

Modeling a Continuous Dynamic Task

Wayne D. Gray, Michael J. Schoelles, & Wai-Tat Fu

Human Factors & Applied Cognition

George Mason University

Fairfax, VA 22030 USA

+1 703 993 1357

gray@gmu.edu

ABSTRACT

Milliseconds matter – when the interface permits, users will adjust their behavior to shave 100-200 msec from low level interactions. Saving milliseconds involves deploying different microstrategies. These strategies are deployed automatically; that is, non-deliberately. To explore the ability of people to maximize performance by deploying different microstrategies, two experiments were conducted on two continuous dynamic tasks. The data support the conclusion that people respond to small differences in their task environment. The data from each experiment were modeled with ACT-R/PM, a theory that combines the ACT-R theory of cognition with a modal theory of visual attention and motor movement. The models suggest that differences in performance between the two experiments stem from the use of different microstrategies that are deployed in response to subtle differences in the interfaces used by the two experiments.

Keywords

ACT-R/PM, microstrategies, attention, motor movement, computational cognitive models

INTRODUCTION

Milliseconds matter – when the interface permits, users will adjust their behavior to shave 100-200 msec from routine, low level interactions (Gray & Boehm-Davis, 1999). Such adjustments to behavior are automatic; that is, non-deliberate. However, they reflect strategic choices rather than inevitable consequences of an underlying cognitive architecture. Because of their scope (100's of milliseconds) and their non-deliberate nature, these choices are termed *microstrategies*.

The initial arguments for microstrategies were illustrated with the descriptive modeling language CPM-GOMS (Gray, John, & Atwood, 1993; John, 1990). In this paper, we extend those arguments by building ACT-R/PM models that interact directly with the interfaces used by the participants in our studies. ACT-R/PM (Byrne, 1998) embodies the cognitive architecture of ACT-R (Anderson & Lebière, 1998) with the hands from EPIC's motor module (Kieras & Meyer, 1997) and a visual attention module (Anderson, Matessa, & Lebière, 1997). In so doing, ACT-R/PM permits motor and visual attention processes to be executed in parallel with each other as well as with cognitive processes.

The current effort was undertaken in the spirit of exploring ACT-R/PM's mechanisms in a fast paced task. In the two experiments reported here, participants moved to and clicked on buttons at the pace of about 1 sec per button for a continuous period of about 32 sec per block. To our knowledge, this makes these the fastest paced continuous tasks for which embodied models have been built.

The two experiments differed in several ways. Modeling the impact of one such difference raised theoretical issues in motor movement and attention. For motor movement, the issue concerned the functional shape and size of a target that is not visible when movement to the target begins. We term this the non-object effect. For attention, the models raise questions about the limits of a pre-attentive feature search (Treisman & Gelade, 1980) and suggest that an interface that was intended to enhance performance actually had the opposite effect.

The next section introduces those aspects of the button task that were common to both experiments. This overview is followed by a description of the procedures, results, and models of experiment 1 (E1) and then the procedures and results of experiment 2 (E2). In the penultimate section, the model built for E2 is discussed and compared with the model for E1. The differences between the two models are related to a difference in the interface used for the two experiments. Modeling differences in performance across these two interfaces revealed structural differences in the microstrategies for each task that are not predicted by extant attention and motor theory.

THE TASK

The two experiments used variations of the procedure shown in Table 1. For both experiments, there are two types of buttons: one home button and multiple target buttons. The home button appears on even numbered trials in a fixed location. The target button appears on odd numbered trials, at random, in one of a small number of known locations. Hence, the microstrategies for moving to the target button must wait for perception to deliver the target button's location to cognition. In contrast, as the location of the home button never varies, the microstrategies for moving to the home button can begin before the home button appears.

Table 1: Experimental procedures

Step	Experiment	Display	Participant
Step 1	Block starts	Home button appears	Moves cursor to and click on home button
Step 2	Timing starts	Home button disappears	
Step 3	Odd numbered trials	Target button appears at random in one of a small number of predefined locations	Moves cursor to and clicks on target button
Step 4		Target button disappears	
Step 5	Even numbered trials	Home button appears	Moves cursor to and clicks on home button
Step 6		Home button disappears	
Step 7	End of trial	Feedback information appears with total time per block and comparison of current block with running average of prior blocks	

For each study, the data reported are for error-free trials only. Any trial on which the participant clicked on the window before clicking on the button was excluded. Also excluded were trials in which the participant moved the cursor through the button, stopped, returned to the button, and then clicked.

