have not [10]. In our research we are interested in automation deficits in response time and the effect of interface style on automation deficit.

Our hypothesis is that direct manipulation interfaces lead to a reduction in automation deficit that is reflected in decreased response times right after automation ceases. The rationale underlying this hypothesis is that decreased semantic distance and improved direct engagement enhance a pilot's ability to monitor a task that is automated and then to quickly resume the task. Besides testing the general hypothesis, we evaluated the importance of each aspect of direct manipulation in minimizing automation deficit.

To test our hypothesis we evaluated the effect of interface styles on a person's ability to resume a task quickly after a period of automation. More specifically, we compared performance in the first few seconds of the manual mode to performance a minute later on, using different types of interfaces.

EXPERIMENTAL DESIGN

Subjects

Twenty subjects were recruited from NRL personnel, with five randomly assigned to each of the four types of interfaces used in the tactical assessment task. All were screened for normal color vision. Two of the subjects were licensed pilots.

General Procedure

The experiment required subjects to perform two tasks, a pursuit tracking task and a tactical assessment task. To establish a setting for adaptive automation, the difficulty of the tracking task alternated between moderate and high throughout the experiment. During the moderate difficulty phases of the tracking task, the subject performed both the tracking task and the tactical assessment task. Each time the difficulty of the tracking task rose to high, the tactical assessment task was automated, and the subject performed the tracking task only. The display screen used in the experiment was partitioned into two nonoverlapping windows, one for the tracking task, the other for the tactical assessment task.

Tracking Task

The tracking task simulated air-to-air targeting of an enemy aircraft using a gunsight similar to the pipper and reticle on a typical head-up display. The target on the display was a graphical representation of an enemy aircraft. The tracking control was a self-centering, displacement joystick. The two levels of tracking difficulty, high and moderate, were produced by changing the movement rate of the target. Performance measures included RMS amplitude calculated for each axis. In addition, the target's and the subject's movements were recorded for later analysis.

Tactical Assessment Task

The second task, tactical assessment, is a critical task in a tactical aircraft and one that has become more challenging with the increased capabilities of modern aircraft. Our hypothesis was tested on interfaces for the tactical assessment task. The simulated tactical situation included three classes of targets - fighters, aircraft, and ground-based missiles - and contacts on the targets by sensor systems. The targets first were designated as possible threats using black color coding, but as they got closer to the *ownship* (the symbol for the aircraft the pilot was in), they were designated as neutral, hostile, or unknown, using blue, red and amber color coding, respectively. The subjects were told that simulated sensor systems were assigning these designations.

The subjects were required to perform two operations, confirm and classify. If the system designated a target as neutral or hostile (i.e., the target was colored blue or red), the subject had to *confirm* the designation by picking the target and then indicating the proper designation, i.e., neutral for blue targets and hostile for red targets. Thus, confirm decisions only required the subject to discriminate colors. If the system designated the target as unknown (i.e., the target was colored amber), the subject had to *classify* the target as hostile or neutral based on its behavior. Table 1 provides the rules for designating a target as hostile or neutral. The target class determines what target attribute the subject uses to determine the target's designation.

Target Class	Hostile	Neutral
Fighter	Constant bearing	Bearing away
Airplane	Air speed ~ 800	Air speed ~ 300
Missile site	Within threat range	Outside threat range

Table 1. Rules for tactical assessment of targets

To classify the amber targets, the subject needed to monitor heading for fighters, speed for aircraft, and projected lateral distance for ground missile threats. The responses were timed and analyzed to produce measures of accuracy and response time.

Training was provided on each task alone and on the dual task without automation. A total of twenty subjects were tested on the intermittent automation, five on each of the four interfaces. Twelve subjects were retested four months later. They received a 3 minute retraining session on both tasks. Further details of the experiment are presented elsewhere [1].

Interfaces for the Tactical Assesment Task

To test our hypothesis, we designed and built four interfaces, using prototyping and iterative development. These four interfaces, which include a direct manipulation interface, a command language interface, and two hybrid interfaces, represent the four combinations of semantic distance and engagement shown in Figure 1. Below, we briefly describe each interface and discuss how each implements some combination of semantic distance and engagement.

