REPORT DOCUMENTATION PAGE

Form Approved OMB NO. 0704-0188

1. REPORT DATE (DD-MM-YYYY)	2. REPORT TYPE	3. DATES COVERED (From - To)
	New Reprint	-

4. TITLE AND SUBTITLE
Which modality is best for presenting navigation instructions?

5a. CONTRACT NUMBER
W911NF-05-1-0153

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
611103

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

6. AUTHORS
Alice F. Healy, Vivian I. Schneider, Blu McCormick, Deanna M. Fierman, Carolyn J. Buck-Gengler, Immanuel Barshi

7. PERFORMING ORGANIZATION NAMES AND ADDRESSES
University of Colorado - Boulder
3100 Marine Street, Room 479
572 UCB
Boulder, CO 80303 -1058

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS (ES)
U.S. Army Research Office
P.O. Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSOR/MONITOR'S ACRONYM(S)
ARO

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
48097-MA-MUR.128

12. DISTRIBUTION AVAILIBILITY STATEMENT
Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES
The views, opinions and/or findings contained in this report are those of the author(s) and should not contrued as an official Department of the Army position, policy or decision, unless so designated by other documentation.

14. ABSTRACT
tThree experiments involved college students receiving and following instructions of various lengths fornavigating in a three-dimensional space displayed on a computer screen. The purpose was to evaluatewhich is the best modality for presenting navigation instructions so that they can be executed suc-cessfully. Single modalities (read, hear, and see) were considered along with dual modalities presentedsimultaneously 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 modalitiesdid not yield better accuracy than information presented twice in a

15. SUBJECT TERMS
modality, navigation

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	15. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Alice Healy
UU	UU	UU	UU		19b. TELEPHONE NUMBER 303-492-5032

Report Title

Which modality is best for presenting navigation instructions?

ABSTRACT

tThree experiments involved college students receiving and following instructions of various lengths fornavigating in a three-dimensional space displayed on a computer screen. The purpose was to evaluatewhich is the best modality for presenting navigation instructions so that they can be executed suc-cessfully. Single modalities (read, hear, and see) were considered along with dual modalities presentedsimultaneously 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 modalitiesdid not yield better accuracy than information presented twice in a single modality. Also, the ordering ofmodalities depended on the extent of practice. Thus, presentation modality does not have a consistentlylarge effect on receiving and following navigation instructions. Repetition and the amount of practiceare much more important variables than is presentation modality in determining how well navigationinstructions are followed.

REPORT DOCUMENTATION PAGE (SF298)
(Continuation Sheet)

Continuation for Block 13

ARO Report Number 48097.128-MA-MUR
Which modality is best for presenting navigation...

Block 13: Supplementary Note

Approved for public release; distribution is unlimited.

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 © 2013 Society for Applied Research in Memory and Cognition. Published by Elsevier Inc. All rights reserved.
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

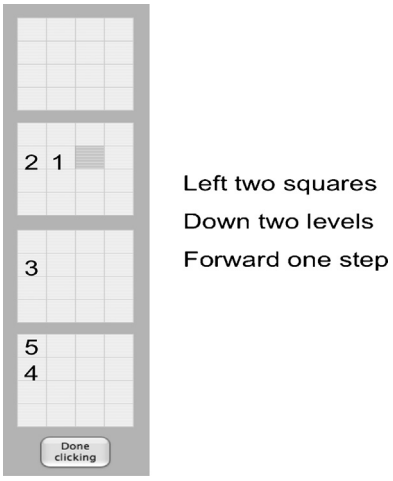

Fig. 1. Sample display showing a message with three commands. The numbers in the matrices show the required clicks; they were not shown to the subjects. The starting point is indicated by the filled-in square.

side of the grid (*read*). Students received 72 different messages varying from one to six commands. We found a disadvantage for reading relative to hearing commands. However, the disadvantage for reading occurred only for longer messages, perhaps because all visual commands were shown simultaneously and subjects did not budget reading time properly.

In another experiment (Schneider et al., 2011), we compared two groups of students: a *hear* group (verbal), in which students listened to messages, and a *see* group (spatial), in which the commands were shown on the computer as simulated movements. We found an advantage for *see* relative to *hear*. However, there was more improvement with practice for *hear* than for *see*, so the advantage for *see* was found only in the first half of the test, with no difference in the last half.

