

Does spatial information impact immediate verbatim recall of verbal navigation instructions?

Vivian I. Schneider¹ · Alice F. Healy¹ · James A. Kole² · Immanuel Barshi³

Published online: 25 September 2017
© Psychonomic Society, Inc. 2017

Abstract The present study addresses the issue of whether spatial information impacts immediate verbatim recall of verbal navigation instructions. Subjects heard messages instructing them to move within a two-dimensional depiction of a three-dimensional space consisting of four stacked grids displayed on a computer screen. They repeated the instructions orally and then followed them manually by clicking with a mouse on the grids. Two groups with identical instructions were compared; they differed only in whether the starting position was indicated before or after the instructions were given and repeated, with no differences in the manual movements to be made. Accuracy on both the oral repetition and manual movement responses was significantly higher when the starting position was indicated before the instructions. The results are consistent with the proposal that there is only a single amodal mental representation, rather than distinct verbal and nonverbal representations, of navigation instructions. The advantage for the *before* condition was found even for the oral repetition responses, implying that the creation of the amodal representation occurred immediately, while the

instructions were being held in working memory. In practical terms, the findings imply that being able to form a mental representation of the movement path while being given verbal navigation instructions should substantially facilitate memory for the instructions and execution of them.

Keywords Navigation and spatial memory · Human memory and learning · Mental models · Amodal mental representation

It has been argued that for navigation instructions there is only a single abstract amodal mental representation that does not include modality-specific features (e.g., Bryant, 1997; Loomis, Klatzky, & Giudice, 2013), rather than the two separate verbal and nonverbal representations included, for example, in Paivio's (1971) dual-coding theory. Nevertheless, it is certainly possible to recall verbal navigation instructions immediately with no reference to a space (i.e., to repeat the words in movement instructions without visualizing the movement path). Presumably the amodal mental representation would take time to develop. The fact that the verbal representation is not always equivalent to the amodal representation is illustrated by a result of Taylor and Tversky (1992). They found that when verifying verbatim or inference locative statements concerning descriptions of naturalistic environments, subjects were faster and more accurate to verbatim than to inference statements. However, both types of statements should be equally fast if the mental representation is amodal. The present study addresses the issue of whether spatial information impacts immediate verbatim recall of verbal navigation instructions. It is possible that the spatial information influences the constructed mental representation eventually but not initially. Thus, verbal instructions might create an abstract amodal mental representation equivalent to that found with perception, but the creation of the amodal representation

Electronic supplementary material The online version of this article (<https://doi.org/10.3758/s13423-017-1379-4>) contains supplementary material, which is available to authorized users.

✉ Alice F. Healy
alice.healy@colorado.edu

¹ Department of Psychology and Neuroscience, University of Colorado Boulder, Muenzinger Building, 345 UCB, Boulder, CO 80309-0345, USA

² School of Psychological Sciences, University of Northern Colorado, Greeley, CO, USA

³ Human Systems Integration Division, NASA, Ames Research Center, Moffett Field, CA, USA

might be effortful and, hence, might not occur immediately (i.e., when the description is held in working memory).

To study this issue, we used a laboratory task in which movements were made in a space consisting of a grid of four stacked 4 × 4 matrices shown on a two-dimensional computer screen but representing a three-dimensional space (Fig. 1). Subjects were given auditory navigation instructions and were required to repeat the instructions before following them, which they did by using a mouse to click on the squares in the grid. In earlier work with this task, Barshi and Healy (2002, 2011) showed that subjects' accuracy in repeating the instructions depended on the space in which the instructions applied, even when the same instructions were given for different spaces. In other words, the subjects' memory for the instructions depended on how they interpreted them. Specifically, two matrix conditions were compared. The instructions in both conditions directed subjects to move left or right (lateral movement) and up or down (vertical movement). However, in one condition (*single-matrix*) the vertical movement involved different rows in a single matrix (thus encouraging only a 2-dimensional representation), whereas in the other condition

