THE EFFECT OF A SPEECH DISCRIMINATION TASK ON NAVIGATION IN A VIRTUAL ENVIRONMENT

Bruce N. Walker and Jeffrey Lindsay
Sonification Lab, School of Psychology
Georgia Institute of Technology,
Atlanta, GA 30332-0170
bruce.walker@psych.gatech.edu

If the input from the visual system is unavailable (e.g., damage to the optic nerves or smoke in a burning building), navigating and avoiding obstacles becomes very challenging. It is therefore desirable to develop a navigation aide for use where visual input has become unavailable. There is a small body of research concerning such navigation aides and their efficacy. However, many issues that may have serious human factors repercussions for such a system are unexplored. This study was conducted in order to examine the effect of an attentionally demanding distractor task on wayfinding performance with an audio only navigation aide, in this case the System for Wearable Audio Navigation (SWAN). A difficult secondary speech task reduced efficiency in navigation, as users switched attentional resources to the speech task. Practical applications are discussed.

INTRODUCTION

When vision is unavailable, navigating through the environment becomes very challenging. This applies not only to the 161 million visually impaired people worldwide (Resnikoff et al., 2004), but also to firefighters, police, and first responders who operate in low-vision situations. There is clearly a role for assistive technology navigation aides for use in such situations.

The SWAN system (see Walker & Lindsay, 2006), has an auditory interface composed of spatialized, non-speech auditory icons that aid users in navigation and awareness of features in the environment. Sounds in SWAN are classified as beacons, objects, and surface transitions. Beacon are placed (virtually) at waypoints along a route. Using this trail of spatialized beacon sounds, SWAN is able to guide users through their environment. Though the system was developed using solid human factors principles and incorporating the limited existing research on sound design for such systems (e.g., Tran, Letowski, & Abouchacra, 2000), there remain important usability questions to be explored. One of these issues is how multitasking while using SWAN will affect movement and navigation performance.

METHODS

Participants

Thirty undergraduates (15 male, 15 female, mean age = 21.5) participated for course credit. All reported normal hearing, and had no previous experience with the tasks.

Navigation task

The SWAN navigation task consisted of moving quickly and accurately along a pre-defined path through a virtual environment (VE). There were no visual cues, and only auditory beacons to guide the listener. This study included the two beacon sounds that led to the best performance in previous studies: a pink noise burst and a sonar pulse. Full details of the sounds and task are in Walker and Lindsay (2006).

Speech discrimination: “Ready Charlie” task

Periodically during a participant’s transit along a path, she would be required to complete a secondary speech discrimination task. This is representative of moving towards a destination and

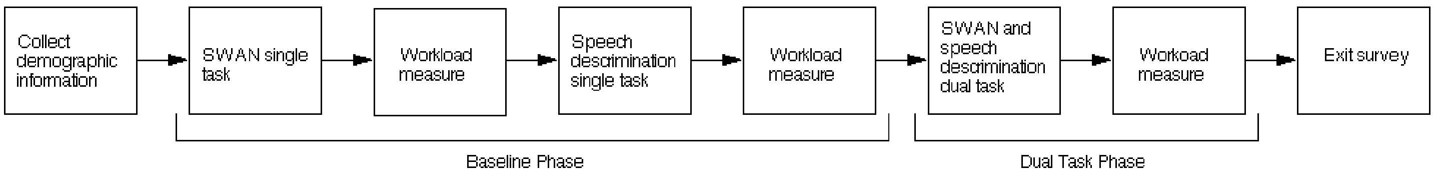

Figure 1. *Study outline. Shows the ordering of each part in the study.*

having a conversation, or cell phone call, or radio communication occur at some point(s) along the way. The speech task stimuli were drawn from the Coordinate Response Measure speech corpus (Bolia, Nelson, Ericson, & Simpson, 2000). The phrases have the format “Ready (call sign) go to (color) (number) now,” such as “Ready Charlie go to green four now.” The task was to monitor several simultaneous streams of this type of stimuli for a specific target call sign, similar to Brungart, Simpson, Ericson, and Scott (2001), and verbally report which color and number were heard.

