Where No Interface Has Gone Before: What Can the Phaser Teach Us About Label Usage in HCI?

Franklin P. Tamborello, II
Phillip H. Chung
Michael D. Byrne
Rice University
Department of Psychology
6100 S. Main, MS-25
Houston, TX 77005
{tambo, pchung, byrne}@rice.edu

Abstract

Most research on how people represent procedures suggests that control labels are central. However, our data suggest that even moderately-experienced users do not rely on labels to locate interface elements.

Keywords

label usage, interface layout

ACM Classification Keywords

H5.m. Information interfaces and presentation (e.g., HCI): GUI, Screen Design

Introduction

Most of us take for granted the importance of labeling controls on an interface. How else could the user know what each control does? Yet when one thinks about how rapidly a skilled user can perform a familiar task using a familiar interface, it hardly seems likely that he/she is taking the time to read the labels on his/her controls.

Surely there are plenty of good reasons for including control labels in an interface. For example, novices need to know what each control does, and expert users still need to be reminded what a seldom-used control does. But anticipating who uses what control under what circumstances is extremely difficult, if not

impossible. Given the above anecdotal evidence that users often disregard labels, it seems plausible that interface labels and other text messages should not be relied upon solely for very important communication from the machine to the user.

We have found experimental evidence suggesting that experienced users of an interface rely very little upon textual labels to perform their tasks. Given that, it seems likely that they are instead using some kind of spatial representation to find the objects they need to perform their actions.

Method
To test our hypothesis that skilled users do not rely on labels for commonly-used controls, we wanted to use a task that 1) incorporates a sufficient training component to ensure that users are skilled at the task when data is collected from them, and 2) incorporates some task components other than reading, such as clicking buttons. Meeting those two design criteria ensures that the task satisfies the two major assumptions of our hypothesis, namely that users are sufficiently familiar with the task and with the interface that they need not search for controls and that there is something for them to do besides reading things. Furthermore, we wanted a contrasting condition that instead varied some spatial aspect of the interface.

To this end we used one of the interactive Star Trek-themed computer tasks, the Phaser[1 – 4]. The Phaser uses a single-screen GUI composed of radio buttons, check boxes, buttons, and a few other elements (Figure 1). Twelve subjects (mean age 25.1 years, age standard deviation 4 years, 5 female) participated in the label removal condition, 18 in the added buttons

condition (mean age 19.5 years, age standard deviation 1.7 years, 11 female).

Participants trained for by reading an instruction manual, then performing the necessary steps once with the manual present as reference, then three times more without the manual. Only trials in which the user committed no errors were counted toward training criteria. Each participant returned approximately one week later for the test session. The test session simply consisted of performing the task with no assistance from the manuals or the experimenter. Participants performed 14 trials of the Phaser task, with the first four considered practice and therefore not included in analyses.

Instructions at the beginning of the test phase noted that the interface may or may not change during the experiment and that the user would be warned of any interface changes. Participants completed the first seven trials of the Phaser using the control version of that task's interface (see Figure 1A). Before the onset of the eighth trial, the experiment displayed instructions warning subjects that the Phaser interface was about to change, but the instructions did not specify how the interface would change. The eighth and all subsequent trials of the Phaser all used the changed interface (see Figure 1B). The changed interface was exactly the same as the control interface with the exception that labels for all controls were replaced with series of "X's." For the contrasting, spatial, condition, the change instead consisted of adding additional, superfluous, buttons to each button cluster, below the used buttons.

A concurrent working memory letter task was also introduced on the day of testing. As in previous studies [1–4], its function was to increase working memory load during task performance. Participants were presented with auditory stimuli in the form of randomly ordered letters spoken through the headphones at a rate of one letter every three seconds. A tone was presented randomly at intervals ranging from nine to 45 seconds, upon which the participants were directed to recall the last three letters in order and type them into the text box that appeared on the screen.

The experiment gave additional motivation to the users to complete Phaser trials quickly and as best they could. Performance in both the Phaser and memory updating tasks were scored, and the three users who earned the top scores were awarded cash prizes.

