RESEARCH REPORT

Retrieval Enhances Route Knowledge Acquisition, but Only When Movement Errors are Prevented

Jonathan W. Kelly, Shana K. Carpenter, and Lori A. Sjolund
Iowa State University

Studies of the testing effect have shown that retrieval significantly improves learning. However, most of these studies have been restricted to simple types of declarative verbal knowledge. Five experiments were designed to explore whether testing improves acquisition of route knowledge, which has a procedural component consisting of actions to be performed at decision points (Golledge, 1991). Participants learned a route through a series of connected rooms in a virtual building. Each room contained multiple doors, only one of which led to the next room. During encoding, participants were shown the correct sequence of doors in a manner similar to global positioning system (GPS) navigation guidance. During subsequent exposures to the route, participants were either shown the correct sequence again or had to recall the sequence from memory. Participants later completed a final test in which they traversed the route without guidance or feedback. Testing improved route memory compared to studying, but only when participants were given feedback about the correct door prior to moving through the room. When feedback occurred after moving to an incorrect door, testing resulted in worse performance compared to studying. These findings parallel work on errorless learning, in which procedural skills are acquired more quickly when errors are minimized during learning.

Keywords: testing effect, retrieval practice, spatial memory, route learning, navigation

Supplemental materials: http://dx.doi.org/10.1037/a0038685.supp

Many studies have shown that memory can be enhanced through retrieval. Trying to recall information leads to memorial benefits over and beyond those of restudying it (for reviews see: Dunlosky, Rawson, Marsh, Nathan, & Willingham, 2013; Rawson & Dunlosky, 2011; Roediger & Butler, 2011; Roediger & Karpicke, 2006). This testing effect has been demonstrated for many types of verbal materials, including word lists (Carpenter, 2009, 2011; Kornell, Bjork, & Garcia, 2011), foreign language vocabulary (Coppens, Verkoeijen, & Rikers, 2011; Kang, 2010; Pyc & Rawson, 2010; Vaughn, Rawson, & Pyc, 2013), trivia questions (Kornell, Hays, & Bjork, 2009; McDaniel & Fisher, 1991), and text (Butler, 2010; Chan, McDermott, & Roediger, 2006; Clark & Svinicki, 2014; Hinze, Wiley, & Pellegrino, 2013; Kang, McDermott, & Roediger, 2007; Marsh, Roediger, Bjork, & Bjork, 2007; Roediger & Marsh, 2005).

Recent studies have begun to explore the benefits of retrieval on learning spatial types of information conveyed through maps or other spatial arrays of objects (Carpenter & Kelly, 2012; Carpenter & Pashler, 2007; Rohrer, Taylor, & Sholar, 2010). However, research on the testing effect has yet to explore what is perhaps the most important aspect of spatial learning—successful navigation. Research on navigation commonly distinguishes between two types of spatial knowledge: knowledge of relative positions of locations in the environment (survey knowledge) and knowledge of sequences of movements required to travel from one location to another (route knowledge). Survey knowledge is required to point toward an unseen landmark or take a novel shortcut, but route knowledge suffices when the same route is traveled repeatedly, such as the drive from home to work. Route knowledge is considered procedural, consisting of actions to be performed at decision points (Gillner & Mallot, 1998; Golledge, 1991). Similar to overlearned motor sequences (Jueptner et al., 1997), route knowledge is believed to require less perceptual processing and cognitive control than survey knowledge. Although survey and route knowledge often develop simultaneously (Montello, 1998), their distinction has been confirmed through neuroimaging studies, which associate survey knowledge with the hippocampus and route knowledge with the caudate nucleus (Harley, Maguire, Spiers, & Burgess, 2003).

When traversing a route through an environment, navigators may have the option of trying to retrieve that route from memory, or following some kind of guidance such as a GPS system or a description conveyed through a map or another person. Based on many studies of the testing effect, it might be expected that
recalling the route from memory is more beneficial for long-term retention than relying on route guidance. However, relying on memory is also more likely to involve errors such as wrong turns and backtracking. Previous research using verbal materials has found that retrieval errors do not impair learning, as long as corrective feedback is provided (e.g., Huelser & Metcalfe, 2012; Kornell et al., 2009). However, the procedural nature of route knowledge (Golledge, 1991) may render it more sensitive to errors. Unlike recalling the wrong word from a list, taking a wrong turn involves movement execution. Even if this error is corrected, memory for the erroneous movement may interfere with memory for the correct movement. In studies involving procedural tasks, benefits have been observed for “errorless learning” conditions. For example, performance on a golf putting task is best when training begins with easier putts and becomes progressively more difficult, compared to training that begins with difficult putts and becomes progressively easier (Maxwell, Masters, Kerr, & Weeden, 2001; Poolton, Masters, & Maxwell, 2005). Procedural memory is typically intact in individuals who suffer impairments in explicit memory (e.g., Gabrieli, Corkin, Mickel, & Growdon, 1993), and individuals with such impairments are particularly receptive to the benefits of errorless learning (e.g., Clare & Jones, 2008; Glisky, Schacter, & Tulving, 1986). Thus, it is possible that using retrieval during route navigation—which occasionally involves navigational errors—may be detrimental to learning.