The present study expands on these earlier investigations. It compares *hear*, *read*, and *see* conditions on their own and in combination, with two different types of *see* conditions, one in which subjects see response locations and the other in which they see symbols representing the movements. Specifically, Experiment 1 compares *hear* and *read* conditions to a condition including *both* hearing and reading simultaneously. However, the *both* condition allows subjects to read the messages at a rate different from that at which they are heard, which could cause an asynchrony in timing, thereby making it difficult to find the expected advantage for message redundancy. To eliminate this problem and thereby enhance the ability to find an advantage for combining modalities, Experiment 2 involves sequential, rather than simultaneous presentation of two modalities. In addition, Experiment 2 adds a *see* condition, in which response locations are shown successively on the grid. Each message is repeated either in the same modality (*read-read*, *hear-hear*, *see-see*) or in two different modalities (*read-hear*, *hear-read*, *read-see*, *see-read*, *hear-see*, *see-hear*). The *see* condition of Experiment 2 differs from the *read* condition in the location where the information is presented (within or to the right of the grid). To eliminate differences between *see* and *read* conditions in terms of stimulus location and thereby allow for a pure comparison of nonverbal and verbal stimuli and, thus, a more controlled test of the picture superiority effect, Experiment 3 introduces a *see* condition involving arrow symbols shown at the side of the grid and contrasts it to a *read* condition, in which words instead are shown at the side of the grid. Single conditions, in which a message is shown only once (*see*, *read*), double conditions, in which a message is repeated in the same modality (*see-see*, *read-read*), and mixed conditions, in which a message is repeated in both modalities (*see-read*, *read-see*), are all examined.

1. Experiment 1

In Experiment 1, three different presentation conditions were compared with all subjects exposed to all conditions: *hear*, *read*, and *both* hear and read simultaneously. The conditions were administered in three different orders such that across subjects each condition occurred in each test position.

1.1. Method

1.1.1. Subjects

Thirty-six native English speaking college-age adults participated for payment ($10/h) (27 subjects) or for introductory psychology course credit (9 subjects).

1.1.2. Design and materials

The study implemented a 3 (message modality) × 3 (message modality order) mixed factorial design. Message modality was a within-subject variable; message modality order was a between-subjects variable, with 12 subjects in each order. The dependent measure was proportion of correct manual movement responses scored on an all-or-none basis ignoring the correctness of the oral repetition responses.

Subjects saw a computer screen with four 4 × 4 matrices stacked vertically. Subjects either heard messages telling them where to move on the grid (*hear*), read messages (*read*), or both heard and read messages simultaneously (*both*). These messages were developed by Barshi and Healy (2002) and were spoken by a male native English speaker. The *read* messages were equated (as closely as possible given software constraints) to the *hear* messages in terms of total presentation time. Message length ranged from one to six commands, which instructed subjects to turn left or right a given number of squares, climb up or down a given number of levels, or move forward or back a given number of steps. The duration of each specific spoken command depended on the exact words used, so the duration differed across commands, although it was fixed for a given command. Because the commands were limited to navigation in three dimensions, message lengths longer than three commands had to repeat dimensions. The command dimensions were always presented in the same order (right/left, up/down, forward/back) because commands to pilots are also always given in a consistent order.

Subjects received 24 trials in each of the three modalities, with four of each message length (one to six commands) in each modality. These message lengths were presented in a pseudorandom order, such that each block of 12 trials consisted of two trials of each length.

The conditions were ordered according to a Latin Square, with one-third of the subjects receiving the set of *read* messages (R) first, then *hear* (H), then *both* (B), one-third HBR, and one-third BRH. The order of the messages was the same for all subjects; for example, Message 1 always consisted of the same information regardless of modality.

Each command in the messages contained four words (e.g., "turn left two squares"). All subjects were to repeat back the messages, as pilots do when receiving ATC instructions. In the *read* condition, a message appeared on the screen with all the commands shown simultaneously, with one command presented per line to the right of the complete empty grid of four 4 × 4 matrices. In the *hear* condition, subjects heard the commands as they viewed the empty grid, and in the *both* condition, subjects simultaneously heard the commands and saw them on the right of the grid. The complete grid was visible on the screen as the subjects received the messages, and it remained visible as the subjects responded to the messages in each condition.

1.1.3. Procedure

Subjects were given six preliminary practice trials before the experimental trials in the first modality, three in the second, and three in the third.

Subjects were given a message in one of the modality conditions specifying where to move on the grid; they were to repeat this message and then click a button marked "Done speaking." At this point the starting square on the grid became highlighted (i.e., it was filled rather than empty); the starting square for the preliminary practice trials is shown in Fig. 1; the starting square for the experimental trials was the mirror image on each dimension. Next, subjects followed the directions by clicking on the squares. Each square in the grid became highlighted once it was clicked and stayed highlighted until the next square was clicked and highlighted. Thus, only one square was highlighted at a time. When finished, subjects clicked "Done clicking," and the next trial began after a 1-s delay.

1.2. Results

The proportion of correct responses was equivalent in the *read* ($M = .477$, $SEM = .029$), *hear* ($M = .491$, $SEM = .028$), and *both* ($M = .488$, $SEM = .029$) conditions. The main effect of modality was not significant, $F(2, 66) < 1$; however, modality did interact significantly with modality order, $F(4, 66) = 6.125$, $p < .001$, $\eta^2 = .271$, because accuracy increased from the first ($M = .444$, $SEM = .029$) to the second ($M = .478$, $SEM = .029$) to the third ($M = .534$, $SEM = .027$) modality position. Thus, performance improved with practice regardless of modality.