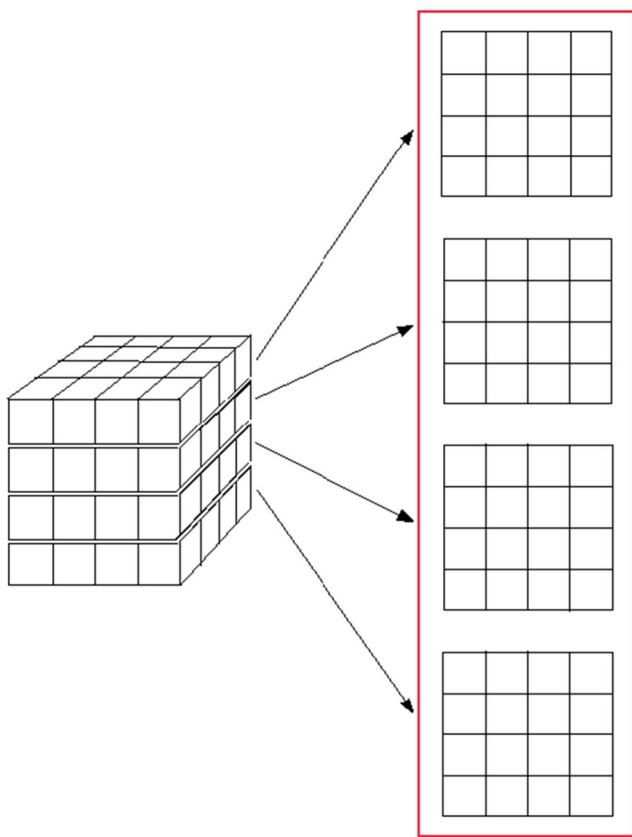


Fig. 1. A grid of four stacked 4X4 matrices shown to subjects on the computer screen (right side) and a cube depicting the three-dimensional space that the grid is meant to represent (left side). Although subjects were not shown the cube on the left side of the figure, they were given a small three-dimensional model representing the space, with dowels connecting four note cards containing the matrices, and this model was left in view throughout the experiment

(*multi-matrix*) it involved the same rows in different matrices (thus encouraging a 3-dimensional representation; Fig. 2). It was found that subjects both repeated and executed the instructions more accurately in the single-matrix condition than in the multi-matrix condition. Thus, there was a significant difference between the two matrix conditions in subjects' ability simply to repeat the instructions even though the words they heard and were to repeat immediately were identical in the two conditions. On the basis of these findings, Barshi and Healy (2002) concluded that the verbal representation of the navigational instructions used for immediate verbatim recall of the instructions depends on the mental representation of the navigational space, which can be thought of as a mental model (Johnson-Laird, 1983) or a situation model (Kintsch, 1988; Zwaan & Radvansky, 1998) of the space. Barshi and Healy (2011) considered and ruled out alternative explanations for these results involving the lack of depth cues, the use of diagrams, and the experimental instructions, all of which might influence the construction of a mental representation of the navigational space. Specifically, there was an effect of matrix condition on the ability to repeat the instructions even when depth cues were added to the display, even when the display was a three-dimensional model instead of a two-dimensional diagram, and even when the experimental instructions did not promote a three-dimensional representation but rather a flat depiction of four checkerboards. Most important, the combined findings of the Barshi and Healy (2011) experiments challenged the claim by Lyon, Gunzelmann, and Gluck (2008) that a subject in this paradigm could “rehearse the sequence of segments verbally, and report them back without ever actually constructing a mental visualization of the path” (p. 125).

In all previous studies with this paradigm, the starting position was fixed. In the present study, the starting position was varied across trials. The subjects were shown the starting position either before or after they received and repeated the instructions. Getting the starting position before the instructions should allow the subjects to visualize a mental path of required moves as the instructions are given. In contrast, with no starting position, such a mental path could not be visualized. Therefore, if the oral repetition of the instructions depends on spatial information, as suggested in the earlier studies, then the subjects' ability to repeat back the instructions should be affected by the timing of the display of the starting position (before or after hearing and repeating the instructions). Specifically, because the mental representation needs to have a starting position to include the path of movements, it is predicted that subjects would be better able to repeat the instructions when the starting position is given before the instructions rather than after them. For the same reasons, manual movement accuracy would be expected to be superior when the starting position is seen before, rather than after, hearing the instructions.