The *direct manipulation interface* (Figure 2a) has direct engagement and low semantic distance. It uses a shared communications medium: both the subject and the computer use the entire tactical assessment window to communicate. This interface simulates a radar display with continuously moving symbols representing the targets. The symbol used to represent a target is an intuitive graphical representation of the target class. Each target symbol is initially colored black but changes to red, blue, or amber once the system assigns the target a designation. A touchscreen overlays the display. The subject confirms or classifies a target by picking a target symbol on the display and selecting one of two strips, labeled HOSTILE and NEUTRAL, located on either side of the display. The subject accomplishes both the pick and the select by touching the appropriate part of the display screen. The words 'HOSTILE' and 'NEUTRAL' in the two side strips are colored red and blue, respectively. For classify decisions, the subject needs to observe the behavior of the graphical symbol that represents the target to determine the proper target designation. For confirm decisions, the subject needs to interpret the color of the target symbol.

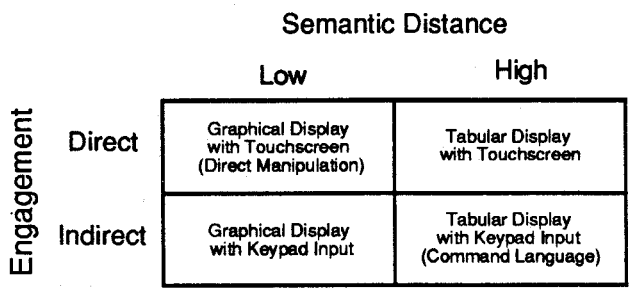

Figure 1. Levels of engagement and semantic distance in the four interfaces for the tactical assessment task.

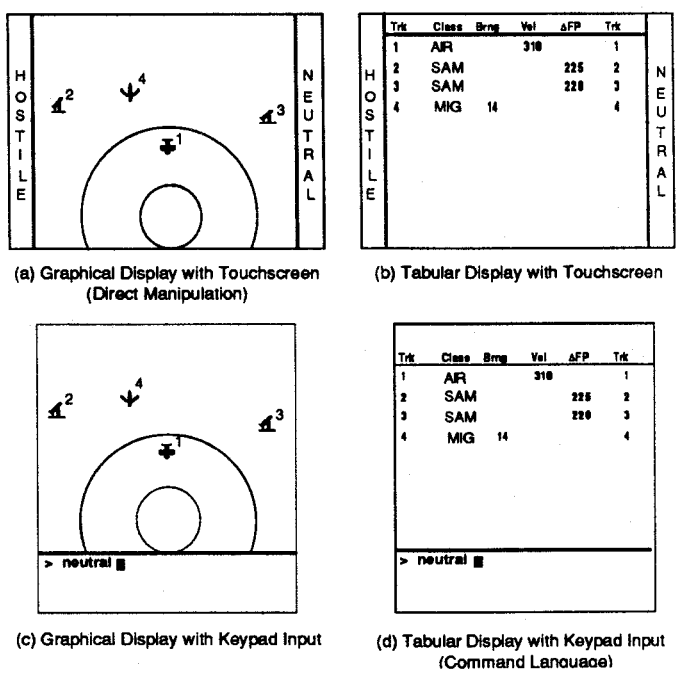

Figure 2. Four interfaces for the tactical assessment task, combining levels of semantic distance and engagement.

The *command language interface* (Figure 2d) has indirect engagement and high semantic distance. This interface uses a split visual medium: the tactical assessment window is partitioned into a top portion, which displays a table of target names and attributes, and a bottom portion, which is for subject input and error feedback. Each entry in the table describes a single target, providing the target's name (an integer), the target's class, and continuously updated data about the target. The name of the target class

carries the system designation; initially black, it changes to red, blue, or amber once the system has assigned a designation. The table is decluttered: i.e., it only presents the critical attribute for the given target class. After the subject has completed a classify or a confirm operation on a target, the system removes the target entry from the table by scrolling the table. The subject uses a keypad to invoke a confirm or classify operation. For each operation, two sequential keypresses are required, one designating hostile or neutral, a second indicating the target number. For classify decisions, the subject needs to interpret the data in the table to determine the appropriate target designation. For confirm decisions, the subject needs to interpret the color of the word identifying the target class.

One important difference between the command language interface described above and the command language interfaces associated with more traditional office systems is that the table of target data is updated continuously. Such an approach is dictated in an aircraft context by the impact of external factors on the domain objects (i.e., the targets) and the real-time demands of the tactical domain. The approach makes less sense in an office system where, in most cases, changes to domain objects are made solely by the user and rapid response times are not as crucial.