Apparatus and Procedure

The navigation task was conducted in a VE built using the Simple Virtual Environments (SVE) software package (Kessler, Kooper, & Hodges, 1998). Participants’ head position and orientation were recorded using a head tracker. A modified joystick controlled a participant’s forward and backward movement. A Creative Labs SoundBlaster Extigy external sound card performed all 3D audio rendering. See Walker and Lindsay (2006) for complete details of the apparatus. Note that the VE has been shown as an excellent (and safer) surrogate for physical movement (Tran, Letowski, & Abouchacra, 2000; Walker & Lindsay, 2006). Sound stimuli from the VE (the beacons) were mixed with the speech task sounds, and played through Sony MDR-7506 headphones. Experimental sessions were videotaped and time stamped. A computerized version of the NASA-TLX (Hart & Staveland, 1988) was administered to assess subjective workload at various stages of the experiment.

Figure 1 depicts the flow of the experiment. In the Baseline phase, participants performed each of the tasks alone. The speech discrimination task stimuli were presented sequentially with a 7 s break between trials (see Brungart et al., 2001). In the Dual Task phase of the study participants performed the two tasks simultaneously. After each of the three task sets (navigation, discrimination, and navigation+discrimination) participants completed a subjective workload measure.

RESULTS

Participants’ position and time were analyzed to compute the dependent variables of path efficiency and time efficiency, as in Walker and Lindsay (2006). These normalized metrics account for different path lengths across maps. In the speech discrimination task, speed and accuracy were used to measure performance, as in Brungart et al. (2001).

The results of the SWAN single task were plotted in terms of time efficiency and path efficiency, split by beacon sound (left portions of Figures 2 and 3). The noise beacon (solid dark lines in Figures 2 and 3) generally led to more efficient performance. Also, participants clearly improved on both metrics with practice. A MANOVA was conducted with beacon type and performance across trials (the effect of practice) as factors. The two dependent measures were path efficiency and time efficiency. Significant main effects were found for both beacon type, $F(2,27) = 4.413$, $p < .05$, *Wilk’s Lambda* = .754 and practice, $F(8,21) = 114.590$, $p < .05$, *Wilk’s Lambda* = .022, moderated by a significant beacon sound x practice interaction, $F(8,21) = 3.106$, $p < .05$, *Wilk’s Lambda* = .458. Examining each dependent measure singly found a significant effect of practice for time efficiency, $F(4,112) = 201.601$, $p < .05$, and path efficiency, $F(4,112) = 30.494$, $p < .05$, but only time efficiency showed a significant beacon type x practice interaction, $F(4,112) = 3.559$, $p < .05$.

The goal of the SWAN single task phase of the study was to get participants to asymptotic performance (see Figures 2 and 3). Then, comparing the last two trials in the SWAN single task to performance during the dual task showed a clear decrease in performance during the dual task phase. The different beacon types showed an unequal rate of decline in performance during the dual task, with the noise beacon not always showing better performance (as was the case in the single task phase). A MANOVA was conducted to test the significance of these observations, with beacon sound type, single versus dual task, and practice as factors. A significant main effect of single versus dual task was found, $F(2,27) = 52.354$, $p < .05$, *Wilk's Lambda* = .205, moderated by a significant interaction between the single/dual task condition and beacon sound, $F(2,27) = 4.046$, $p < .05$, *Wilk's Lambda* = .769. No significant main effect of beacon type was found. Further analysis examining each dependent measure singly found a significant effect of single versus dual task for path efficiency, $F(1,28) = 43.851$, $p < .05$, but no such effect for time efficiency. A significant interaction for single versus dual task and beacon type was also found for time efficiency, $F(1,28) = 7.324$, $p < .05$.