EXPERIMENT 1

The first experiment used a five-button display with the home button in the center (see Figure 1). As shown in the figure, when a button was not on the screen its position was marked by an outline. E1 investigated the effect of a delay between the disappearance of one button and the appearance of the next one. We expected that a delay between the disappearance of the home button and the appearance of the target button (i.e., between steps 2 and 3 of Table 1) would lead to an increase in response time that directly corresponded to the length of the delay. In contrast, as participants knew where the home button would appear, the interval between the disappearance of the target button and the appearance of the home button (i.e., between steps 4 and 5 of Table 1) should be absorbed by a microstrategy that began to move the cursor to the home button's location before the home button could appear and be recognized.

Method

Participants. Thirty-six George Mason University undergraduates participated in the study for course credit. The experiment took approximately 30 minutes to complete. About halfway through the study, the error rate for three of the participants soared and remained steady. These participants were dropped from the study. As a result, the 0, 200, and 500 delay conditions had 12, 11, and 10 participants respectively.

Material and software. Both experiments were written in Macintosh Common Lisp (MCL 4.2) and run on a Macintosh PowerPC computer with a 17" monitor. The mouse movement settings used were the default setting for the operating system. The following mouse events were recorded and saved to a log file with 16.67 msec accuracy: mouse down and mouse up times on the window or a particular button, the time that the mouse

entered a button, and for those cases in which the cursor overshot the button, the mouse leave time.

Procedure. The basic procedure outlined in Table 1 was followed. No more than one button appeared on the screen at a time. When a particular button was not on the screen, its location was represented by dotted lines that outlined its shape (as per Figure 1).

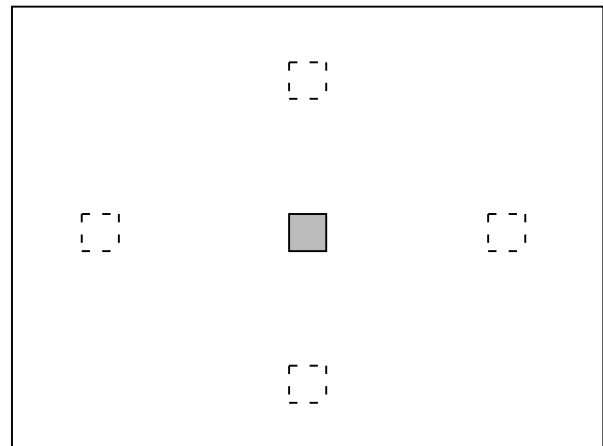


Figure 1: Screen layout for E1 showing the home button and the location of the four target buttons. When a button was not shown, its location was indicated by an outline

There were three between-subjects delay conditions: 0-msec, 200-msec, and 500-msec. The session began with a practice block that required clicking on 8 target and 8 home buttons. After the practice block ended, the experimenter left the room and the participant continued for 35 blocks of 16 target and 16 home buttons per block. For each block, the left and right buttons each appeared 6 times; the top and bottom buttons each appeared twice. Hence, the top and bottom buttons functioned as distracters that served to discourage participants from anticipating whether the next target button would appear to the right or left of the home button. The analyses reported below include the times to move to and from the left and right button. Movement times to and from the top and bottom buttons are excluded.

Results

As we did not wish to model learning, data from the first ten blocks were discarded. As a preliminary analysis had shown no difference in moving from home to either the left or right target button, or in moving from the left and right buttons to home, this factor was collapsed into a to-target versus to-home factor (target-home). For each participant, the median mousedown to mousedown time for to-target and to-home was calculated for blocks 10-35.

The overall results are illustrated by the solid lines in Figure 2. There was a main effect of direction – moving to-home was faster than moving to-target [$F(1, 30) = 1332.44, p < .0001, MSE = .003$].