The two dependent variables were step completion times and error rates per step. However, some steps generated no meaningful completion times because of intervening events. For example, "stop charging" was clicked immediately after waiting for a thermometer-style gauge to fill up to a point indicating the battery was charged. so the time for this step was not determined by the user's actions

Results: Label Removal
Removing labels had negligible impact on user performance of the Phaser task, as measured by error rate or click time, per step (see Figure 2); main effect
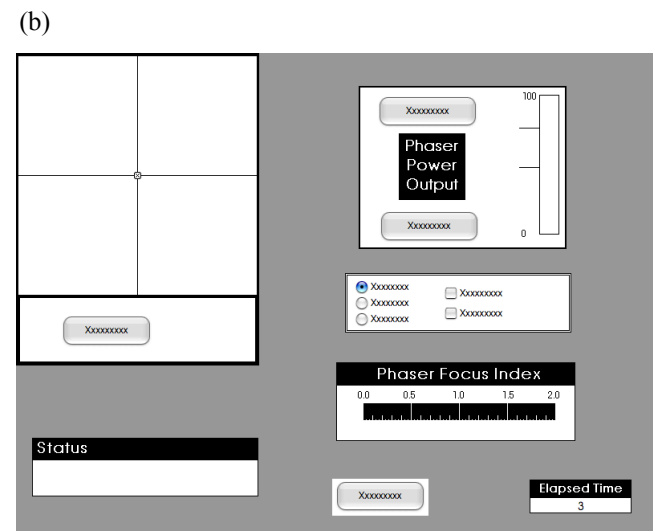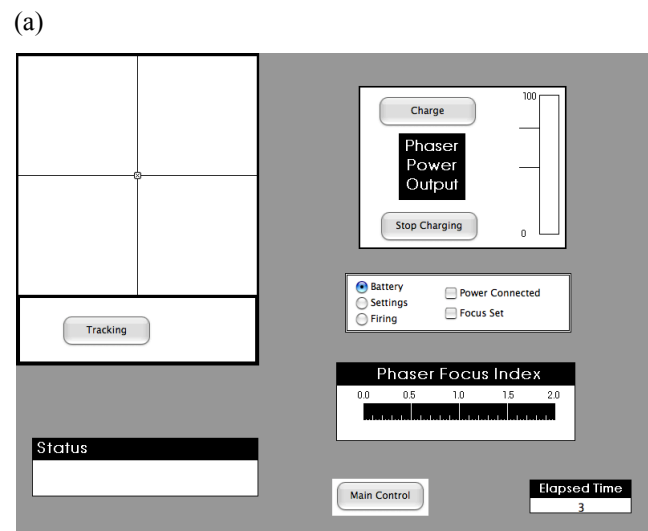
(a)

(b)

figure 1: Phaser control (a) and labels removed (b) interfaces.

of interface change on error rate, $F(1, 11) = 1.36$, $p = 0.27$, interaction with step number, $F(9, 99) = 1.67$, p 0.11, click time main effect, $F(1, 11) = 2.70$, $p = 0.13$, interaction with step number, $F(6, 66) = 1.78$, $p = 0.12$.

Results: Button Addition
In contrast to removing labels, simply adding extraneous buttons did impact performance (Figure 3):

change by step interaction on error, $F(9, 153) = 3.99$, $p < 0.001$. There was no similar effect on step completion times, indicating the effect on error rates was not simply a speed-accuracy trade-off.