In this experiment, we varied matrix condition as well as starting position condition using a 2 × 2 factorial design.

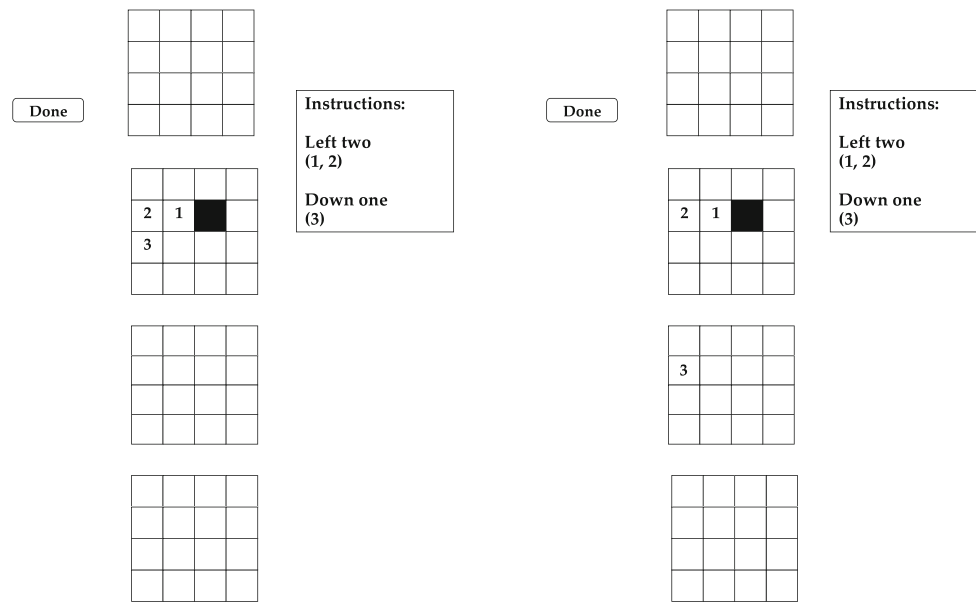


Fig. 2. Sample displays plus instructions for single-matrix condition (left panel) and multi-matrix condition (right panel). The digits indicate the movements to be made; subjects did not see the digits or the written instructions; instead they heard the instructions. The filled-in square is the starting position. Note that movements (i.e., digits) occurred only in one of the four matrices in the single-matrix condition (because in that

condition *up* and *down* referred to different rows of one matrix) but in more than one matrix in the multi-matrix condition (because in that condition *up* and *down* referred to different matrices). Different groups of subjects were exposed to the two conditions, and the instructions were the same for the two conditions; only the required movements were different

Examining both of these variables allows us to assess the relative magnitude of the two manipulations. It is important to note that the movements made in the two starting position conditions (*before* and *after*) are identical. In contrast, as in the study by Barshi and Healy (2002), the movements made in the two matrix conditions (single-matrix and multi-matrix) differ because in the single-matrix condition subjects make only short movements between adjacent squares within a matrix, whereas in the multi-matrix condition subjects make both short movements within a matrix and relatively long movements from one matrix to another. Thus, the comparison between the two starting position conditions is better controlled than that between the two matrix conditions and therefore should provide more conclusive insights into the relationship between spatial information and the mental representation of the instructions. In particular, a difference in performance between the starting position display time conditions (i.e., better performance for the *before* relative to the *after* condition) could reveal whether the mental representation includes the specific navigational path, rather than just the abstract space to which the navigation instructions apply.

Method

Subjects

A total of 24 subjects were tested, 6 in each of 4 subgroups. All of these subjects were undergraduate students at the University

of Colorado taking a class in General Psychology, and all were native speakers of English. They earned course credit for their participation in this experiment. An additional subject was tested (in the *single-matrix before* condition), but the data from that subject were not included in the statistical analyses, because the subject got only 5 trials correct out of 36 trials total. According to a power calculation based on the main effect of matrix condition in Barshi and Healy (2002) for the oral repetition responses ($\eta^2 = 0.289$), only 14 subjects would be needed to achieve 0.8 power for that main effect.