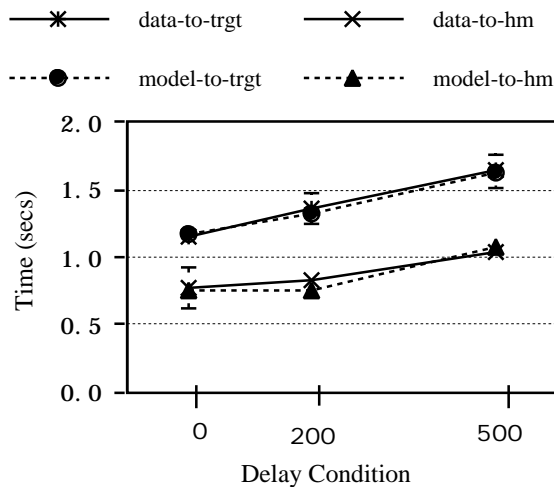


Figure 2: Data and model fit for E1. Time to move to and click on target button versus home button for participants (solid lines) and model (dotted lines). Error bars show the standard deviation for the empirical data.

Not surprisingly, increasing the delay before a button's appearance from 0 to 200 to 500 msec increased response time [$F(2, 30) = 31.80, p < .0001, MSE = 0.024$]. However, of more interest is the interaction of delay with direction (i.e., to-target versus to-home) [$F(2, 30) = 25.52, p < .0001, MSE = .003$].

The interaction suggests that there are two, largely independent, effects contributing to the results of the current study. Movement time to-target is greatly influenced by delay time; movement time to-home is not. For to-target, the nearly straight line in Figure 2 shows that the increased time to move to the target button is completely accounted for by the increased delay in the button's appearance. In contrast, for to-home the small increases in response time do not mirror the large increases in delay.

Model

The fit of the model to the data is shown by the dotted lines in Figure 2. As the figure suggests, the fit is quite good with an r^2 of 0.985 and a root mean square deviation (rmsd) of 0.041.

For movement time, model parameters for calculating Fitts law were set to the parameters used by the participants. These included screen resolution (832 x 624), pixels per inch (76), and distance from screen (20"). The ACT-R parameter *enable rational analysis* was turned on and the *base level learning* parameter was set to the default (0.500). The approach-width for moving to-home or to-target was set to the actual width of the button. ACT-R/PM has a number of parameters that govern the motor and visual modules; however, none of these parameters were altered from their default settings.

The model used a two-deep goal hierarchy. A subgoal was pushed to click on the home button. When this task was achieved, the goal was popped and a new goal was pushed to move to the target button. The button-clicking task was performed by two conceptually similar sets of three productions; one set per subgoal -- shift-attention, move-cursor, and clickOn-button.

The productions for shifting attention were functionally identical except for how they decided when to move attention. For the target button, the shift-attention production does a pre-attentive search to determine if it can detect a visual feature of a button. In contrast, the production for shifting attention to-home attempts to retrieve a declarative memory element that contains the location information for the home button.

Both productions shift attention to a location. The key difference is that to shift attention to-target, a target button must be on the screen. In contrast, attention can begin to shift to the home location almost immediately after the subgoal for moving to-home is set. Although the memory must be retrieved, the home button does not need to be on the screen before attention begins to shift.

The key condition for moving the cursor to-home is identical to that for shifting attention to-home – a declarative memory element is retrieved that contains the location of the home button. However, unlike the pre-attentive feature search conducted to shift attention to-target, moving the cursor to-target requires that a visual object that is a target button be on the screen now. In ACT-R/PM, visual objects are created when attention shifts to a location and focuses on a feature. Hence, for the target button, the move cursor production cannot fire until after the shift attention production completes.

The two clickOn-button productions are essentially identical. Both test to determine whether there is a button object on the screen and whether the cursor feature is within the bounds of the button object. If so, then the motor module is sent a command to click. More details on the model and on the interleaving of the model's

Gray, W. D., Schoelles, M. J., & Fu, W.-t. (2000). Modeling a continuous dynamic task. In N. Taatgen & J. Aasman (Eds.), *Proceedings of the Third International Conference on Cognitive Modeling* (pp. 158-168). Veenendaal, The Netherlands: Universal Press.

cognitive, attention, and motor components can be found in Gray, Schoelles, and Fu (2000).

