

# When Using The Tool Interferes With Doing The Task

Susan S. Kirschenbaum  
 Naval Undersea Warfare Center Division  
 Code 2214, Building 1171/1  
 Newport, RI 02841  
 +1-401-841-3354  
 kirsch@c223.npt.nuwc.navy.mil

Wayne D. Gray, Brian D. Ehret, & Sheryl L. Miller  
 George Mason University  
 m/s 3f5  
 Fairfax VA  
 +1-703-993-1357  
 gray/behret/smillerb@gmu.edu

## ABSTRACT

How much time the user spends working on a task versus fiddling with the tool is an important aspect of usability. The concept of the *ratio* and *distribution* of *tool-only* operations to total operations is proposed to capture this aspect.

## Keywords

problem space, submarines, usability

## INTRODUCTION

Using a tool to perform a task requires the user to translate goals from a task space into subgoals and operations that are executed in a tool space (Moran, 1983; Payne, Squibb & Howes, 1990; Young, 1981). Task space goals can be as basic as withdrawing some amount of money from your savings account (GET-CASH) or as complex as employing a nuclear submarine to locate an enemy without being detected yourself (LOCATE-SUB).

Ideally, the task space subsumes the tool space. Sometimes, however, the tool itself becomes problematic. Then, the goal temporarily changes from performing the task to managing the tool. For example, if you insert your card in the ATM and it is returned immediately, you must change your focus from the task goal, GET-CASH, to the tool goal, GET-ATM-TO-ACCEPT-CARD. You might try a different orientation of the card in the slot; try a different ATM; or go inside the bank to see if the card is damaged. Indeed, for the moment, task-goals recede and tool-goals with their concomitant tool-only operations dominate. (When this recently happened to one of the authors, she was so busy trouble shooting the ATM card problem that she forgot to complete the task goal, GET-CASH.)

For tasks that are executed on the computer, all operations on the task, of necessity, entail an operation using the tool and can be referred to as *tool-task* operations. However, as with the ATM example, some operations are tool-only and have no direct relationship to accomplishing the task. The

current study examines the *ratio* and *distribution* of tool-only operations to all operations (tool-only + tool-task). The *ratio* shows the relative effort that the person spends on managing the tool rather than using it to perform the task. The *distribution* refers to how the tool-only operations are distributed among the tool-task operations. Our working hypothesis is that isolated tool-only operations represent minor tweaks that distract minimally from task performance. In contrast, sequences of tool-only operations represent problem solving in the tool-only space that may disrupt task performance.

## METHOD

(A fuller report can be found in Gray, Kirschenbaum & Ehret, in preparation.)

## Task & Participants

Our data were collected as part of an effort to investigate the pre-decision making processes of Submarine officers as they attempt to locate an enemy submarine. For this report we include an hour of data from 5 approach officers (AO). Each hour represents either two 30 min. scenarios or one 60 min. scenario. All AOs had recently served as commanding officers or executive officers on nuclear submarines.

## Procedure

The study used a high-fidelity simulation of the ocean environment that was capable of creating a generic submarine and targets. During the scenario, the AO controlled the situation by requesting information and ordering actions much as they would on their own ships. The simulation's keyboard and mouse was controlled by one of the experimenters who is an expert user of the simulation. (She is referred to in this report as *own ship operator*, or OS-op.)

## RESULTS

### Encodings

All verbalizations and some screen events were transcribed and segmented. Each of the transcripts was independently encoded by two raters. (The full details of this procedure are given in Gray et al., in preparation.)

### Reliability of encodings

Across the five-hour corpus, 2,653 segments were encoded as one of nine types of operations. Inter-rater reliabilities (Cohen's kappa, *K*, corrects for chance matches) ranged from

Permission to make digital/hard copies of all or part of this material for personal or classroom use is granted without fee provided that the copies are not made or distributed for profit or commercial advantage, the copyright notice, the title of the publication and its date appear, and notice is given that copyright is by permission of the ACM, Inc. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires specific permission and/or fee.

CHI '96 Companion, Vancouver, BC Canada  
 © 1996 ACM 0-89791-832-0/96/04..\$3.50

a low of  $K = .61$ ,  $Z = 16.51$  to a high of  $K = .75$ ,  $Z = 21.95$ .

### Task, tool, and NA

For current purposes, the nine types of operations form three categories: tool-only, tool-task, and NA. The 972 NA utterances were irrelevant to both tool and task and were excluded from further consideration.

