Journal of Applied Research in Memory and Cognition 2 (2013) 192–199

Contents lists available at ScienceDirect

Journal of Applied Research in Memory and Cognition

journal homepage: www.elsevier.com/locate/jarmac

Which modality is best for presenting navigation instructions?

Alice F. Healy[a,*], Vivian I. Schneider[a], Blu McCormick[a], Deanna M. Fierman[a], Carolyn J. Buck-Gengler[a], Immanuel Barshi[b]

[a] *Department of Psychology and Neuroscience, University of Colorado Boulder, United States*
[b] *Human Systems Integration Division, NASA, Ames Research Center, United States*

ARTICLE INFO

Article history:
Received 18 April 2013
Received in revised form 25 July 2013
Accepted 25 July 2013
Available online 7 August 2013

Keywords:
Modality
Navigation instructions
Multiple resources
Memory

ABSTRACT

Three experiments involved college students receiving and following instructions of various lengths for navigating in a three-dimensional space displayed on a computer screen. The purpose was to evaluate which is the best modality for presenting navigation instructions so that they can be executed successfully. Single modalities (*read*, *hear*, and *see*) were considered along with dual modalities presented simultaneously or successively. It was found that when there were differences between single modalities, generally execution accuracy was best for *see* and worst for *read*. Information presented in two modalities did not yield better accuracy than information presented twice in a single modality. Also, the ordering of modalities depended on the extent of practice. Thus, presentation modality does not have a consistently large effect on receiving and following navigation instructions. Repetition and the amount of practice are much more important variables than is presentation modality in determining how well navigation instructions are followed.

Following navigation instructions is a common task in everyday life where individuals must navigate through buildings, neighborhoods, construction sites, parking structures, and space. Consider, for example, getting directions about where to find merchandise in a large multi-story department store (e.g., Macy's or Harrods). In other, more serious cases, following navigation instructions can have life-or-death consequences. For instance, errors in communication between air traffic control (ATC) and pilots can have severe repercussions. Even small differences in the accuracy of following navigation instructions could lead to serious accidents. For example, if ATC tells pilots to turn right, and they turn left instead, a collision might occur resulting in casualties. Finding ways to minimize these communication errors is thus a critical question for research. One concern is which modality would be best to present navigation instructions so that the recipient can understand, remember, and execute those instructions with minimal errors. Messages are usually presented by ATC in the auditory modality, with pilots hearing the messages as spoken commands. However, with current technology (e.g., *data link*; Kerns, 1991; Lancaster & Casali, 2008) the visual modality can be used instead, with pilots reading the messages as written commands. A third possibility also involves visual presentation, but in this case without words, with commands shown as pictures or symbols (see, e.g., Tversky, 2003, for a discussion of the use of such graphics). Pilots often see electronic displays and navigational charts, but ATC does not currently present navigation instructions in that manner. Another option would be to present navigation instructions in more than one modality either simultaneously or sequentially. We consider each of these alternatives in the present study, in which we use an experimental paradigm simulating a communication situation in which individuals such as pilots receive and then follow navigation instructions like those from ATC.

In the present study, we compared the comprehension of navigation instructions that were heard, read, or seen by the subjects, with the three presentation modes equated in message presentation time to permit a pure assessment of modality effects. In contrast, in the actual implementation of data link, the visual presentation is essentially permanent. When this visual data link procedure was compared to an auditory procedure (Helleberg & Wickens, 2003), the visual mode was better than the auditory mode in terms of following navigation instructions, presumably because of the differences in the permanence of the presentation. However, a more recent study involving data link by Lancaster and Casali (2008) found a disadvantage for the visual mode relative to the auditory mode in terms of both increased time to respond and ratings of workload. Furthermore, in a data link study McGann, Morrow, Rodvold, and Mackintosh (1998) found an advantage for the auditory mode over the visual mode in terms of actions related to clarification of navigation messages. Moreover, without the

* Corresponding author at: Department of Psychology and Neuroscience, Muenzinger Building, 345 UCB, University of Colorado, Boulder, CO 80309-0345, United States. Tel.: +1 303 492 5032; fax: +1 303 492 8895.
E-mail addresses: alice.healy@colorado.edu, ahealy@psych.colorado.edu (A.F. Healy).

2211-3681/$ – see front matter
http://dx.doi.org/10.1016/j.jarmac.2013.07.004

differences in permanence between the auditory and visual modes, Wickens, Sandry, and Vidulich (1983) found that pilots receiving navigation instructions performed better with auditory than with visual presentation. An advantage for auditory compared to visual presentation modalities has also been found in many studies involving memory for word lists across short or long retention intervals (e.g., Crowder & Morton, 1969; Gardiner, Gardiner, & Gregg, 1983; Goolkasian & Foos, 2002; Murdock, 1967). Explanations for these modality effects include Penney's (1989) suggestion that auditory and visual items are processed in separate streams, with a code for acoustic material longer lasting than that for visual material. Nairne's (1988) feature model provides an alternative explanation in terms of subjects' general preference for auditory over visual features as recall cues.