Summary

The data and model for the first experiment support the proposition that microstrategies can be deployed that shave milliseconds from the performance of a routine interaction. Moving to-target requires moving to a variable location. Before the move can begin, perception must deliver position information to cognition. In contrast, when moving to a fixed location, location information can be retrieved from memory. As discussed at greater length in Gray, Schoelles, and Fu (1999), the ability to retrieve location information enables a microstrategy that begins to shift attention to the home location even before the mouse down on the preceding target button has been completed.

EXPERIMENT 2

The second experiment used nine buttons arranged in three rows of three columns. The home button was at bottom center. In contrast to E1, when a button was not on the screen, its position was not indicated in any way. Another difference between studies is that in E2 we did not manipulate delay; that is, for all participants, there was a 0-msec delay between the disappearance of one button and the appearance of the next one.

E2 co-varied the distance and angle of approach between the home and target buttons. Both of these parameters are important in Fitts law calculations (e.g., see MacKenzie, 1992); both are modeled by ACT-R/PM's motor module (Byrne, 1999) which is based on EPIC's motor module (Kieras & Meyer, 1996). As Figure 3 illustrates, the approach-width for an object varies with angle of approach.

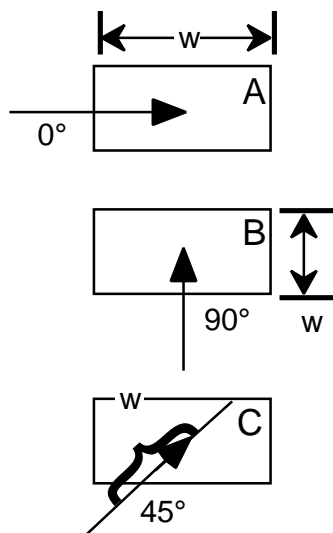


Figure 3: A-C illustrate how three different angles of approach to the same object result in three different approach-widths (w).

Unlike E1, for E2, until a button appears, there is no reminder of its location on the screen. We expected that the absence of a reminder would test the generality and accuracy of the microstrategies that relied on memory of button location. We suspected that for E1, the presence of the button's outline when the button itself was absent might have influenced the ease of moving to-home. As discussed below, the absence of outlines on the display had an affect, but not the one we had anticipated. Modifying the model to account for the new data forced us to make subtle, but theoretically interesting, changes to the model.

Method

As discussed above, there were nine button positions, arranged in three rows of three columns. The home button was in the bottom-middle position. Other than the arrangement of the display, the materials and software were the same as in E1. E2 included a between-subjects manipulation of two different button icons. As this manipulation produced no effect ($F < 1$) it is ignored here.

Subjects. Thirty George Mason University undergraduates participated for course credit. The experiment took approximately 30 minutes to complete. Data from one of the participants were discarded as partway through the study he began making a high number of errors.

Procedures. The basic procedure outlined in Table 1 was followed. No more than one button appeared on the screen at a time. When a particular button was not on the screen, there was no visible reminder of its location.

Results

As for E1, data from the first ten blocks were discarded. For each participant, the median mousedown to mousedown time for moving to each target and moving from each target to home was calculated for blocks 10-35.

For each button, the first bar in Figure 4 shows the mean time for moving to and clicking on that button. The third bar shows the mean time for moving from that button and clicking on home. As suggested by the figure, an ANOVA showed a main effect of direction – moving to the home button was faster than moving to the target button [$F(1, 28) = 752.79, p < .0001$ ($MSE = .004$).

By Fitts law, movement time is a function of distance and approach-width. As the distance and approach-width of the eight target buttons from home varied, it was expected that the main effect of target type would be significant [$F(7, 196) = 159.21, p < .0001, MSE = .002$]. Of more interest was the significant interaction of target type with direction (to-target versus to-home) [$F(7, 196) = 12.28, p < .0001, MSE = .001$]. This interaction was unexpected. Indeed, according to the model fitted to E1, the difference in time for direction should be the same for all targets.

MODELING EXPERIMENT 2

We attempted to model E2 using the same productions and parameters as used for E1. However, the fit of the

location and create a declarative memory element encoding the object at that location.