The third interface (Figure 2c), the *graphical/keypad interface*, combines the low semantic distance of the first interface with the less direct engagement of the second interface. Like the command language interface, this interface splits the tactical assessment window into two portions. The top portion contains the simulated radar display; the bottom portion is for subject input and error feedback. The subject uses the keypad to enter his classify and confirm decisions.

Finally, the fourth interface (Figure 2b), the *tabular/pointer interface*, combines high semantic distance with direct selection of the tactical targets on the display using a touchscreen. The subject picks a target by touching the appropriate table entry. He designates the target by touching either the HOSTILE or NEUTRAL strip at the sides of the display. This last interface is similar to a menu interface, except that the table items are updated dynamically. Scrolling in this interface occurs just after the subject completes entry of the confirm or classify decision and is thus associated with the completion of a user action.

Distance and Engagement in the Interfaces

Although the four interfaces intuitively represent different combinations of semantic distance and engagement, it is important to understand the theoretical rationale for the level of distance and engagement in each interface. Metaphorically, the direct manipulation interface represents a model world of the task domain, the command language interface a verbal description. A graphical representation more closely matches the way that a pilot thinks about the tactical situation. More importantly, these two interfaces support the user's goals differently. We distinguish two user goals: to remain aware of the current tactical configuration, and to perform the assigned task. The low distance display was designed to support both goals. To support the first goal, the display continuously provided a graphical representation of the target's location and how the target was moving. To support the second goal, all relevant information about each target was encapsulated by this graphical representation.

The high distance display was designed to support only the second goal, user performance of the assigned task. In developing the high distance display, considerable effort was required to design a table that effectively supports the assigned task. For example, the target's spatial coordinates (x, y positions) were not provided because they are not relevant to the task and would have made the table harder to interpret. Moreover, the color code indicating the type of decision required was shown in the class column only, thus separating the system-assigned designation from the target attribute information. Finally, the columns were arranged to support efficient eye movements.

The levels of engagement can also be considered from several perspectives. We provide a pointing device (i.e., a touchscreen) for high engagement and a keypad for low engagement. The keypad uses a mode shift for two keys in order to preserve a common aspect of command language interfaces and to avoid introducing direct engagement with labelled keys for each action and object, a feature that Shneiderman associates with direct manipulation [15].

The theoretical difference between the levels of engagement in the interfaces is based upon the notion of a shared medium. In the direct engagement interfaces (the direct manipulation interface and the tabular/pointer interface), both the user and the computer system use a shared communications medium; that is, they both operate on the same objects. In the direct manipulation interface, the shared medium is the spatial display. The objects to be operated on are the target symbols and the strips labeled 'HOSTILE' and 'NEUTRAL'. In the tabular/pointer interface, the shared medium is the table, and the objects to be operated on are the table entries. In both direct engagement displays, the objects to be operated on and the strips share the same color code. Thus, for example, red in either the spatial display or the table of target attributes indicates that the subject should select the strip with the red wording.

In the indirect engagement interfaces (the command language interface and the graphical/keypad interface), the computer communicates to the user through one medium (i.e., section of the tactical display) and set of objects, while the user communicates to the computer through another medium (a keypad and another section of the tactical display) using a different set of objects. Thus there is a separation of the user input and computer output.

RESULTS

We found considerable support for our hypothesis: automation deficit was least with the direct manipulation interface and greatest in the interfaces that lacked one component of direct manipulation. We assessed automation

deficit by comparing subject performance on the first decision after the tactical assessment task was resumed to performance on the seventh decision. This effect was significant in the 12 subjects tested twice, $F(3,8) = 11.9$, $p < .002$. Similar results were found in the initial testing with the larger set of 20 subjects. As shown in Figure 3, with the direct manipulation interface, initial performance was as good as later performance. In other words, virtually no automation deficit was found with the direct manipulation interface. In contrast, automation deficit was clearly present in the two hybrid interfaces. Later performance was improved if either component of direct manipulation was present. This is shown by the reduction in response time for the later response in the two hybrid interfaces. If neither component of direct manipulation was present, as in the command language interface, both initial and later performance were poor. Further analysis has suggested that there may still be a deficit after a minute or so in handling events at a high rate with the command language interface when the tactical task is completely automated [1].