1.3. Discussion

The combination of modalities (*both* condition) did not yield higher performance than each modality presented separately (*read* or *hear* condition), perhaps because of the asynchrony between the timing of reading a given command and the timing of hearing the same command for the combined modalities.

Neither hearing alone nor reading and hearing together elevated performance over reading alone. Instead, performance improved with practice; performance was worst for the first condition tested and best for the last condition tested.

2. Experiment 2

In Experiment 1, subjects in the *read* condition might have read the stimuli at a rate different from that in which the stimuli were presented in the *hear* condition, despite the equivalent presentation times. Thus, an advantage for the dual-mode condition might be evident only if the two modes are presented separately in time. Hence, in Experiment 2 sequential (rather than simultaneous) dual-mode conditions were compared to single-mode conditions. To provide an appropriate baseline in terms of message repetition, the single-mode conditions involved repetition of the same stimuli in the same modality. Also, a *see* condition was added to the *hear* and *read* conditions. The comparison of conditions was made between subjects, rather than within subjects as in Experiment 1. The between-subjects manipulation also allowed for equal amounts of practice in each modality, so that the amount of practice could not mask any modality effects.

2.1. Method

2.1.1. Subjects

Seventy-two native English speaking college students participated for introductory psychology course credit.

2.1.2. Design and materials

The study implemented a 6 (block of trials) × 9 (message modality order) mixed factorial design. Block of trials was a within-subject variable; message modality order was a between-subjects variable, with 6 or 12 subjects in each modality order. The dependent measure was again proportion of correct responses.

Subjects saw the same computer screen as in Experiment 1. They either heard messages that told them where to move on the grid (*hear*), read messages (*read*), or saw the required movements on the grid (*see*). In the *see* condition, as in Schneider et al. (2011), each square was highlighted on the grid in the same order that subjects were to click when following the *read* or *hear* messages. The time to see the movements in the *see* condition was equated to the time to hear the messages in the *hear* condition. The messages in the *read* condition were shown one command at a time (rather than showing all commands simultaneously as in Experiment 1), with the duration of each command also equated to that in the *hear* condition. Each command was presented on a single line to the right of the grid. A given command was presented and then removed before the next command appeared, and successive commands appeared on different lines, with each command appearing on a line below the previous command.

The set of commands used were those employed by Schneider et al. (2011), who compared *see* and *hear* conditions. All *read* and *hear* commands included only three words (e.g., left two squares). Instead of presenting commands in a fixed order, this set of commands was constructed so that the number of movements along each movement dimension (right/left, up/down, forward/back) was equated at each serial position of the commands.

Subjects received 72 trials, divided into six blocks of 12, including two of each of the six message lengths (one to six commands) in a pseudorandom order.

There were three single-mode conditions (*read-read*, *hear-hear*, and *see-see*), in which a given message was presented twice in succession in the same modality, with 12 subjects in each condition. There were six dual-mode conditions (*read-hear*, *read-see*, *hear-read*, *hear-see*, *see-read*, and *see-hear*), in which a given message was presented in one modality and then repeated in another modality, with 6 subjects in each condition. The messages were repeated in the single-mode conditions so that performance could be appropriately compared to that in the dual-mode conditions, where the messages were necessarily repeated. The order of the 72 trials was the same for each of the nine conditions.

2.1.3. Procedure

The same procedure was used as in Experiment 1 except that subjects did not repeat back the messages or see any button marked "Done speaking" because oral repetition responses are not appropriate for the *see* condition. At the start of the session, subjects were given six preliminary practice trials for their specific condition. Also, the starting square was denoted with a red asterisk and shown at the start of each trial.

2.2. Results

Two analyses were conducted; the first, restricted to the single-mode conditions, allows for the assessment of the relative merits of each modality on its own. The second, an overall analysis including all nine conditions, allows for a comparison of single-mode and dual-mode conditions.

2.2.1. Single-mode conditions only

The proportion correct was highest in *see-see* ($M = .804$, $SEM = .015$), lowest in *read-read* ($M = .684$, $SEM = .019$), and intermediate in *hear-hear* ($M = .756$, $SEM = .017$). The main effect of modality was significant in the analysis restricted to the single-mode conditions, $F(2, 33) = 4.710$, $p = .016$, $\eta^2 = .222$. Fisher's PLSD post hoc tests indicated that only the difference between *see-see* and *read-read* was significant. There was also a significant interaction of modality and block (see Fig. 2), $F(10, 165) = 1.947$, $p = .042$, $\eta^2 = .106$, because performance on *see-see* was at least somewhat better than that on *hear-hear* in all but the last block, in agreement with Schneider et al. (2011), and *read-read* was worse than the other two conditions on all but the first block, where it was somewhat better than *hear-hear*.