Materials

Subjects were given 36 messages, 6 of each of 6 lengths (1 to 6 commands). Each command included two words (e.g., left two) that provided both the direction and magnitude of the movement. Only two movement dimensions were included: right/left and up/down. A fixed order of messages was used for all subjects, with the 36 trials broken down into six blocks of six trials, including one of each message length in a random order. The same messages were given to subjects in all four groups, with only the starting position different for the single-matrix and multi-matrix conditions on some trials (due to the constraints on the possible moves within the space). There also was a set of six practice trials, one of each message length, with the same starting position for the single-matrix and multi-matrix conditions.

Design

The computer screen displayed a grid of four stacked 4×4 matrices (Fig. 1, right side), which included a total of 64 squares. Subjects were instructed that the grid represents a three-dimensional space (Fig. 1, cube on the left side, which they were not shown) and were shown a small three-dimensional model representing that space, and this model was left in view throughout the experiment. The starting position was a single filled-in square, with all other squares empty. As each square was clicked, it was filled in and all of the other squares (including the starting position and all previously clicked squares) became empty. Only the numbers one, two, and three were used in the commands. None of the commands led subjects to “fall off” the grid. There was a consistent structure and a fixed alternating order to the commands: (left, right) (one, two, three), followed by (up, down) (one, two, three). For example, a message with six commands was “left one, up two, right three, down two, left three, up two.”

Two different matrix conditions were compared, which used the same verbal navigation instructions but required different executions of these instructions. Specifically, the conditions varied in how the subjects were to move given the commands to move up or down. In the *multi-matrix* condition, *up* and *down* referred to different matrices of the grid, so subjects moved from one matrix to another, whereas in the *single-matrix* condition, *up* and *down* referred to different rows of just one matrix so subjects moved from one row to another within one matrix (Fig. 2).

Unlike previous studies (e.g., Barshi & Healy, 2002, 2011), the starting position varied across trials and in fact was different on every trial. The space, including all 64 squares of the stacked matrices, was displayed at all times (including both before and after the subjects received and repeated the instructions) for both conditions. In addition, the presence of a filled-in square indicating the starting position was shown until movements were made in the *before* condition. However, the filled-in square was shown only after the subjects received and repeated the instructions in the *after* condition, with only empty squares shown before the subjects received and repeated the instructions. Thus, the only difference between the *before* and *after* conditions was when the starting square was filled in.

A mixed factorial design was employed including two between-subjects variables—starting position display time (before, after) and matrix condition (multi, single)—and one within-subjects variable—message length (1–6). Two dependent variables were examined separately—accuracy on the oral repetition responses and accuracy on the manual movement responses. Scoring was all-or-none (i.e., a correct response required that all commands in a message were repeated/followed correctly). The entire path had to be correct; it was not sufficient for the final location to be correct.

Procedure

Subjects heard messages with one to six commands and two words per command. Each message was followed by a beep. After the beep was heard, the subjects’ first task was to repeat the message aloud and then click with the computer mouse on a button labeled DONE, which was to the left of the grid. The subjects’ oral repetition responses were audio taped. Their next task was to follow the navigation instructions by clicking with the mouse on the appropriate squares in the grid. To move right or left, they were to move horizontally within the same matrix. To move up or down, they were to move vertically within the same matrix in the single-matrix condition, but to move vertically to a different matrix in the multi-matrix condition. For example, in response to the command in the single-matrix condition “down one,” subjects were to click on the cell immediately below the one they were on in the same matrix (numeral 3 in Fig. 2, left panel). In response to the same command, subjects in the multi-matrix condition were to click on the same cell as the one they were on in the matrix below the one they were on (numeral 3 in Fig. 2, right panel). Subjects were required to click every cell they passed. Thus, subjects were to make the same number of clicks as the number in the instructions (i.e., “two” = two clicks, “three” = three clicks). At the end of their manual movement responses in the grid, the subjects were to click again on the DONE button. Separating trials was a 2-s pause.