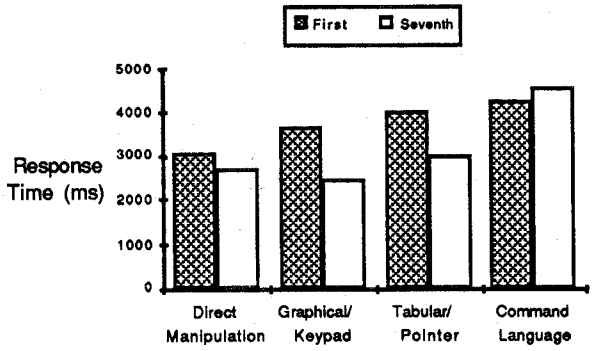

Figure 3. Interface effect on automation deficit in response time.

We also analyzed the effect of the type of decision and the type of display on automation deficit. We found that automation deficit was related significantly to the interaction between the type of decision and the type of display, $F(1,16) = 7.89$, $p < .02$. On classification decisions, automation deficit was greater with the tabular displays. On confirmation decisions, the deficit was greater with the graphical displays. The interaction is best illustrated by calculating the difference between the first response and the seventh response (see Figure 4). This pattern was also seen in the retesting four months later, although it was not as strong.

Looking at all the responses, not just the first and a later comparison response, we found no significant differences either in response time or in accuracy between the four interfaces. The reason is that responses with the non-direct manipulation interfaces improved during periods when the event rate was lower (i.e, fourth through sixth targets). Thus the four interfaces supported comparable performance in "normal" operation. In addition, although response times were slower in the retesting four months later, the effect was not significant and was not related to interface style. Accuracy was related to the type of decision and the type of information that had to be interpreted. Accuracy for the confirmation decisions was 95% and for the classification decisions was 78%. Accuracy was lowest for classification decisions which depended upon monitoring whether a number was changing. This occurs when the subject monitors the bearing of a fighter.

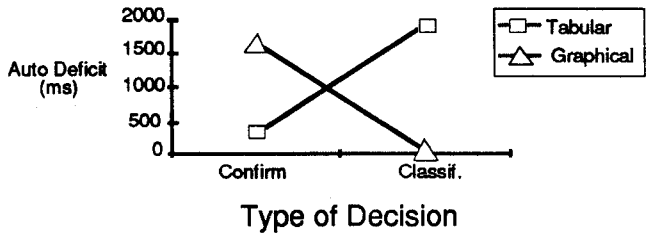

Figure 4. Automation deficit for different decisions with two types of tactical displays

Intra-Task Effects of Interface

In a multiple task domain, the interface for one task might have effects on other tasks. An interesting intra-task effect of engagement (keypad versus touchscreen) was found when the performance on the tracking task was examined. Those using the keypad for the tactical assessment task had better tracking in the initial phase of resuming the tactical assessment task than those using the touchscreen. To understand this result, it is useful to consider touchscreen usage as a form of tracking, and initial performance of this additional tracking task may interfere with making required adjustments to the other (joystick) tracking task. This result suggests that the touchscreen in the tactical assessment task induces an automation deficit in the tracking task. This occurs even though the subjects have been continually doing the tracking task.

Questionnaire Results

Twenty-four rating scales were used to obtain subjective judgments about feelings of control, feelings of awareness, preferences for the interface, judgments of the difficulty in learning and performing the tasks and specific aspects of the tasks, and ability to anticipate the changes in automation. Significant results were found on five scales. The most interesting results were the ratings of ability to anticipate changes in automation and awareness of the tactical situation at the end of automation. Ability to anticipate the changes in automation was dependent upon the type of tactical display. Those with the graphical display felt that they were able to anticipate the changes more often than those with the tabular interface. Furthermore, those with the graphical display felt that they were significantly more aware of the tactical situation at the end of automation. Debriefing confirmed that subjects with the graphical interface noticed the ebb and flow in activity during automation (i.e., activity picks up just before the task switches from automatic to manual), but those in the two tabular interfaces did not. This effect occurred despite the fact that tactical events were appearing

in both types of display at the exact same time, that the number of items in both is always the same at any particular moment, and that the ebb and flow of activity is exactly the same in each type of display.

These results show that subjects using the graphical display were able to monitor events during automation. Their ability to anticipate the changes could have produced improved performance, at least for those who had the graphical display and the touchscreen. Several other questions were asked about activity during automation, and subjects accurately described some global characteristics of what had occurred during automation (i.e, how many targets had been present), but not details.

DISCUSSION

Our research has implications for the theory of direct manipulation as well as for the design of interfaces for dynamic, multitask systems. The theoretical implications are based upon both empirical results as well as observations we made during the course of developing the interfaces and conducting the experiment.