2.2.2. All conditions

When all conditions were considered, *see-see*, *see-read*, and *hear-read* yielded the best performance, and *read-read* the worst (see Fig. 3 for the means and standard errors in each condition). The analysis including all conditions yielded a significant main effect of condition, $F(8, 63) = 2.119$, $p = .047$, $\eta^2 = .212$. Fisher's PLSD post hoc tests indicated that along with the significant difference between *see-see* and *read-read*, there were significant differences between *see-read* and *read-read* and between *hear-read* and *read-read*. Thus,

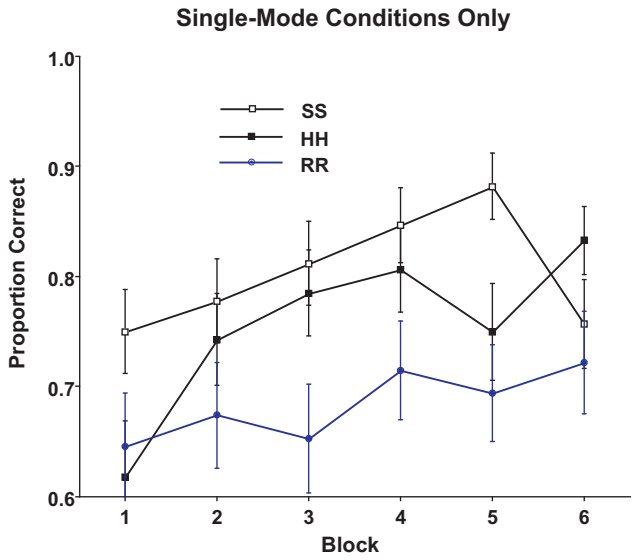

Fig. 2. Proportion of correct responses in Experiment 2 as a function of modality and block for the single-mode conditions only. *Note*. RR = read-read, HH = hear-hear, SS = see-see. Error bars represent standard errors of the mean.

reading alone was not conducive to accurate performance, but reading preceded by either hearing or seeing was.

2.3. Discussion

Reading after either of the other two modes (i.e., *hear-read* or *see-read*) and *see-see* were best, whereas *read-read* was worst. Finding that *see-see* was one of the best conditions implies that words are more difficult to encode than moves. Also, finding that *read-read* was the worst condition implies that reading is worse than hearing or seeing. Importantly, reading was beneficial only when preceded by the hearing or seeing modes.

Why is the *read* mode particularly helpful when it follows another mode? The *read* mode requires looking at the words on the side of the grid rather than where the moves are made in the grid. Therefore, the *read* mode might interfere with following the moves in the grid if subjects look at or point to each successive location as they receive the commands (Brooks, 1967). Thus, one possibility is that subjects follow the moves in the grid on the first presentation but do not follow the moves in the grid on the second, so the *read* mode does not interfere after presentation in a different mode and actually might serve as a valuable confirmation of what was already seen or heard. This explanation is consistent with the proposal by Lu et al. (in press) that redundancy helps guarantee security.

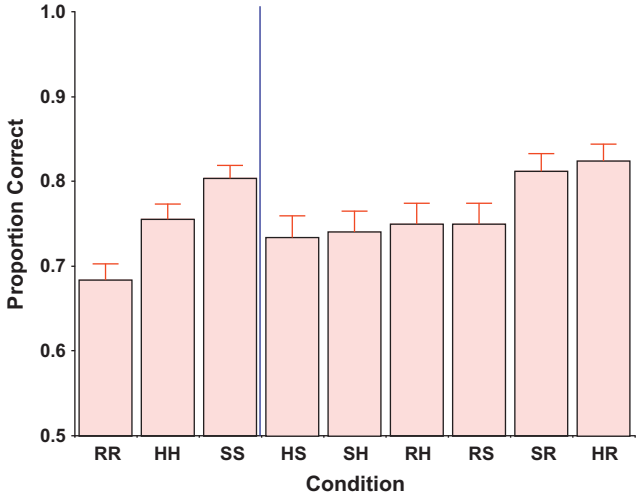

Fig. 3. Proportion of correct responses in Experiment 2 as a function of condition. *Note*. RR = read-read, HH = hear-hear, SS = see-see, HS = hear-see, SH = see-hear, RS = read-see, SR = see-read, RH = read-hear, HR = hear-read. The single-mode conditions are separated from the dual-mode conditions by a vertical line, and both sets of conditions are ordered from lowest to highest accuracy. Error bars represent standard errors of the mean.

Contrary to the dual-coding theory of memory (Paivio, 1971) and theories of multimedia learning (e.g., Mayer, 2001), these results showed that two modalities are not always better than one for processing information (see Lee & Bowers, 1997, for a similar finding in multimedia learning).