Results

The results concerning the oral repetition responses are summarized in the top panel of Fig. 3. These data were analyzed with a $2 \times 2 \times 6$ mixed factorial analysis of variance, including the factors of display time, matrix condition, and message length. As mentioned earlier, the dependent variable was all-or-none accuracy, requiring that all movements be correct. As in previous studies, accuracy declined with increases in message length, $F(5, 100) = 111.200$, $MSE = 0.174$, $\eta_p^2 = 0.848$, $p < 0.001$, with the largest drops between Lengths 3 and 4 and Lengths 4 and 5. Most important is the finding that accuracy was better overall when the starting position was given before the instructions (0.692) than when it was given after the subject repeated the instructions (0.572), $F(1, 20) = 5.564$, $MSE = 0.562$, $\eta_p^2 = 0.218$, $p = 0.029$. Contrary to previous findings (e.g., Barshi & Healy, 2002), there was, however, no difference in accuracy between the single- (0.637) and the multi- (0.627) matrix conditions, $F < 1$. Furthermore, none of the interactions involving display time, matrix condition, and message length were significant for the oral repetition responses (see *Supplemental Materials*).

The same pattern of results was found for the all-or-none accuracy of manual movement responses (Fig. 3, bottom panel). Specifically, accuracy declined with increases in message

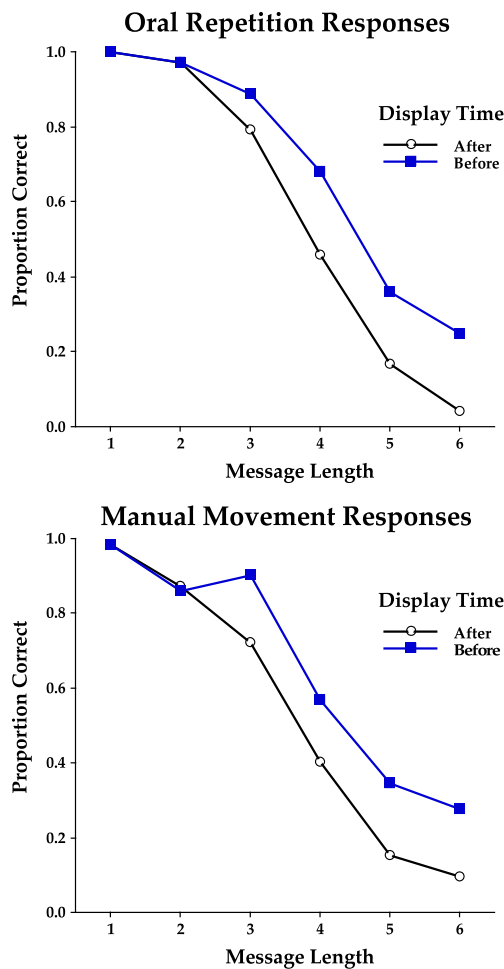


Fig. 3. Proportion of correct oral repetition (top panel) and manual movement (bottom panel) responses as a function of display time and message length

length, $F(5, 100) = 105.689$, $MSE = 0.156$, $\eta_p^2 = 0.841$, $p < 0.001$. Again, most important is the finding that accuracy was better overall when the starting position was given before the instructions (0.657) than when it was given after the subject repeated the instructions (0.539), $F(1, 20) = 4.571$, $MSE = 0.659$, $\eta_p^2 = 0.186$, $p = 0.045$. The difference in accuracy between the two matrix conditions was not significant, $F(1, 20) = 1.914$, $MSE = 0.659$, $\eta_p^2 = 0.087$, $p = 0.182$, although numerically accuracy was somewhat higher with the single-matrix condition (0.637) than with the multi-matrix condition (0.560), as in previous studies. Furthermore, as with the oral repetition responses, none of the interactions involving display time, matrix condition, and message length were significant for the manual movement responses (see *Supplemental Materials*).