The button outlines in E1 were intended to make it easier

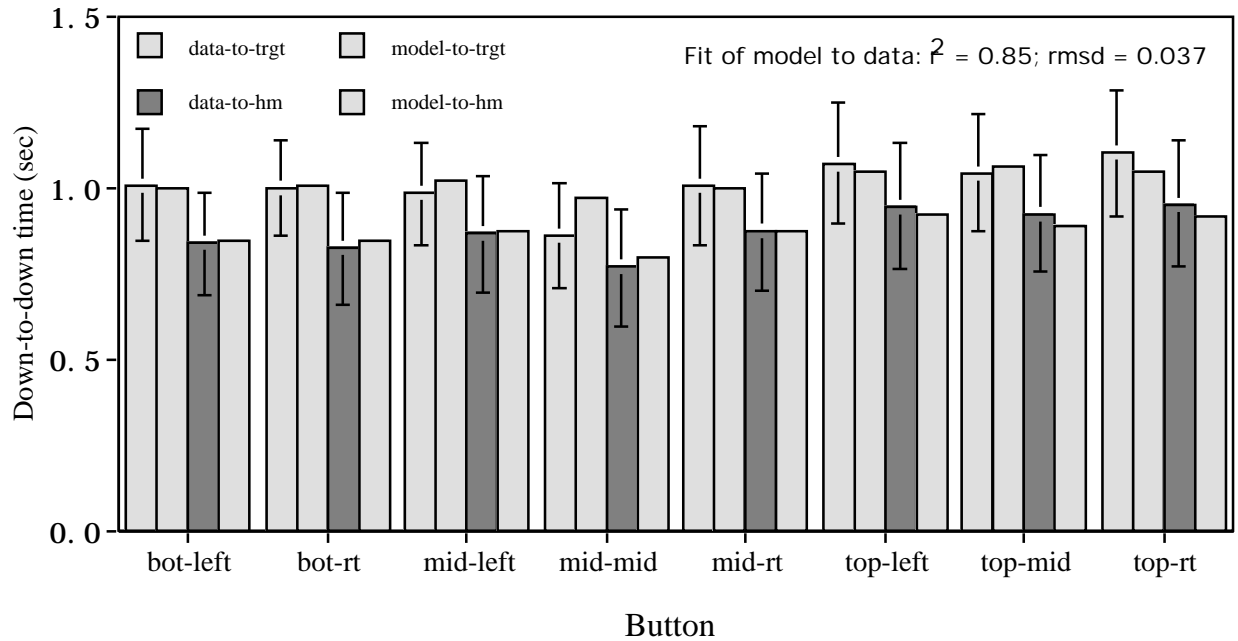


Figure 4 Mouse down to mouse down for E2. Error bars show the standard deviation

model was not very good. This misfit caused us to take a closer look at the empirical data.

For the 0-msec delay condition in E1 the mean time to move to the target was 1.150 sec whereas the mean time to move to home was 0.782. For E2, the mean time to move to the target was faster than for E1, 1.013 sec, whereas the mean time to move to home was slower, 0.877. These differences suggest that the two experiments posed different tasks to their participants.

To summarize, in moving to-target, E1 was slower than E2, but in moving to-home, E1 was faster than E2. We believe that the two differences are independent of each other but stem from the same simple difference between the interfaces used in the two experiments. In E1, when the buttons were not present, their location was marked by an outline. In E2, in the absence of a button there was no visible reminder of its location.

Detecting features versus objects: Explaining E1 versus E2 differences in moving to-target

As discussed above, the model for E1 suggested a conservative criterion for moving the cursor to the target button. Rather than moving to a location based on the results of a pre-attentive search, the cursor would not be moved until a button object was created. The default parameter in ACT-R/PM assumes that after the production fires, it takes 135 msec to shift attention to a

for participants to adopt a strategy of moving to the home location before the home button appeared. However, an unwitting side effect of the outlines may have been to provide feature clutter. The paradigmatic example of pre-attentive search is looking for the letter “T” in a display of “O’s.” In this case, the angular features of the T make it seem to “pop out” from the O’s. In contrast, a more effortful search is required if the T is embedded in a display of F’s and L’s. In this case, there is an overlap in features such that the solitary T does not pop out.