3. Experiment 3

The *see* condition used in Experiment 2 (and in Schneider et al., 2011) displayed every move on the computer screen as subjects would make the moves themselves. This procedure might have given an unfair advantage to the *see* condition and also made the *read* and *see* conditions differ in other ways than those involving verbal or nonverbal presentation of the messages. Specifically, the *read* and *see* conditions differed in terms of the location of the stimuli (to the right of the grid or within the grid, respectively). Stimuli on the side of the grid might require scanning back and forth from the stimuli to the grid if subjects follow the movements in the grid as they receive the navigation commands. To investigate differences between verbal and nonverbal messages without any irrelevant, confounding differences between conditions, in Experiment 3 we employed a *see* condition that used symbols (arrows) to convey the direction and magnitude of the required movements, with the arrows occurring to the right of the grid just like the words in the *read* condition. According to earlier research examining spatial compatibility effects (e.g., Miles & Proctor, 2012), arrow symbols should be more compatible than words with spatial locations, and thus, the *see* condition might be expected to lead to better execution performance than the *read* condition in the present spatial navigation task. In this case, we examined *see* and *read* conditions with both a single presentation (*see* and *read*) and two repeated presentations (*see-see* and *read-read*) as well as combined conditions (*see-read* and *read-see*). No *hear* conditions were included.

3.1. Method

3.1.1. Subjects, design, materials, and procedure

The subjects, design, materials, and procedure for Experiment 3 were equivalent to those in Experiment 2, except as noted here.

The study employed a 6 (block of trials) × 3 (presentation type) mixed factorial design. Again, the dependent measure was proportion of correct responses. Block of trials was a within-subject variable; presentation type was a between-subjects variable. Presentation type included single (*read* or *see*) versus double (*read-read* or *see-see*) versus mixed (*read-see* or *see-read*) presentation of messages, with 12 subjects in each of the six specific presentation types.

The *read* commands were presented in the same way as in the *read* condition of Experiment 2 (see Fig. 1). The *see* commands were presented as colored arrows and boxes (see Fig. 4). After the initial command appeared, each subsequent command was timed to appear on the screen after 1.25 s.

3.2. Results

Two analyses were conducted, an overall analysis including all three presentation types (single, double, or mixed) and a second

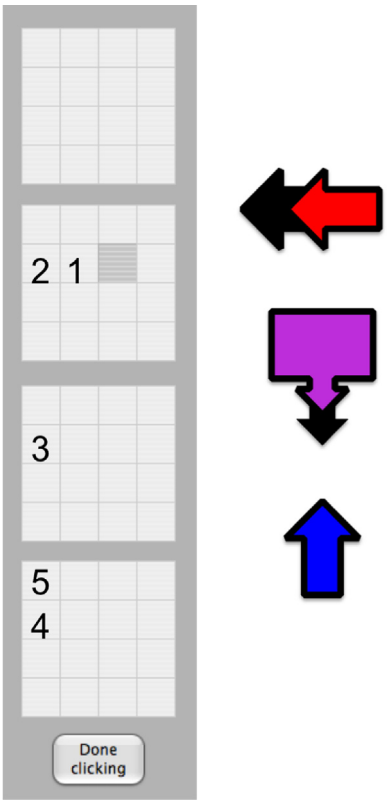

Fig. 4. Sample display showing a message with three commands in the *see* modality of Experiment 3. The numbers in the matrices show the required clicks; they were not shown to the subjects. The starting point is indicated by the filled-in square.

analysis excluding the mixed condition but including the variable of modality (*read* or *see*). Two separate analyses were needed because the variables of modality and presentation type were not fully crossed. The analysis excluding the mixed group allowed modality (*read* versus *see*) to be crossed with presentation type (single versus double presentation).

3.2.1. Overall ANOVA

Performance was better for double presentation ($M = .715$, $SEM = .013$) and mixed presentation ($M = .644$, $SEM = .014$), when messages were viewed twice, than for single presentation ($M = .542$, $SEM = .014$). The main effect of presentation type was significant, $F(2, 69) = 11.317$, $p < .001$, $\eta^2 = .247$. Fisher's PLSD post hoc tests revealed that although double was higher than mixed presentation, the difference was only marginally significant ($p = .056$), but the advantage for twice-viewed messages over single presentation held when the messages were viewed in one modality ($p < .001$) or in mixed modalities ($p = .007$).

3.2.2. ANOVA with modality variable

Performance for doubly presented messages ($M = .715$, $SEM = .013$) was superior to that for singly presented messages ($M = .542$, $SEM = .014$), $F(1, 44) = 25.235$, $p < .001$, $\eta^2 = .364$. Performance on the *read* ($M = .634$, $SEM = .014$) and *see* ($M = .623$, $SEM = .014$) modalities was virtually equivalent, $F(1, 44) < 1$, and did not depend on presentation type, $F(1, 44) = 2.373$, $p = .131$, $\eta^2 = .051$. However, subjects improved more consistently across blocks for *read* than for *see*, so that by the last two blocks there was actually a numerical advantage for *read* relative to *see* (see Fig. 5); the interaction between modality (*read* and *see*) and block was significant, $F(5, 220) = 2.545$, $p = .029$, $\eta^2 = .055$.

