In theProceedings of the 7th International Conference on Human-Computer Interaction jointly with the 13th Symposium on Human Interface (Japan), HCI International '97, August 24-29, 1997, San Francisco, CA.

Supporting Task Performance: Is Text or Video Better?

Brian E. Norris and William B.L. Wong

Multimedia Systems Research Laboratory
Department of Information Science, University of Otago
PO Box 56, Dunedin, New Zealand.

1. INTRODUCTION

Multimedia technology allows a variety of the presentation formats to portray instructions for performing a task. These formats include the use of text, graphics, video, aural, photographs, used singly or in combination (Kawin, 1992; Hills, 1984; Newton, 1990; Bailey, 1996). As part of research at the Multimedia Systems Research Laboratory to identify a syntax for the use of multimedia elements, an experiment was conducted to determine whether the use text or video representations of task instructions was more effective at communicating task instructions (Norris, 1996). This paper reports on the outcome of that study.

The repair and assembly environment of a local whiteware manufacturer provided the study domain. The task chosen for the study was the replacement of a heating element in a cooktop oven. As there were no task instructions available from the manufacturer, the study was conducted in two phases: Phase I was a cognitive task analysis of service technicians to determine the steps as well as the cues and considerations of the assembly task; and in Phase II we evaluated the text and video representations of the task instructions. The next sections briefly describe the methodology and the results from the experiment.

2. METHODOLOGY

In the first phase of the study, the Cognitive Task Analysis (CTA) resulted in the identification of important cues and considerations that service technicians used to replace the element in the cooktop (Klein 1993, Vicente, 1995). Two techniques were used in the CTA: The first was controlled observation in which the service persons were observed performing the task to identify the overt steps in the process. 12 steps were identified in the process. The second technique was a retrospective interview using cognitive probes to identify important cues and considerations of the service persons during the assembly of the heating element. These cues, e.g. little tricks to pop open the back of the, were incorporated into the instructions. In the second phase, the experiment consisted of two sets of trials to evaluate the text and video formats (Dumas & Reddish, 1993). A deliberately simplified user interface that was common to both trials was used to present the instructions for the 12 task steps. The instructions were presented in the same locations on the user interface as either text or as video.

As this was a pilot study ten participants were involved in the two trials. There were a total of six males and four females, divided equally between the two test conditions. They were all undergraduate students with little or no computing background. None of the participants had any prior experience in replacing cooktop elements.

In each trial, the participants were asked to observe the text or video-based instructions for one step and then to perform that step before proceeding to the next step. Participants were told they could refer or re-play the instructions for each step as many times as they consider necessary in order for them to correctly perform the task. The participants' performance was observed by the experimenter sitting in the same room, and was measured on the following dimensions (Nielsen, 1993): (i) Efficiency, (ii) Number of errors, and (iii) Learnability.

Efficiency was measured in terms of time taken to perform each step of the task. For each step the time that was recorded was in two parts. The first part was the time the participant took to look at the information on the screen. This time is termed "time looking". The second part was the time taken to actually perform the step. This component of time is referred to as "time doing". An error is defined as any action that deviated from that specified in the task instruction. The participant was observed performing the individual steps of the task, and any time the participant did not perform the task exactly as the computer system had instructed, this was counted as an error. Learnability was measured in terms of repeats, i.e. the number of times a participant referred to the instructions for the step in order to carry out that step. If the participant turned to look at the instructions on the screen again, this was counted as a repeat.

3. RESULTS

The results from the experiment indicate that the video format was more effective than the text format for supporting task performance. This is briefly reported below.

3.1 Efficiency.

The average total time (time looking plus time doing) for completing the entire task was less for the text format than for the video format. But on closer examination, the results indicate that actual task execution is better with video instructions that with text. Figure 1 shows that while it took longer to view the video instructions than the text instructions, participants using video were able to complete their tasks in 12% less time than those presented with text. While not a significantly larger advantage, the other measures to be presented next suggest there is a quality difference in task performance through the use of video.