For E1, the overlap in color and shape features between the buttons and their outlines may have precluded the use of pre-attentive search as a reliable microstrategy. Participants may have adopted the more conservative microstrategy of waiting the extra 135 msec before moving rather than false alarming by moving to an outline rather than a button. (Our data include only perfect trials. Any trial in which a participant clicked on another location prior to clicking on the button was excluded.) However, in E2, there was no visual clutter. In E2, the screen was always blank except for the current button. Hence, in E2 moving the cursor to a location based on a pre-attentive feature detection process (as per the shift-attention production) would be a safe and fast microstrategy. This microstrategy would enable the participant to shave 135 msec from the time needed to move to the target. Except for this change, all other productions in the two experiments were identical.

Changing the production for moving the cursor to the target produced a reasonably good fit between model and the to-target data. This fit can be seen in Figure 4 by comparing the first two bars for each button. However, this change did not affect the fit of model to the to-home data.

The non-object effect: Explaining the variability in moving to-home

The better fit in moving to the target button was achieved by a simple change in one production that decreased the response times for moving to all target buttons by a constant amount. However, it soon became apparent to us that merely shifting all to-home times up or down by a constant amount, would not improve the fit of model to the to-home data.

The interaction of target type with direction (see above) provided the clue needed to obtain a better fit. This interaction suggested that the difference in time to go to-target versus to-home was not a simple constant but varied somehow with the target type.

The key to a better fitting model was the approach-width calculation. As shown by Figure 3, the approach-width to a given object varies with the angle of approach. By the standard calculation, the approach-width in moving from home to any given target was identical to moving from that target to home. Hence, along with distance, the approach-width varied button by button. (Actually, the distance and approach-width were constant for the following pairs of buttons: top-left and top-right, middle-left and middle-right, and bottom-left and bottom-right.)

The standard calculation of approach-width makes a standard assumption; namely, that the target is visible when movement begins. However, in moving to-home, the participants began to move the cursor before they could have attended to the home button. In E1, although the home button was not present, its form and shape was clearly marked by the outline. The question for E2 can be stated as, "what is the shape of a square when the square is not there?"

A working answer to this riddle is the *non-object effect*; namely, the shape of a square when the square is not there is round and small. Assuming a constant approach-width to-home from any target reproduced the effect of the interaction in the empirical data. When the approach-width in moving to a target is based on the target's actual width, but the approach-width in moving to-home (from a given target) is based on a constant width, then the difference between moving to versus moving from is not a constant but varies button by button. Hence, assuming a constant approach-width enabled the model to mimic the interaction found in the empirical data. Having a constant approach-width is functionally equivalent to assuming that the target is round.

When the approach-width was fixed to the actual width of the button (i.e., the length of the button's X-axis), the

model's times to-home were faster than the empirical data. As ACT-R/PM follows Fitts law, increasing or decreasing the fixed width for home, decreased or increased the model's times for all to-home movements by a constant amount. By trial and error, we found that the best fit was obtained by assuming that the button *that is not there* was not only round but was 2/3's the width of the actual button.

E2 Model Summary

As shown in Figure 4, the model provided a reasonable fit to the empirical data with $r^2 = 0.85$ and $rmsd = 0.037$. All but one of the production rules used in E2 were the same as those used in E1. The only parameter differences between the two models were the approach-width and size of the home button. In E1, the approach-width and size of the home button was the same as the target button. In E2, the approach-width for each target button was calculated based on approach angle; however, for moving to-home a constant approach-width was used regardless of approach angle. This approach-width assumes that the functional shape of the home button was round and its functional size was one third smaller than its actual size.

DISCUSSION AND CONCLUSIONS

Critics tend to dismiss cognitive modeling as post-hoc fitting of data that produces no new insights and no surprises. "After all," so the argument goes, "if you built the model how could it do anything that surprises you?" There are many solid ways of answering this criticism – several of which are represented by the current paper.

E1 used a fast and continuous speeded response paradigm to test whether models of embodied cognition could simulate the response times of participants and, in so doing, provide an explanation of the microstrategies used by participants. Using the ACT-R/PM architecture that provides for the interleaving of cognition, motor movements, and shifts of visual attentional, various microstrategies emerged. The model supports the hypothesis that the difference between moving to-target versus to-home can be explained by the time needed to perceive and verify that a button object exists. As the delay was increased, the increases in time to move to-target were mimicked by the model as the inevitable delay in the time required to wait and see where the target appeared before movement could begin. In contrast, the much smaller increase in the time to move to-home was accounted for by the cursor arriving at the home location before the button appeared. When the task is stable and predictable, participants are able to move to the home location before it is possible for perception to deliver the button's location to cognition.