Discussion

The present study addresses the issue of whether spatial information influences the mental representation of verbal

navigation instructions and does so immediately or only eventually with effort. To investigate this issue, subjects heard, repeated, and followed the same verbal navigation instructions in two starting position display time conditions (*before* and *after*) and, importantly, the movements were identical in these two conditions, with the only difference between conditions involving when the starting square was filled in. Better performance for the *before* than for the *after* condition would imply that the mental representation includes the specific navigational path, not just the abstract space referred to by the instructions. In fact, subjects' performance was better, even for the oral repetition responses, in the *before* condition than in the *after* condition. These findings are consistent with the proposal that there is only a single abstract amodal mental representation (Bryant, 1997; Loomis et al., 2013), rather than distinct verbal and nonverbal representations (Paivio, 1971), of the navigation instructions, as suggested by the findings of Taylor and Tversky (1992). The fact that the advantage for the *before* condition was found even for the oral repetition responses implies that the creation of the abstract amodal mental representation was not effortful but rather occurred immediately, while the instructions were being held in working memory. In contrast, the present findings are inconsistent with the claim made by Lyon et al. (2008) that subjects could report the sequence of segments verbally without constructing a mental visualization of the path. Whether or not the subjects could form a mental visualization of the movement path at the time when they repeated the instructions (which was possible in the *before* condition but not in the *after* condition) clearly influenced their oral repetition responses, which were always made before the manual movement responses.

In contrast to the effect of display time, the effect of matrix condition was not significant in the present experiment, as it had been in numerous previous experiments (e.g., Barshi & Healy, 2002, 2011). It might seem surprising that the effect of manipulating the location of the starting position and its presentation time could trump the strong matrix effect found previously. Starting position was fixed in the previous experiments and was known in advance of hearing the navigation instructions. In contrast, starting position was varied in the present experiment, so that even in the single-matrix condition subjects did not know in advance which one of the four matrices was relevant to the instructions. This difference in starting positions could explain why the effect of matrix condition is diminished in the present experiment: It seems that more than just mentally representing the abstract space to which the navigation instructions apply, subjects in these studies construct a mental representation of the specific path along which they must navigate. Varying the starting position was likely to have prevented the subjects from representing the path and from forming a stable mental representation of the instructions. The stability of the mental model seems to be necessary to demonstrate the more subtle differences between

single- and multi-matrix conditions, such as their respective involvement with two or three spatial dimensions. The multi-matrix condition but not the single-matrix condition encourages a three-dimensional mental representation of the space. Because we did not find significant effects of matrix condition in the current study, we have not gained any insights concerning the debate about how a three-dimensional space is represented (Jeffery, Jovalekic, Verriotis, & Hayman, 2013), although future research involving the effects of matrix condition in the present paradigm might shed new light on this important issue.

In any event, the effect of starting position display time in the present study provides converging evidence (coupled with that from our previous studies) that performance in both repeating and executing navigation instructions depends on spatial information and not just on the words in the instructions. In addition, the manipulation of starting position, unlike that of matrix condition, has the advantage of holding constant the actual movements that were made. Because the mental representation of the path is practically impossible to construct when the starting position is unknown, the starting position manipulation is even more powerful than the matrix condition manipulation.

Although earlier studies did not vary the timing of the starting position for movement responses, they did vary the presence of other information that could help the participant form a mental representation of the movement space or movement path. For example, Morett, Clegg, Blalock, and Mong (2009) used a driving simulator to study route learning and navigation. Route information was presented either in a verbal narration, consisting of verbal directions concerning how to move in the space, or in a map, or in both formats either simultaneously or sequentially using both orders (narration then map, map then narration). Route knowledge was tested both by tasks requiring explicit recall of the visual-spatial information and by a navigation task. In the navigation task, subjects were required to navigate in the simulator through both the studied routes and novel routes. It was found that presenting a map in addition to the narration facilitated both recall and navigation. Furthermore, for the task of navigating through novel routes, viewing a map before receiving the verbal narration showed a trend of better performance relative to receiving the verbal narration before viewing the map. The authors suggested that the map might “assist in mental consolidation” (Morett et al., 2009, p. 46) of the learned information. Specifically, they explained their findings, in part, in terms of conjoint retention theory (Kulhavy, Lee, & Caterino, 1985), which is focused on the influence of the order of presentation of verbal and visual sequential information on learning. By this theory, there is better performance when visual information precedes verbal information than when verbal information precedes visual information, so that visual stimuli can guide the interpretation and consolidation of verbal information. The conjoint retention theory is also

consistent with the findings of the current experiment if the starting position is viewed as visual information and the navigation instructions as verbal information, so that presenting visual information before verbal information facilitates performance. The starting position presents the crucial piece of visual information allowing subjects to form a representation of the route or movement path in the present paradigm.