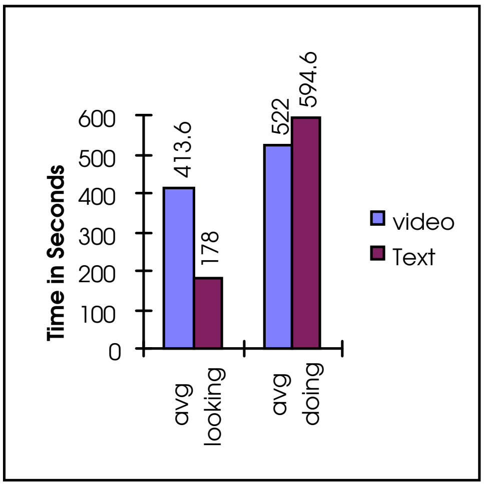

Figure 1 Average Times for Looking and Doing

3.2 Number of errors.

Errors give an indication of how well the instructions for the element replacement procedure is understood and replicated. Participants presented with the video instructions made an average of 2.6 errors during the entire trial, compared with an average of 5.6 errors committed by participants who were presented with text instructions. See Figure 2.

3.3 Learnability

Measuring how quickly a participant learnt the replacement procedure is the purpose of this dimension. Participants who were presented with video instructions needed to refer to the video only on the average 0.8 times during the entire 12-step task. Whereas participants presented with text needed to refer to the instructions on an average of 3.8 times, or about four times more often than the participants who used the video instructions. Figure 3 illustrates this.

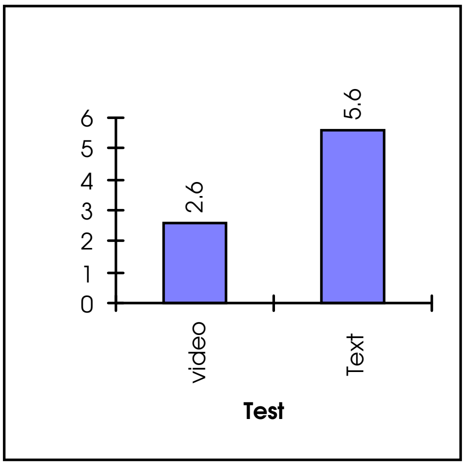

Figure 2 Average Number of Errors

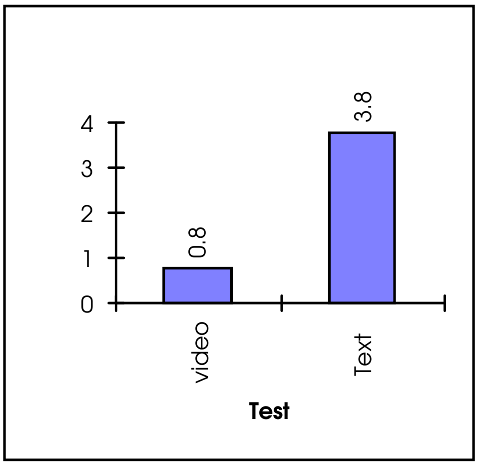

Figure 3 Average Number of Repeats

4. DISCUSSION

Although the results from the experiment seem to suggest that higher efficiency can be obtained through the presentation of text-based instructions, the effectiveness of the two presentation formats must also be viewed in terms of process quality. Text-based instructions resulted in more than twice the error rate experienced then when using video instructions. While we have not investigated the cause of the difference in error rates, it is plausible that the errors are a result of different cognitive processing: Text requires more decoding and interpretation, while video relies on the more powerful perceptual-cognitive systems to understand how a procedure is carried out. The number of repeats represent how quickly one is able to learn the procedure for replacing the heating element. Learning procedures appears to be about four times faster with video instruction than text. Again, the cause of this learning advantage was not investigated in the study, but it the findings suggests that procedural type information is better assimilated by a viewer by showing (video) than by describing (text).