Discussion
The lack of effect of label removal is not likely to simply be a matter of an ineffectual experimental paradigm, as the additional buttons condition illustrates. Nor does it
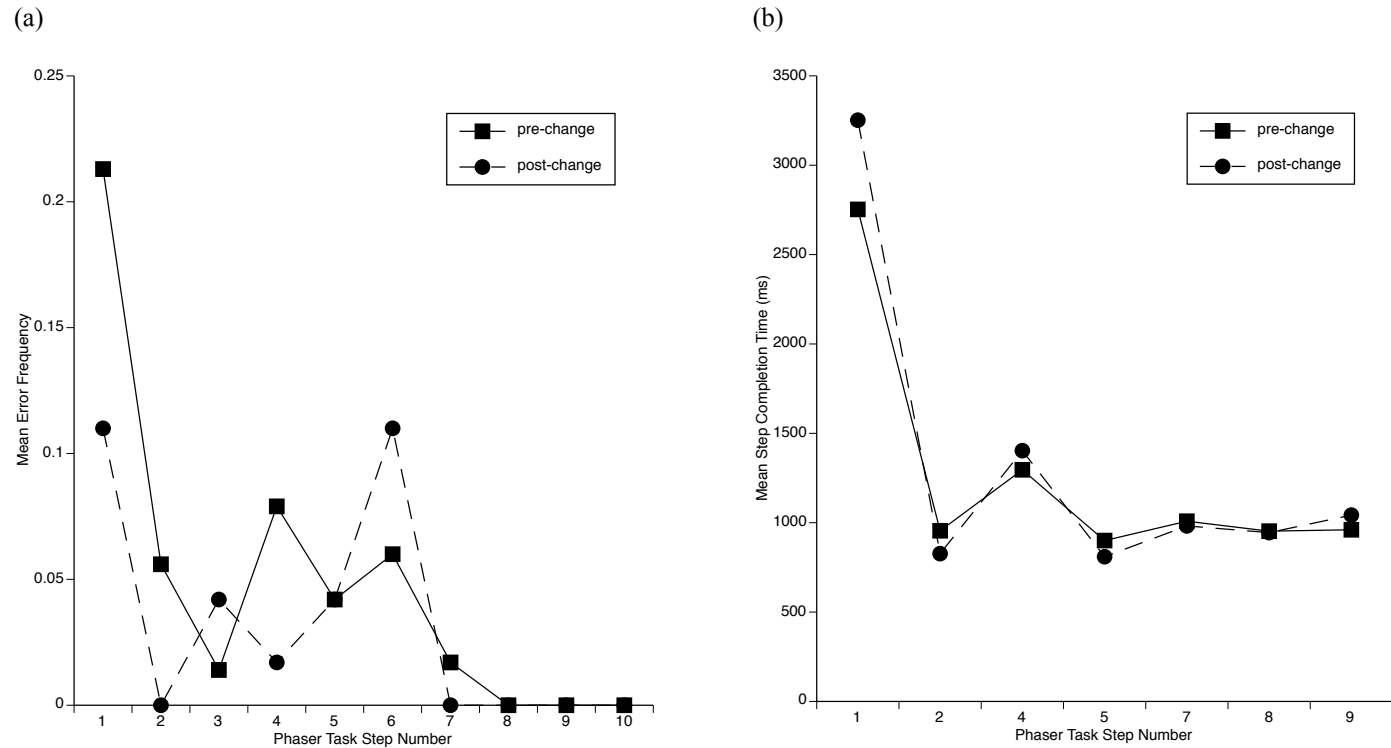
(a)

(b)

figure 2: Phaser error rates (a) and click times (b), labels crossed-out.

seem to be a lack of statistical power, given the results from the button addition condition.

Why should adding objects to the interface impact performance, particularly when those objects are never used in the task? The reason(s) is not clear to us, but it seems plausible that the appearance of the extra objects could cause some kind of perceptual/motor interference. Furthermore, it seems striking that something so seemingly inconsequential as adding

buttons that do not even work would disrupt task performance more than removing all button labels.