On the positive side, we found that the theoretical predictions that we made were generally supported. This result is noteworthy for several reasons. First, this research is a rare example of designing interfaces to test a theory explicitly. Previous studies of direct manipulation and command language interfaces have used interfaces for established applications which may not fairly represent the theoretical concepts. Second, our predictions concern a specific aspect of performance (automation deficit) in a complex, multitask situation. Either challenge---specificity of prediction or complexity of context---would put demands on a theory. Both were present in this research, which makes the successful predictions of the theory especially impressive.

However, we also found that the theory has limitations. First, the theory does not address interfaces which include a mixture of interface styles and which are probably the rule more than the exception in complex applications. The reason is that complex applications involve different types of tasks. A single interface style may not support all tasks in an optimal manner. In the HHN theory, a general interface for the application is assumed. This requires choosing a representation that is suitable for most tasks. But it may not be optimal for certain tasks. Thus choosing a single interface style for a complex application may produce suboptimal performance on some aspects of the application.

This point is important because it is based not only on observation but on empirical results. In our data we found evidence that the optimal display for reducing automation deficit depends upon the type of decision. Simple decisions were served better by the tabular display, complex decisions by the graphical display. In terms of theoretical predictions, the shortcoming of the HHN theory is that it (and we) did not make predictions about the simple decisions. In retrospect, it is evident that the theory would have to be modified to address decision complexity. It is likely that the confirmation decisions were best supported by the tabular display because the user did not need complete information about the object but simply needed to know the value of a single parameter. If the model world metaphor is implemented faithfully, then different representations for different decisions are not directly possible. Thus an extension of the theory should be considered to support different levels of representation for different requirements.

Second, we found that the theory does not always help with detailed aspects of interface design. Our goal was to evaluate interfaces that had different levels of distance and engagement. The iterative design process we used forced many decisions about details of each of the four interfaces. Many of these decisions were based upon performance considerations and could not be based upon logical derivations from the tenets of the theory. Furthermore, the performance constraints were related to the specific application. For example, the relative placement of the two windows (horizontal or vertical) had an impact on how easy it was to use hands dedicated to the two tasks. This is a stimulus-response compatibility issue that the theory does not address. In essence, the theory is not performance based as are other formal models such as GOMS. It is most relevant in dealing with aspects of the interface that relate to its cognitive complexity.

Finally, we found that distance and engagement are difficult terms to define operationally and to evaluate. Our experiment required interfaces that combined different levels of distance and engagement. In other words, these were design requirements for the interfaces. One of the problems is how to distinguish between distance and engagement. Our empirical results suggest that they are not independent, in that the degree of automation deficit in the command line interface was not a combination of the deficits in the two hybrid interfaces, which each lacked an aspect of direct manipulation. HHN themselves point out that engagement is only present when both semantic and articulatory directness is present.

The interfaces that we produced represented combinations of different levels of distance and engagement. What is not clear is how much distance and engagement were actually present. It is apparent that any interface that allows the person to perform a task successfully has bridged the distance of the gulfs of execution and evaluation as HHN discuss them. The command language interface we produced supported the user's goal of performing the task and therefore reduced semantic distance to a greater degree than an interface which would not support this goal. And yet, it did not provide a view of the model world as a pilot would normally think of it, so considerable distance still remained. Better precision about the degree of distance and engagement in an interface would be helpful.

CONCLUSION

Based upon our findings, we expect that intermittent operation of complex tasks will be more effective with direct manipulation interfaces in a variety of dynamic, real-

time systems. Although our results were found in a cockpit application, extension to other systems is appropriate, particularly systems in which the operator is intermittently moving from one task to another. To envision potential generalizations, it is helpful to characterize our application in abstract terms. The dual task application we tested included 1) a continuous task with simple perceptual demands, rigorous manual demands and minor cognitive complexity; and 2) an intermittent task with varying cognitive and perceptual complexity and minimal manual demands. The cognitive complexity of the intermittent task was manipulated by changing the interfaces and by changing the decisions. The results can be interpreted at an abstract level: increases in the cognitive complexity of an interface adversely affect the resumption of its use after a period of automation. This principle certainly holds for systems that include the two types of tasks. The principle would probably hold for systems which have greater complexity on the continuous task. In fact, the effects of interface would probably be greater. The key to appropriate generalization is that there was relatively little cognitive interaction between the two tasks. There was some manual interaction as noted below.