### Tool-only ratio and distribution

Because of the nature of this task, all task operations used the tool. Thus, the tool-only ratio was computed as tool-only/total operations (where total includes both tool-only and tool-task operations). The mean tool-only ratio was 420/1661 or 0.25 (SD = .05) and ranged from 0.21 to 0.34. (The data to calculate these ratios are included as columns 2 and 3 of Table 1.)

**Table 1: Total operations (col. 2), ratio of tool-only/total (col.3), and run length of tool-only operations (cols. 4-8), by AO.**

AO	operations		Proportion tool-only operations by run length				
	total	ratio	1	2	3	4	≥5
s05	472	0.24	0.29	0.29	0.19	0.07	0.16
s06	355	0.23	0.26	0.27	0.11	0.15	0.21
s07	271	0.21	0.43	0.21	0.21	0.00	0.14
s08	346	0.29	0.26	0.12	0.21	0.04	0.36
s10	217	0.34	0.23	0.11	0.25	0.11	0.30
total/ mean	332	0.25	0.29	0.20	0.19	0.07	0.24
SD	21.6	0.05	0.08	0.08	0.05	0.06	0.09

As shown in Table 1, only 0.29 of the tool-only (TO) operations occurred in isolation; that is, surrounded by tool-task (TT) operations (e.g., TT->TO->TT). Most tool-only operations occurred in runs of two or more. For example, 0.20 of the tool-only operations occurred in runs of two (TT->TO->TO->TT); 0.19 in runs of 3 (TT->TO->TO->TO->TT); 0.07 in runs of 4, and 0.24 in runs of ≥5.

The mean tool-only ratio in this study was .25 and over half of the tool-only operations (0.51) occurred in runs of 3 or greater. Despite an expert OS-op and the generally low standard deviations, the highest to lowest tool-only ratios varied by a factor of 1.6 (from 0.21 for s07 to 0.34 for s10). The distribution of tool-only operations varied as well. For example, for runs of 5 or more the variation between AOs is 2.5 to 1, from 0.14 for s07 to 0.36 for s08. With only five subjects we cannot say whether these extremes are due to individual differences or due to situation specific events that arose while using the simulation.

### SUMMARY & CONCLUSIONS

We introduced the notions of the *ratio* and *distribution* of *tool-only* operations and we advanced the claim that these are important additions to how we measure and think about usability. Our task required the interaction between two people, a task expert (AO) and a tool expert (OS-op) and as such was well constituted to elicit the type of verbal protocol data required for our analyses.

As this is an exploratory effort, our conclusions must be modest. We have no idea whether the tool-only ratios we report, when compared to other tools and tasks, are high, low, or medium. Likewise, while we can report that long sequences of tool-only operations take attention away from task performance, we have no idea whether the distributions we found are high, low, or about average. To pursue these issues requires collecting much data over many different tool and task combinations. Likewise, we are intrigued by the finding that task experts interacting with the simulation via a tool expert should vary so greatly in both the tool-only ratio and distribution of tool-only operations. This second set of issues requires correlating task performance with tool disruptions. As a working hypothesis we assume that both higher ratios and longer runs of tool-only operations are correlated with poorer task performance. However, our modest contribution consists of asking the question, not providing the answer.

### ACKNOWLEDGMENTS

Susan S. Kirschenbaum's work has been jointly sponsored by Office of Naval Research (ONR) (Program element 61153N) and by Naval Undersea Warfare Center's Independent Research Program, as Project A10328. The work at George Mason University was supported in part by a grant from ONR (#N00014-95-1-0175) to Wayne D. Gray.

Approved for public release: Distribution statement A

### REFERENCES

- Gray, W. D., Kirschenbaum, S. S., & Ehret, B. D. (in preparation). *Subschemas for submariners: Cognitive models of situation assessment* (C-A-T-96-01): George Mason University.
- Moran, T. P. (1983). Getting into a system: External-internal task mapping analysis, *Proceedings of CHI'83 Conference on Human Factors in Computing Systems*, (pp. 45-49). New York: ACM.
- Payne, S. J., Squibb, H. R., & Howes, A. (1990). The Nature of Device Models: The Yoked State Space Hypothesis and Some Experiments with Text Editors. *Human-Computer Interaction*, 5(4), 415-444.
- Young, R. M. (1981). The Machine Inside the Machine: Users' Models of Pocket Calculators. *International Journal of Man-Machine Studies*, 15(1), 51-85.