When the material to be remembered consists of navigation instructions, another important consideration is that the movement space is visual so that auditory presentation provides a mixed mode, which has been found to reduce cognitive load (e.g., Mousavi, Low, & Sweller, 1995). Related to this finding is the demonstration by Brooks (1967) of a conflict between reading verbal messages and imagining the spatial relations that the messages describe; such a conflict was not found when the verbal messages were heard rather than read. Visual messages presenting the spatial information directly or with symbols, rather than with words, might be another way to avoid this conflict. Such visual messages would also benefit from the well-established picture superiority effect (e.g., Nelson, Reed, & Walling, 1976; Paivio & Csapo, 1973; Shepard, 1967; Snodgrass & McClure, 1975), whereby there is better memory for items presented as pictures than as words. The advantage for such visual messages is also consistent with the stimulus/central-processing/response (S–C–R) compatibility model of a pilot's task based on the assumption that spatial tasks (e.g., moving in a space) are more compatible with visual inputs than with auditory inputs (Wickens & Hollands, 2000; Wickens, Vidulich, & Sandry-Garza, 1984).

Many existing psychological theories, including the theory of multiple resources in cognitive processing (e.g., Wickens, 2008), the dual-coding theory of memory (Paivio, 1971, 1991), and theories of multimedia learning (Mayer, 2001; Mayer & Gallini, 1990; Mayer & Sims, 1994), predict that two modalities would be better than one for learning information. (It should be noted, however, that the theory of multiple resources is relevant to concurrent processing of the various resources rather than to sequential processing of them, and the theories of multimedia learning do not predict superior performance for two modalities when those modalities compete for attention.) Thus, presenting navigation instructions in more than one modality simultaneously, and perhaps also sequentially, might be expected to improve comprehension and memory for the instructions relative to presenting them in a single modality. However, in their study of the data link procedure, Helleberg and Wickens (2003) compared a redundant condition, in which both visual and auditory information was presented simultaneously, to single-modality conditions, in which information was presented in either a visual or an auditory format. They found that flight task performance (i.e., measured as deviations from ATC instructions) was best for the visual condition, worst for the auditory condition, and intermediate for the redundant condition. They attribute the poor performance in the auditory and redundant conditions to auditory preemption effects, which interrupt the continuous visual tasks. In a related study, Wickens, Goh, Helleberg, Horrey, and Talleur (2003) also found no advantage in lateral or vertical tracking for a redundant condition, although auditory presentation was modestly superior to visual presentation in that case, attributed to the head-down nature of the visual scanning. Also, Lancaster and Casali (2008) found no advantage for a redundant condition over a pure auditory condition in terms of time to respond and workload, with both of those conditions superior to the pure visual condition by those measures. In a recent meta-analysis, Lu et al. (in press) point out that although redundant modality combinations have traditionally been considered beneficial, there can be a cost reflecting competition for attentional resources. Their analysis revealed that redundant auditory-visual tasks were generally more accurate but slower than tasks using a single modality (auditory or visual). They attribute this pattern of results to the fact that redundancy helps guarantee security (i.e., redundancy provides more opportunities for the information to be noticed) but does so at the expense of efficiency. These earlier studies of redundancy were restricted to simultaneous presentation of multiple modalities. Some of the costs of redundancy (e.g., those concerning interruption and competition for attention) would not accrue when the modalities are presented sequentially instead of simultaneously, at least when there is a control for the total duration of the presentation.

To investigate the relative benefits of various modalities and combinations of modalities for presenting navigation instructions, we used an experimental laboratory paradigm that isolates the navigation task and has already revealed many relevant findings (e.g., Schneider, Healy, & Barshi, 2004; Schneider, Healy, Barshi, & Kole, 2011). In our experiments, college students see a computer screen showing a grid of four matrices stacked on top of each other and are given messages instructing them to move in the grid by clicking with a computer mouse (see Fig. 1). This task is analogous to the aviation task as well as to the task mentioned in the Introduction of getting and following directions about where to find merchandise in a large multi-story department store (with the matrices corresponding to different floors). The measure of performance is accuracy in following the messages. Because of the large effects of message length on memory for navigation commands (e.g., Loftus, Dark, & Williams, 1979), the messages vary substantially in length. In some cases, simulating pilot behavior, students are also asked to repeat back the instructions before executing them. This paradigm differs from the pilot task in many important respects (e.g., experienced pilots presumably have much more extensive practice and there are high visual demands on pilots). However, in a study with certified pilots who received realistic voice ATC navigational and operational instructions while flying a flight simulator, Mauro and Barshi (1999) found results consistent with those found with college students in the present task.

In one experiment (Schneider et al., 2004), we compared two groups of students, an auditory group receiving auditory messages (*hear*) and a visual group receiving visual messages shown on the

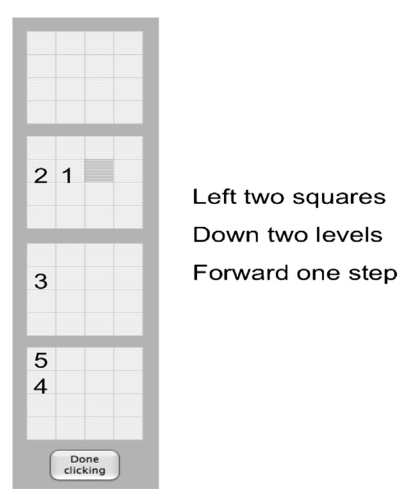

Fig. 1. Sample display showing a message with three commands. The numbers in the matrices show the required clicks; they were not shown to the subjects. The starting point is indicated by the filled-in square.