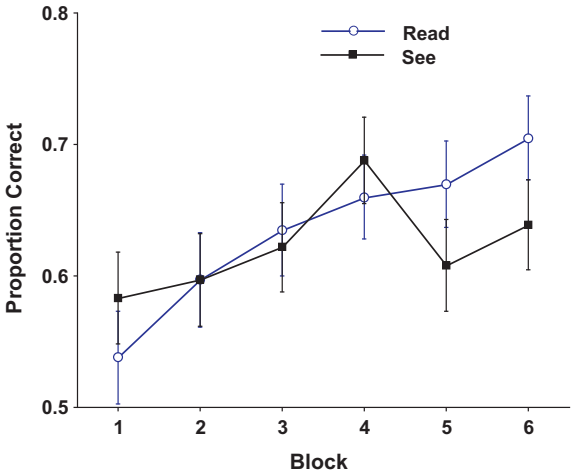

Fig. 5. Proportion of correct responses in Experiment 3 as a function of modality and block. Note that the *see* condition in Experiment 3 involved arrow symbols. Error bars represent standard errors of the mean.

3.3. Discussion

Repeating messages in either modality (*read* or *see*) aided performance relative to presenting messages only once. However, repeating messages in two different modalities was no better than repeating them in a single modality, which is consistent with the proposal by Lu et al. (in press) that redundancy is valuable because it helps guarantee security if repetition in the same modality is at least as effective as repetition in two different modalities for guaranteeing security. In addition, performance did not differ for the *read* and *see* conditions despite the better spatial compatibility for the arrow symbols used in the *see* condition than for the words used in the *read* condition (Miles & Proctor, 2012). Thus, the *see* condition in Experiment 3 does not seem to be as effective as the *see* condition in Experiment 2, which displayed every move on the computer screen as subjects would make the moves themselves and which did yield significantly better performance than the *read* condition in that experiment. As speculated earlier, the procedure used for the *see* condition in Experiment 2 might have given it an unfair advantage because subjects in that condition did not have to scan between the commands on the side of the grid and the grid itself.

Also, in agreement with the earlier findings of Schneider et al. (2011), we found relatively little improvement with practice for the *see* modality, although performance did improve for the *read* modality. Thus, there was an interaction, with performance on the *see* modality somewhat better than on the *read* modality at the start of practice, but the opposite pattern at the end of practice; the modality ordering depended on the amount of practice. We speculate that this difference between the *see* and *read* conditions in the improvement due to practice could be a function of the difference between those conditions in spatial compatibility. Because of the greater spatial compatibility of the symbols used in the *see* condition than of the words used in the *read* condition, the *see* condition might require less skill development for the present task involving navigation to locations in space than would the *read* condition.

4. General discussion

The present study assessed which modality maximizes execution performance on following navigation instructions. The results generally show that the modalities are ordered from least to most effective as *read*, *hear*, *see*, but this ordering was not found in all cases. In Experiment 2, when there were significant differences

observed between *see*, *hear*, and *read*, they were ordered such that *see* was best and *read* was worst. However, as made clear in Experiment 1, practice has an effect that overwhelms modality of presentation, and in Experiment 2 the effects of modality depended on practice block. Likewise, in Experiment 3 the effect of practice was greater for the *read* modality than for the *see* modality, so the ordering of the *read* and *see* modalities depended on the amount of practice.

As in previous studies (e.g., Crowder & Morton, 1969; Gardiner et al., 1983; Goolkasian & Foos, 2002; Murdock, 1967), we tended to find an advantage for *hear* relative to *read*. We can explain this difference using the same theoretical proposals (e.g., Nairne, 1988; Penney, 1989) advanced for explaining the similar findings in the standard memory tasks of an advantage for auditory relative to visual presentation. However, we also showed that visual presentation tends to be better than auditory presentation when the visual presentation involves *see* rather than *read*. The previous findings all involved memory for word lists, whereas the present findings involve memory for navigation instructions, where there was expected to be a conflict between reading the verbal messages and imagining the spatial relations that the messages describe (Brooks, 1967). Such a conflict is diminished when the visual messages are seen as movements or symbols rather than read as words, and messages seen as movements or symbols might benefit from the well-established picture superiority effect (Nelson et al., 1976; Paivio & Csapo, 1973; Shepard, 1967; Snodgrass & McClure, 1975) and from stimulus/central-processing/response compatibility given the spatial nature of the required movements (Wickens & Hollands, 2000; Wickens et al., 1984). The present findings are in agreement with the earlier studies using the present paradigm, which also showed an advantage for *hear* relative to *read* (Schneider et al., 2004) and for *see* relative to *hear* (Schneider et al., 2011).

Consistent with prior theory (e.g., Mayer, 2001; Paivio, 1971; Wickens, 2008), two different modalities should have an advantage over a single modality at least when the two modalities are presented simultaneously (the theory of multiple resources applies only to concurrent, not sequential, processing). However, no advantage for redundant simultaneous dual modalities was found in Experiment 1 (in accordance with the findings of Helleberg & Wickens, 2003; Lancaster & Casali, 2008; Wickens et al., 2003), and, when controlling for the number of repetitions in Experiments 2 and 3, no advantage for dual modalities was evident when the modalities were presented sequentially. These findings are consistent with the proposal by Lu et al. (in press) that redundancy aids performance by guaranteeing security if it is also the case that repetition in a single modality is at least as good as repetition in two different modalities in terms of guaranteeing security.