Generalization may not be warranted if the system includes multiple tasks which use similar cognitive processes. In a multitask application, there may be different forms of expressions to the various tasks; the interaction of these expressions is an important issue. Direct engagement in particular may introduce incompatibilities. We found that tracking performance was adversely affected in the initial seconds of resuming pointing with the touchscreen. The cause was an incompatibility between the two forms of manual manipulation. The important issue is whether direct manipulation interfaces to different tasks could compete. According to Wickens [13], the answer is yes. In his resource theory, competition for attentional resources occurs whenever information to the user is in similar modalities or is in a similar code (e.g., spatial or verbal). Competition also occurs whenever responses are similar. Thus two direct manipulation interfaces which both have spatial graphical displays, and which both require pointing devices could produce competition for attentional resources. Thus the generalization of our results to other multiple task systems should be made with consideration given to possible competition between aspects of the direct manipulation interface.

ACKNOWLEDGMENTS

We acknowledge Rob Jacob for contributing to the initial idea for the hypothesis and Rob Carter and Diane Damos for contributions to the design. Many individuals reviewed the interfaces and contributed to their development. We thank the subjects for their participation. This research was supported by the Office of Naval Technology.

REFERENCES

1. Ballas, J. A., Heitmeyer, C. L., and Pérez, M. A. *Direct Manipulation and Intermittent Automation in Advanced Cockpits*. Technical Rep. 9375. Naval Research Laboratory, Washington, D. C. (in press).
2. Beltracchi, L. A direct manipulation interface for heat engines based upon the Rankine cycle. *IEEE Transactions on Systems, Man and Cybernetics*. **17**(3), pp. 478-487 (1987).
3. Benson, C. R., Govindaraj, T., Mitchell, C. M. and Krosner, S. P. Effectiveness of direct manipulation interaction in the supervisory control of FMS parts movement. *Proc. IEEE International Conference on Systems, Man, and Cybernetics*. Cambridge, MA, pp. 947-952 (Nov. 14-17, 1989).
4. Bernotat, R. K. Man and computer in future on-board guidance and control systems of aircraft. In B. Shackel (Ed.) *Man-computer interaction: Human factors aspects of computers and people*. Sijthoff & Noordhoff, Rockville, MD (1981).
5. Draper, S. W. Display managers as the basis for user-machine communication. In D. A. Norman & S. W. Draper (Eds.) *User-centered system design*. Erlbaum Associates, Hillsdale, NJ, pp. 339-352 (1986).
6. Hutchins, E., Hollan, J. D. and Norman, D. A. Direct manipulation interfaces. in D. A. Norman & S. W. Draper (Eds.) *User-centered system design*. Erlbaum Associates, Hillsdale, NJ, pp. 87-124 (1986).
7. Karat, J. Evaluating user interface complexity. *Proceedings of the Human Factors Society 31st Annual Meeting*. Human Factors Society, Santa Monica, CA, pp. 566-570 (1987).
8. Kieras, D. E. An overview of human-computer interaction. *Journal of the Washington Academy of Sciences*. **80**(2), pp. 39-70 (1990).
9. Marshak, W. P., Kuperman, G., Ramsey, E. G., and Wilson, D. Situation awareness in map displays. *Proceedings of the Human Factors Society 31st Annual Meeting*. Human Factors Society, Santa Monica, CA pp. 533-535 (1987).
10. Parasuraman, R., Bahri, T., Deaton, J. E., Morrison, J. G. and Barnes, M. *Theory and design of adaptive automation in aviation systems*. Cognitive Science Laboratory, The Catholic University of America, Washington, D. C. (1990).
11. Reising, J. M. and Hartsock, D. C. Advanced warning/caution/advisory displays for fighter aircraft. *Proceedings of the Human Factors Society 33rd Annual Meeting*. Human Factors Society, Santa Monica, CA. pp. 66-70 (1989).
12. Whiteside, J., Jones, S., Levy, P. S. and Wixon, D. User performance with command, menu, and iconic interfaces. *Proc. ACM CHI'85 Human Factors in Computing Systems*. ACM, New York. pp. 185-191 (1985).
13. Wickens, C. D. and Liu, Y. Codes and modalities in multiple resources: A success and a qualification. *Human Factors*. **30**(5), pp. 599-616 (1988).
14. Wiener, E. L.. Beyond the sterile cockpit. *Human Factors*. **27**(1), pp. 75-90 (1985).
15. Ziegler, J. E. & Fahnrich, K. P. Direct manipulation. In M. Helender (Ed.) *Handbook of human-computer interaction*. Elsevier Science Publishers, North-Holland, pp. 123-133 (1988).