Without the models, the results of E1 and E2 would have been cited as simply supporting our general claim that microstrategies exist and are the key to understanding interactive behavior [as per \Gray, 1999 #988]. However, the best model for each study suggested interesting

differences that, if confirmed, will deepen our understanding of how interactive behavior emerges from the constraints and opportunities provided by the interaction of human cognition with task goals and the artifact used to perform the task (see also Gray, in press; Gray & Altmann, in press).

The changes that we made to the model were constrained by the modeling architecture. Hence, the changes were not arbitrary and each change pointed to an existing theory – one of pre-attentive search and one of motor movement – that was underspecified. For moving to-target, the absence of screen clutter in E2 as opposed to the presence of button outlines in E1, suggested an E1 versus E2 difference in the viability of a microstrategy that relied on pre-attentive search to guide cursor movement. For moving to-home, unlike E1, in E2 the home location was completely blank when movement to-home began. The microstrategy that best fit the data assumed that the approach-width to a non-object was the same from any angle.

The changes made to the model are suggestive not definitive, and no other conclusion regarding these changes should be drawn from the current data. The changes made to model the E2 data are hypotheses that we plan to pursue in future empirical and modeling studies. However, both changes stem from the absence in E2 of an interface feature used in E1, button outlines, that was intended to facilitate performance. The differences in microstrategies between models have deepened our understanding of the subtle opportunities that exist for unintended consequences of interface design.

ACKNOWLEDGEMENTS

The work reported was supported by a grant from the Air Force Office of Scientific Research AFOSR#F49620-97-1-0353 as well as by the National Science Foundation (IRI-9618833).

REFERENCES

Anderson, J. R., & Lebière, C. (Eds.). (1998). *Atomic components of thought*. Hillsdale, NJ: Erlbaum.

Anderson, J. R., Matessa, M., & Lebière, C. (1997). ACT-R: A theory of higher-level cognition and its relation to visual attention. *Human-Computer Interaction*, 12(4), 439-462.

Byrne, M. D. (1998). Perception and action. In J. R. Anderson & C. Lebière (Eds.), *Atomic components of thought* (pp. 167-200). Hillsdale, NJ: Erlbaum.

Byrne, M. D. (1999). *ACT-R Perceptual-Motor (ACT-R/PM): A users manual*. <http://www.ruf.rice.edu/~byrne/RPM/docs/index.html>.

Gray, W. D. (in press). The nature and processing of errors in interactive behavior. *Cognitive Science*.

Gray, W. D., & Altmann, E. M. (in press). Cognitive modeling and human-computer interaction. In W. Karwowski (Ed.), *International encyclopedia of ergonomics and human factors*. New York: Taylor & Francis, Ltd.

Gray, W. D., John, B. E., & Atwood, M. E. (1993). Project Ernestine: Validating a GOMS analysis for predicting and explaining real-world performance. *Human-Computer Interaction*, 8(3), 237-309.

Gray, W. D., Schoelles, M. J., & Fu, W.-t. (1999). Modeling microstrategies in a continuous dynamic task. *Manuscript submitted for publication*.

John, B. E. (1990). Extensions of GOMS Analyses to Expert Performance Requiring Perception of Dynamic Visual and Auditory Information. *Proceedings of the ACM CHI'90 Conference on Human Factors in Computing Systems* (pp. 107-115). New York: ACM Press.

Kieras, D. E., & Meyer, D. E. (1996). *The EPIC architecture: Principles of operation* (Unpublished manuscript from <ftp://ftp.eecs.umich.edu/people/kieras/EPICarch.ps>). Ann Arbor: University of Michigan.

Kieras, D. E., & Meyer, D. E. (1997). An overview of the EPIC architecture for cognition and performance with application to human-computer interaction. *Human-Computer Interaction*, 12(4), 391-438.

MacKenzie, I. S. (1992). Fitts' law as a research and design tool in human-computer interaction. *Human-Computer Interaction*, 7(1), 91-139.

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