4.1. Practical application

The present results make it clear that the ordering of modalities for optimal presentation of navigation instructions depends on factors such as practice. In fact, the present study demonstrates that repetition and the amount of practice are much more important variables than is presentation modality for maximizing performance at receiving and following navigation instructions like those received by pilots from ATC. As mentioned earlier, however, experienced pilots presumably have much more extensive practice than was provided in the present experiments, and there are high visual demands on pilots, especially during takeoff and landing. Thus, future research on this topic should include longer practice as well as other aspects of the task facing pilots but missing from the present paradigm. Following navigation instructions, though, is an everyday task not limited to pilots (e.g., as mentioned in the Introduction, finding merchandise in a multi-story department store), and individuals might be much less practiced and might not have high visual demands in those tasks. Thus, the lessons learned in this study might be applicable to many real-world situations in which individuals receive and execute instructions about where to make movements in an environment. Specifically, to maximize performance, individuals should be given sufficient practice at the navigation task, and the navigation messages should be presented twice in a row. However, the modality in which the messages are presented (e.g., in written or spoken form) is not crucial, so that the most convenient modality can be employed.

Acknowledgements

This research was supported in part by National Aeronautics and Space Administration Grants NNA05CS42A, NNA07CN59A, and NNX10AC87A, Army Research Institute Contract DASW01-03-K-0002, and Army Research Office Grant W911NF-05-1-0153 to the University of Colorado. A preliminary version of Experiment 1 was presented at the Symposium on Memory Dynamics and the Optimization of Instruction at the 2007 annual convention of the American Psychological Association, a preliminary version of Experiment 2 was presented at the 2008 annual meeting of the Psychonomic Society, and all three experiments were reviewed at the 2013 biennial meeting of the Society for Applied Research in Memory and Cognition. We would like to thank Lyle Bourne for thoughtful suggestions and advice concerning this research.

References

Barshi, I., & Healy, A. F. (2002). The effects of mental representation on performance in a navigation task. *Memory and Cognition*, *30*, 1189–1203. http://dx.doi.org/10.3758/BF03213402

Brooks, L. R. (1967). The suppression of visualization by reading. *Quarterly Journal of Experimental Psychology*, *19*, 289–299. http://dx.doi.org/10.1080/14640746708400105

Crowder, R. G., & Morton, J. (1969). Precategorical acoustic storage (PAS). *Perception and Psychophysics*, *5*, 365–373. http://dx.doi.org/10.3758/BF03210660

Gardiner, J. M., Gardiner, M. M., & Gregg, V. H. (1983). The auditory recency advantage in longer term free recall is not enhanced by recalling prerecency items first. *Memory and Cognition*, *11*, 616–620. http://dx.doi.org/10.3758/BF03198286

Goolkasian, P., & Foos, P. W. (2002). Presentation format and its effect on working memory. *Memory and Cognition*, *30*, 1096–1105. http://dx.doi.org/10.3758/BF03194327

Helleberg, J. R., & Wickens, C. D. (2003). Effects of data-link modality and display redundancy on pilot performance: An attentional perspective. *International Journal of Aviation Psychology*, *13*, 189–210. http://dx.doi.org/10.1207/S15327108IJAP1303_01

Kerns, K. (1991). Data-link communication between controllers and pilots: A review and synthesis of the simulation literature. *International Journal of Aviation Psychology*, *1*, 181–204. http://dx.doi.org/10.1207/s15327108ijap0103_1

Lancaster, J. A., & Casali, J. G. (2008). Investigating pilot performance using mixed-modality simulated data link. *Human Factors*, *50*, 183–193. http://dx.doi.org/10.1518/001872008X250737

Lee, A. Y., & Bowers, A. N. (1997). *The effect of multimedia components on learning* Proceedings of the Human Factors and Ergonomics Society 41st annual meeting, (pp. 340–344).

Loftus, G. R., Dark, V. J., & Williams, D. (1979). Short-term memory factors in ground controller/pilot communication. *Human Factors*, *21*, 169–181. http://dx.doi.org/10.1177/001872087902100204

Lu, S. A., Wickens, C. D., Prinet, J. C., Hutchins, S. D., Sarter, N., & Sebok, A. (in press).Supporting interruption management and multimodal interface design: Threemeta-analyses of task performance as a function of interrupting task modality.Human Factors, doi:10.1177/0018720813476298.

Mauro, R., & Barshi, I. (1999). Effect of emotion on memory for communication in simulated flight and analog tasks. In R. S. Jensen (Ed.), *Proceedings of the Tenth International Symposium of Aviation Psychology* (pp. 1331–1336). Columbus, OH: The Ohio State University.