In the speech discrimination ("Ready Charlie", or RC) task, accuracies during the first dual task block (mean = 15%, SE = 1%) started out lower than the end of the single task phase (mean = 24%, *SE* = 1% and mean = 30%, *SE* = 2% in Blocks 1 and 2 respectively) However, by the second block of dual task trials accuracy (mean = 30%, *SE* = 2%) had risen back to Block 2 of the single task (see Figure 4). The reaction time data tell a similar story.

NASA-TLX measures showed that the workload during the SWAN single task (mean = 39.4) was significantly less than that reported during the RC single task (mean = 68.0) and the dual task (mean = 68.9).

DISCUSSION

Performance in the SWAN single task was similar to performance found in prior work (Walker & Lindsay, 2006). The noise burst was once again found to be a better auditory navigation beacon than the sonar pulse, presumably due to the broadband spectrum of the noise aiding in sound localization. Also, excellent navigation performance levels here again demonstrate that people are capable of

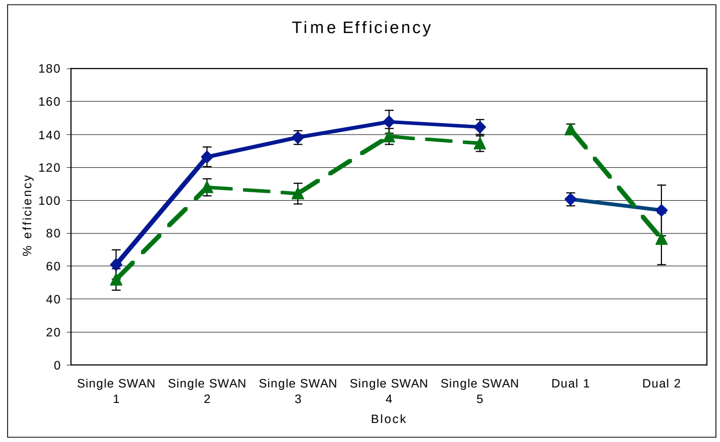

Figure 2. *Time efficiencies by beacon type for the SWAN navigation task in the single and dual task phases. The noise beacon is represented by the darker solid line, and the sonar beacon is represented by the lighter dashed line.*

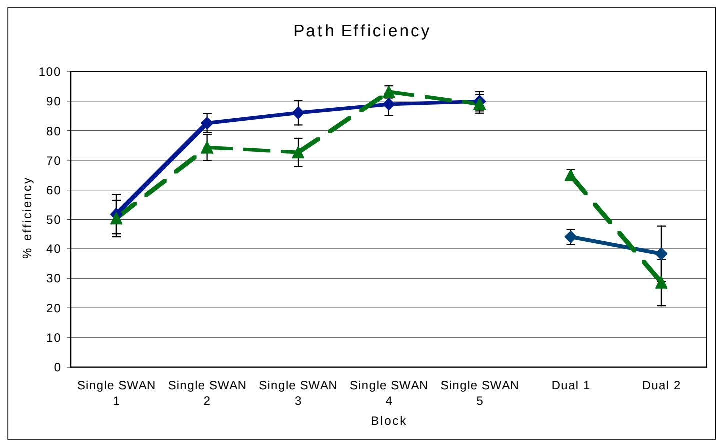

Figure 3. *Path efficiencies by beacon type for the SWAN navigation task in the single and dual task phases. The noise beacon is represented by the darker solid line, and the sonar beacon is represented by the lighter dashed line.*

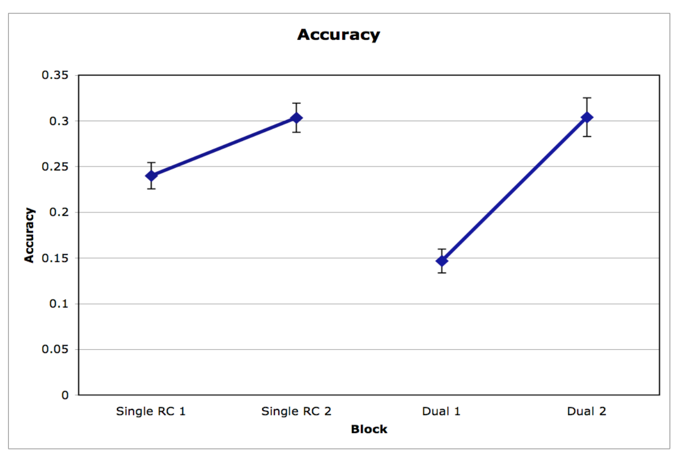

Figure 4. *Speech discrimination ("Ready Charlie") task accuracy, in percent correct.*

performing well on the SWAN navigation task with little practice.