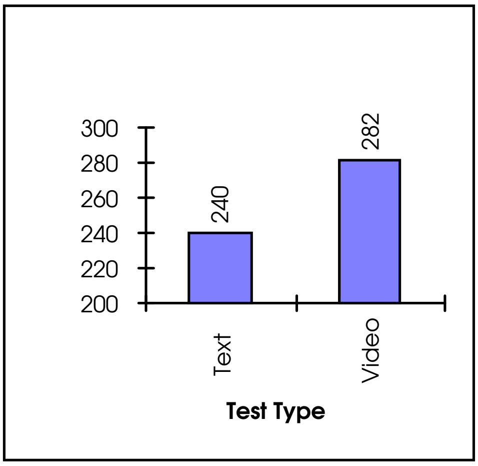

Figure 5 Productivity

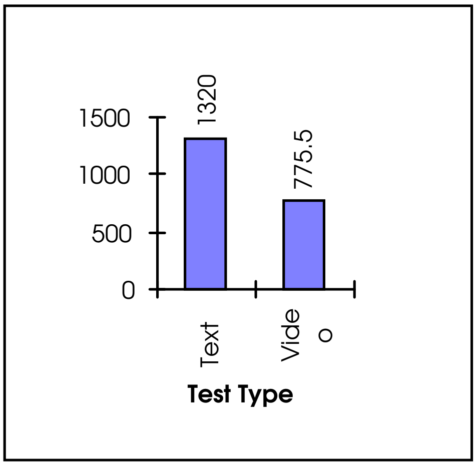

Figure 6 Corrections

To facilitate discussions with our clients at the whiteware manufacturing plant, these results were translated into more meaningful figures to them: Production per week, Corrections needed per week, and Time between interruptions. These measures suggested that a worker working with text-based instructions would produce about 240 units while his or her counterpart working with video-based instruction would produce about 280 units (Figure 5). This is a general improvement in production capacity. This representation was relevant to our clients as their factory floor workers are rotated through different jobs during the week, and hence some re-learning will be needed. The real gains to our client are, however, not in increased production but in the dramatically fewer corrections as Figure 6 suggests. Based on the same number of units produced, the number of corrections needed per week per worker using text was estimated at 1300 corrections, while the expected number of corrections for the worker using video is estimated at 800 corrections. This translates into reduced warranty and product recall costs. Another factor that affects productivity is the time between interruptions. It appears that workers using text-based systems can expect to interrupt their task to refer to the instructions by as much as once every 160 seconds. However, workers using a video-based system are expected to be interrupted only once every 650 seconds to review their instructions. While these figures do not account for learning, they have been useful in demonstrating the differences between text- video-based task support.

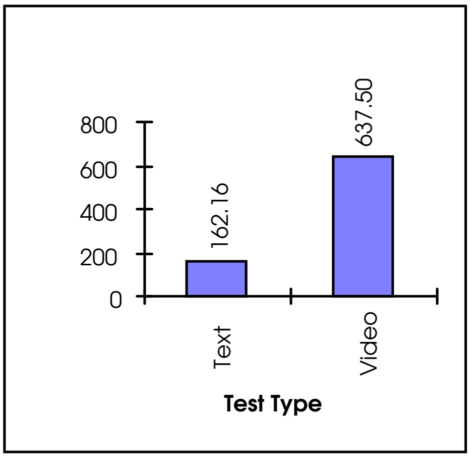

Figure 7 Time Between Interruptions

5. CONCLUSION

While the samples used were small, the results are indicative of better task support performance using video than text-based instruction. More work with larger samples is being planned to improve the generalisability of these results. These planned studies would also investigate why such differences in performance exist, and how these results could be explained within the context of a multimedia syntax.

REFERENCES

Bailey R W. (1996). Human Performance Engineering: Designing High Quality Professional User Interfaces for Computer Products, Applications and Systems. Third Edition. Prentice-Hall Inc. New Jersey.

Dumas J S, Redish J C. (1993). A Practical Guide to Usability Testing. Ablex Publishing Corporation. New Jersey.

Hills P J. (1984). Video Production in Education and Training. Croom Helm Ltd. Kent.

Kawin B F. (1992). How Movies Work. University of California Press. California.

Klein G. (1993). Naturalistic Decision Making: Implications for Design. Klein Associates Inc. Ohio.

Newton D P. (1990). Teaching with Text: Choosing, Preparing and Using Textual Materials for Instruction. Kogan Page Limited. London.

Nielsen J. (1993). Usability Engineering. Academic Press Inc.

Norris, B. (1996). A comparative analysis of a Task Support Environment: Text vs Video. A Pilot Study. Unpublished dissertation, Department of Information Science, Otago University.

Vicente K J. (1995) Task Analysis, Cognitive Task Analysis, Cognitive Work Analysis: What's the Difference? Proceedings of the Human Factors and Ergonomics Society 39th Annual Meeting. Pages 534-537.