Mayer, R. E. (2001). *Multimedia learning*. Cambridge: Cambridge University Press.

Mayer, R. E., & Gallini, J. K. (1990). When is an illustration worth ten thousand words? *Journal of Educational Psychology*, *82*, 715–726. http://dx.doi.org/10.1037/0022-0663.82.4.715

Mayer, R. E., & Sims, V. K. (1994). For whom is a picture worth a thousand words? Extensions of a dual-coding theory of multimedia learning. *Journal of Educational Psychology*, *1994*, 389–401. http://dx.doi.org/10.1037/0022-0663.86.3.389

McGann, A., Morrow, D., Rodvold, M., & Mackintosh, M.-A. (1998). Mixed-media communication on the flight deck: A comparison of voice, data link, and mixed ATC environments. *The International Journal of Aviation Psychology*, *8*, 137–156. http://dx.doi.org/10.1207/s15327108ijap0802_4

Miles, J. D., & Proctor, R. W. (2012). Correlations between spatial compatibility effects: Are arrows more like locations or words? *Psychological Research*, *76*, 777–791. http://dx.doi.org/10.1007/s00426-011-0378-8

Mousavi, S. Y., Low, R., & Sweller, J. (1995). Reducing cognitive load by mixing auditory and visual presentation modes. *Journal of Educational Psychology*, *87*, 319–334. http://dx.doi.org/10.1037/0022-0663.87.2.319

Murdock, B. B., Jr. (1967). Auditory and visual stores in short-term memory. *Acta Psychologica*, *27*, 316–324. http://dx.doi.org/10.1016/0001-6918(67)90074-1

Nairne, J. S. (1988). A framework for interpreting recency effects in immediate serial recall. *Memory and Cognition*, *16*, 343–352. http://dx.doi.org/10.3758/BF03197045

Nelson, D. L., Reed, V. S., & Walling, J. R. (1976). Pictorial superiority effect. *Journal of Experimental Psychology: Human Learning and Memory*, *2*, 523–528. http://dx.doi.org/10.1037/0278-7393.2.5.523

Paivio, A. (1971). *Imagery and verbal processes.* NY: Holt, Rinehart & Winston.

Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology*, *45*, 255–287. http://dx.doi.org/10.1037/h0084295

Paivio, A., & Csapo, K. (1973). Picture superiority in free recall: Imagery or dual coding? *Cognitive Psychology*, *5*, 176–206. http://dx.doi.org/10.1016/0010-0285(73)90032-7

Penney, C. G. (1989). Modality effects and the structure of short-term verbal memory. *Memory and Cognition*, *17*, 398–422.

Schneider, V. I., Healy, A. F., & Barshi, I. (2004). Effects of instruction modality and readback on accuracy in following navigation commands. *Journal of Experimental Psychology: Applied*, *10*, 245–257. http://dx.doi.org/10.1037/1076-898X.10.4.245

Schneider, V. I., Healy, A. F., Barshi, I., & Kole, J. A. (2011). Following navigation instructions presented verbally or spatially: Effects on training, retention, and transfer. *Applied Cognitive Psychology*, *25*, 53–67. http://dx.doi.org/10.1002/acp.1642

Shepard, R. N. (1967). Recognition memory for words, sentences, and pictures. *Journal of Verbal Learning and Verbal Behavior*, *6*, 156–163. http://dx.doi.org/10.1016/S0022-5371(67)80067-7

Snodgrass, J. G., & McClure, P. (1975). Storage and retrieval properties of dual codes for pictures and words in recognition memory. *Journal of Experimental Psychology: Human Learning and Memory*, *1*, 521–529. http://dx.doi.org/10.1037/0278-7393.1.5.521

Tversky, B. (2003). Structures of mental spaces: How people think about space. *Environment and Behavior*, *35*, 66–80. http://dx.doi.org/10.1177/0013916502238865

Wickens, C. D. (2008). Multiple resources and mental workload. *Human Factors*, *50*, 449–455. http://dx.doi.org/10.1518/001872008X288394

Wickens, C. D., Goh, J., Helleberg, J., Horrey, W. J., & Talleur, D. A. (2003). Attentional models of multitask pilot performance using advanced display technology. *Human Factors*, *45*, 360–380. http://dx.doi.org/10.1518/hfes.45.3.360.27250

Wickens, C. D., & Hollands, J. G. (2000). *Engineering psychology and human performance* (3rd ed.). Upper Saddle River, NJ: Prentice Hall.

Wickens, C. D., Sandry, D. L., & Vidulich, M. (1983). Compatibility and resource competition between modalities of input, central processing, and output. *Human Factors*, *25*, 227–248. http://dx.doi.org/10.1177/001872088302500209

Wickens, C. D., Vidulich, M., & Sandry-Garza, D. (1984). Principles of S–C–R compatibility with spatial and verbal tasks: The role of display-control location and voice-interactive display-control interfacing. *Human Factors*, *26*, 533–543. http://dx.doi.org/10.1177/001872088402600505