We therefore believe that users experienced in the use of a certain interface to perform a specific task generally do not read labels of control objects such as buttons,. We think this suggests a potential benefit that may be utilized: If experienced users do not read control labels, then substantial space within an interface may saved by simply leaving labels out or
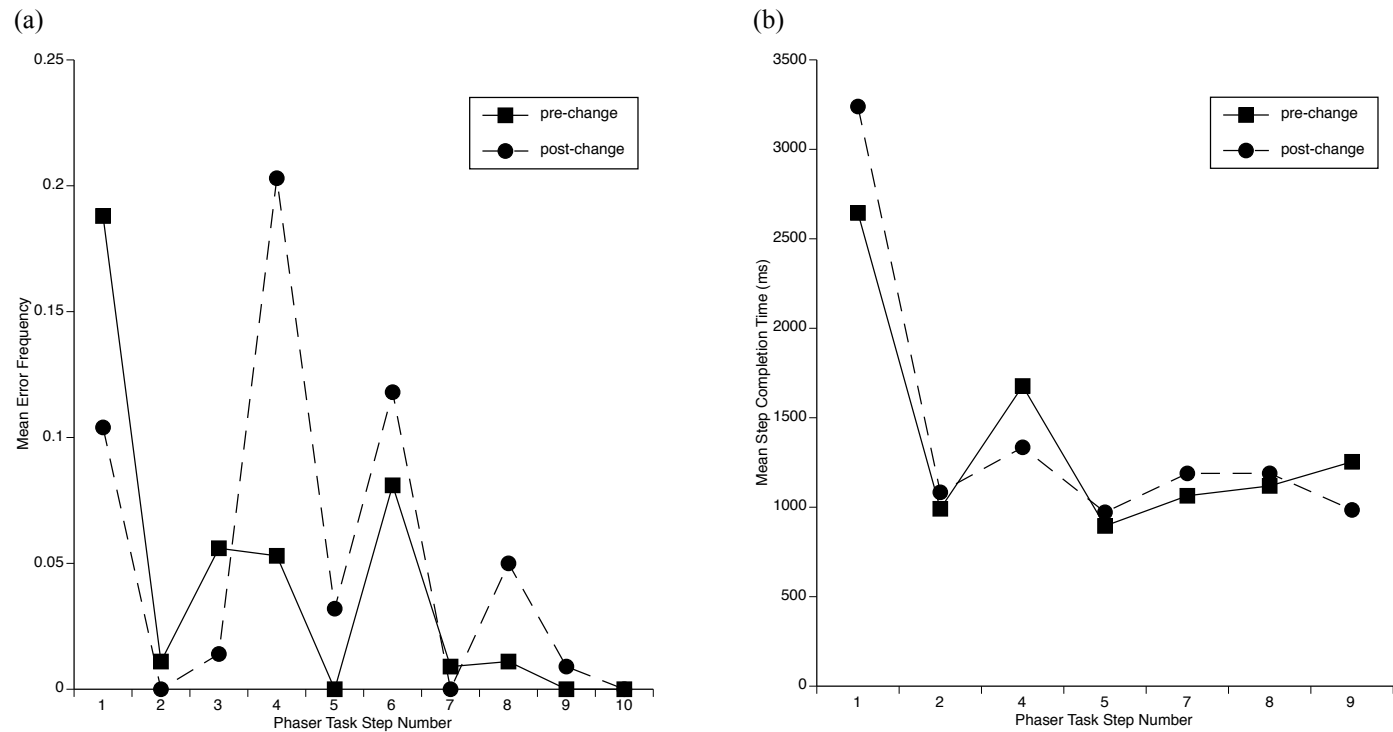
(a)

(b)

figure 3: Phaser error rates (a) and click times (b), extraneous buttons added.

reducing their size when conditions allow it (obviously not for novice users).

But if users do not read labels, how do they navigate an interface and know where each action needs to be performed? To answer that question, we are engaged in a program of research designed to map the spatial factors used in finding one's way around an interface. Preliminary data suggests that people may not simply memorize a global position of a control device within the entire display. They also seem to not rely upon landmarks within the interface that are not also controls themselves. We think experienced users of interfaces may be using some kind of hierarchical, relative representation of space that is closely coupled with their representations of action sequences.

There seems to exist a real hole in user interface models like GOMS where factors like spatial representations come into play. GOMS, for example, makes predictions about how long it should take to read a label, and may make gross estimates of visual search latencies. But it says nothing about what to expect of human performance when you move buttons within a cluster and also move button clusters. Nor does it tell us how a human can navigate an interface without verbal labels. To this end, we are pursuing investigations designed to tell us something about how people represent space within an interface and what ramifications those representations may have for human performance, particularly of skilled tasks. In particular, we intend to take an integrated cognitive modeling approach so that we may say something substantive about the complex interaction of perceptual-motor, cognitive, task structure, and interface structure factors interacting when people perform some task like the Phaser.

References
[1] Byrne, M. D., & Bovair, S. (1997). A working memory model of a common procedural error. *Cognitive Science, 21*, 31-61.

[2] Byrne, M. D., & Davis, E. M. (2006). Task structure and postcompletion error in the execution of a routine procedure. *Human Factors, 48*, 627-638.

[3] Chung, P. H. (2006). *Changing the Interface with Minimal Disruption: The Roles of Layout and Labels*. Doctoral dissertation, Rice University, Houston, TX.

[4] Chung, P. H., & Byrne, M. D. (in press). Cue effectiveness in mitigating postcompletion errors in a routine procedural task. *International Journal of Human-Computer Studies.*