Subjects in this paradigm use a computer mouse to move through a small space on a computer screen. Nevertheless, the results should have important practical implications for remembering navigation instructions in a space of any size in the real world. As in previous studies (e.g., Barshi & Healy, 2002), the present results, showing steep declines in performance between message Lengths 3 and 5, imply that when giving navigation instructions no more than three commands should be provided in one message. In addition, the present results extend previous findings, because they imply that being able to form a mental representation of the movement path while being given verbal navigation instructions should substantially facilitate memory for the instructions and execution of them. Specific navigation instructions can be given to individuals before they are even exposed to the space to be traversed. The current findings imply that remembering and following navigation instructions in such cases would be quite difficult but would be aided by receiving enough information to allow for visualizing the path to be taken in the space.

Acknowledgements This research was supported in part by National Aeronautics and Space Administration Grants NCC2-1310, NNA07CN59A, NNX10AC87A, and NNX14AB75A. The authors thank Antonia Hamilton and three anonymous reviewers for their helpful suggestions concerning an earlier version of this manuscript.

References

- Barshi, I., & Healy, A. F. (2002). The effects of mental representation on performance in a navigation task. *Memory & Cognition*, *30*, 1189–1203. doi: <https://doi.org/10.3758/BF03213402>
- Barshi, I., & Healy, A. F. (2011). The effects of spatial representation on memory for verbal navigation instructions. *Memory & Cognition*, *39*, 47–62. doi: <https://doi.org/10.3758/s13421-010-0024-5>
- Bryant, D. J. (1997). Representing space in language and perception. *Mind & Language*, *12*, 239–264. doi: <https://doi.org/10.1111/j.1468-0017.1997.tb00073.x>
- Jeffery, K. J., Jovalekic, A., Verriotis, M., & Hayman, R. (2013). Navigating in a three-dimensional world. *Behavioral and Brain Sciences*, *36*, 523–543. doi: <https://doi.org/10.1017/S0140525X12002476>
- Johnson-Laird, P. N. (1983). *Mental models*. Cambridge, MA: Harvard University Press.
- Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review*, *95*, 163–182. doi: <https://doi.org/10.1037/0033-295X.95.2.163>
- Kulhavy, R. W., Lee, J. B., & Caterino, L. C. (1985). Conjoint retention of maps and related discourse. *Contemporary Educational Psychology*, *10*, 28–37. doi: [https://doi.org/10.1016/0361-476X\(85\)90003-7](https://doi.org/10.1016/0361-476X(85)90003-7)

- Loomis, J. M., Klatzky, R. L., & Giudice, N. A. (2013). Representing 3D space in working memory: Spatial images from vision, hearing, touch, and language. In S. Lacey & R. Lawson (Eds.), *Multisensory imagery* (pp. 131-155). New York: Springer. doi: https://doi.org/10.1007/978-1-4614-5879-1_8
- Lyon, D. R., Gunzelmann, G., & Gluck, K. A. (2008). A computational model of spatial visualization capacity. *Cognitive Psychology*, *57*, 122-152. doi: <https://doi.org/10.1016/j.cogpsych.2007.12.003>
- Morett, L. M., Clegg, B. A., Blalock, L. D., & Mong, H. M. (2009). Applying multimedia learning theory to map learning and driving navigation. *Transportation Research Part F: Traffic Psychology and Behaviour*, *12*, 40-49. doi: <https://doi.org/10.1016/j.trf.2008.07.001>
- Paivio, A. (1971). *Imagery and verbal processes*. New York: Holt, Rinehart & Winston.
- Taylor, H. A., & Tversky, B. (1992). Spatial mental models derived from survey and route descriptions. *Journal of Memory and Language*, *31*, 261-292. doi: [https://doi.org/10.1016/0749-596X\(92\)90014-O](https://doi.org/10.1016/0749-596X(92)90014-O)
- Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin*, *123*, 162-185. doi: <https://doi.org/10.1037/0033-2909.123.2.162>