In looking at the effect of the secondary task on performance of the SWAN task, a decrease in performance was observed in the time and path efficiencies of the SWAN task as well as the accuracy and response time in the speech discrimination task. The largest drop in performance was observed in the SWAN task. This was unexpected, given that the stimuli in the RC task are transient; it had been hypothesized that performance in the RC task would suffer more than the SWAN task (which has persistent stimuli). One possible explanation for this result is that participants, who were instructed to do as well as possible on both tasks during the dual task phase, made a decision (conscious or not) to sacrifice SWAN performance in order to allow more cognitive resources to be devoted to the RC task.

The subjective workload ratings were found to be highest for the dual task portion of the experiment. The reported workload for the RC single task was almost as high as that of the dual task. This suggests that participants, having performed both of the single tasks, did not find the dual task to be significantly harder than the RC task alone. It could indicate that a majority of the workload in the dual task phase was due to the RC task.

There are important practical implications of these results for the use of the SWAN interface. While participants found the SWAN task had a relatively low workload, it still required enough cognitive resources (or enough of the same cognitive resources) that introducing a secondary auditory task interfered with a user's ability to perform with the interface. This is obviously a potential cause for caution in further design and implementation of such interfaces. However, there are some important caveats. Firstly, it should be noted that the RC task used here was a very hard task, which is made apparent both in participants' reported workloads as well as their measured performance. In many common SWAN usage scenarios, it is unlikely that a user would be required to perform another auditory task of this difficulty while using the SWAN, especially for any extended periods. Additionally, users in some (but not all) cases would simply be able to stop moving and focus on the other auditory task and afterwards they could continue navigating with the SWAN interface. Also while performance declined during the dual task phase, it is not clear that this drop would have any practical significance in terms of a user's ability to navigate using the SWAN interface. Performance in terms of path and time efficiencies in SWAN single task situations indicate that users are extremely adroit at following the paths indicated by the interface. The moderate drop in performance during the dual task phase observed in this study may be relatively negligible in terms of real world user performance.

REFERENCES

Bolia, R. S., Nelson, W. T., Ericson, M. A., & Simpson, B. D. (2000). A speech corpus for multitalker communications research. *Journal of the Acoustical Society of America*, 107, 1065-1066.

Brungart, D. S., Simpson, B. D., Ericson, M. A., & Scott, K. R. (2001). Informational and energetic masking effects in the perception of multiple simultaneous talkers. *Journal of the Acoustical Society of America*, 110(5): 2527-2538.

Hart, S. G., & L. E. Staveland (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P. A. Hancock and N. Meshkati (Eds.) *Human Mental Workload*. Oxford, England: North-Holland (pp. 139-183).

Kessler, D. R., Kooper, R., & Hodges, L. (1998). *The Simple Virtual Environment Library: Version 2.0 User's Guide. Graphics, Visualization, and Usability Center Technical Report GIT-GVU98-13.*

Resnikoff, S., Pascolini, D., Etya'ale, D., Kocur, I., Pararajasegaram, R., Pokharel, G., et al. (2004). Global data on visual impairment in the year 2002. *Bulletin of the World Health Organization*, 82, 844-885.

Tran, T. V., Letowski, T., & Abouchacra, K. S. (2000). Evaluation of acoustic beacon characteristics for navigation tasks. *Ergonomics*, 43(6), 807-827.

Walker, B. N., & Lindsay, J. (2006). Navigation performance with a virtual auditory display: Effects of beacon sound, capture radius, and practice. *Human Factors, 48*, 286-299.