In the current study, participants navigated a route through a virtual environment by relying on retrieval or guidance. The task was designed to capture the procedural elements of route learning (Golledge, 1991) by requiring participants to learn a sequence of movements through a series of virtual rooms containing multiple doors. Only one door in each room led participants into the next room. A visual depiction of this task can be found in the supplemental video. During encoding all participants were shown the correct sequence of doors by highlighting the correct door in each room and directing participants to move to the highlighted door. They then moved through the rooms again, this time choosing the correct doors from memory (Test) or selecting the correct doors that were highlighted for them again (Study). Afterward, all participants completed a final test by moving through the same rooms without guidance or feedback.

Based on numerous studies of the testing effect, it might be expected that Test is more effective than Study for learning the correct door sequence. If these effects parallel the results of verbal learning studies, then significant benefits of testing may emerge regardless of the fact that testing can involve more navigation errors than restudying (e.g., Kornell et al., 2009). However, given the procedural nature of route learning, and the effect of errors on procedural learning (e.g., Maxwell et al., 2001), the benefits of testing may not be expected to occur, particularly under conditions in which participants make erroneous movements during retrieval.

**Experiments 1a and 1b**

Experiments 1a and 1b compared the effects of Test versus Study on the learning of a route through a series of virtual rooms. In Experiment 1a, Study participants completed four encoding trials followed by the final test, and Test participants completed one encoding trial followed by three test trials and then the final test. Due to near-ceiling final test performance of Study participants, the number of trials was modified in Experiment 1b. This time, Study participants completed three encoding trials followed by the final test, and Test participants completed two encoding trials followed by a single test trial and then the final test. The experiments were identical in all other respects.

**Method**

**Participants.** Sixty-two students (30 in Experiment 1a and 32 in Experiment 1b) at Iowa State University participated in exchange for course credit. Participant gender was balanced across condition within each experiment.

**Stimuli.** The virtual environment consisted of 30 adjacent rooms arranged in a row (Figure 1), displayed on a 22 in. monitor updated at 60 Hz. Graphics were rendered using Vizard software (WorldViz, Santa Barbara, CA) running on a computer with Intel Core2 Quad processors and Nvidia GeForce GTX 285 graphics card.

Each room was 10 × 10 × 2.5 m and contained three identical doors on one wall and a unique landmark in the center (Figure 2, first panel). One door in each room led to the next room; the other two doors were locked. The participant used the arrow keys on a keyboard in order to initiate smooth movements through the virtual rooms, similar to how movement is controlled in video games. Use of the arrow keys initiated the visual experience of moving through space, such that participants “passed by” the object in the center of the room in order to reach one of the doors (see Figure 3). Examples of the virtual environment, and the way in which the experience of movement was rendered, are shown in the supplemental video.

**Design and procedure.** Participants were seated at the computer and given verbal instructions before starting. During encoding, the correct door in each room was highlighted with a green border (Figure 2, first panel) and participants were instructed to use the arrow keys to move toward the highlighted door in each room. Upon reaching a location within 50 cm of the correct door, participants entered a new room, and the view of the new room was frozen for 1.5 s to ensure sufficient time for participants to encode the correct door, which was highlighted. Participants again moved toward the correct door in the new room, and this process repeated until all 30 rooms had been traversed.

In Experiment 1a, participants in the Study group completed this encoding task four times. Participants in the Test group completed the encoding task once followed by three test trials over the sequence of rooms. In Experiment 1b, participants in the Study group completed the encoding task three times and participants in the Test group completed the encoding task twice followed by one test trial over the sequence of rooms. In both experiments, Test procedures required participants to select the correct door in each

![Figure 1. Overhead view of the route learned by participants.](image-url)
room using their memory (Figure 2, second panel). Participants accomplished this by moving through the virtual room toward the remembered door. Once participants reached the door toward which they moved, they were informed about whether their choice was correct or incorrect. If correct, they were immediately transported into the next room. If incorrect, red Xs appeared on the two incorrect doors (Figure 2, third panel) and a green border appeared around the correct door. Participants were then required to back up from the incorrect door, and move through the room until they reached the correct door, highlighted in green. Upon reaching this door, participants advanced to the next room. Thus, in Experiments 1a and 1b, participants moved through the rooms to make their door selections during retrieval. We therefore refer to the Test conditions as “Test-Move,” to distinguish this type of trial from those in later experiments in which movement was not initiated until after feedback was provided.

After completion of encoding (Study) or encoding and initial testing (Test-Move), all participants were given a 10 minute dis-

![Figure 3](https://example.com/figure3.png)

Figure 3. Views of one room taken at three separate times during movement toward the correct door. See the online article for the color version of this figure.

![Figure 2](https://example.com/figure2.png)

Figure 2. Perspective view of one room along the route. Each room contained a landmark and three doors. During encoding, the participant was shown the correct door (first panel). During initial and final testing, the participant selected a door from memory (second panel). Feedback was provided during initial testing if the participant selected an incorrect door. In the Test-Move conditions of Experiments 1a, 1b, and 2b, initial test feedback was withheld until the participant moved through the virtual environment and arrived at the selected door (third panel). In the Test-Click conditions of Experiments 2a, 2b, and 3, initial test feedback was provided after the participant clicked on the selected door with a mouse, but before the participant moved through the environment (fourth panel). The landmarks were a bench, tennis racket, shovel, trash can, refrigerator, soccer ball, world map, bird, plant, briefcase, car, lamp, hat, bowling pins, pyramid, duck, train, chair, barrel, staircase, apple, shark, airplane, baseball bat, barbeque, teapot, boat, couch, guitar, and unicycle. See the online article for the color version of this figure.
Experiment 1a, worse in the Test-Move condition than in the Study condition, research on testing effects, final test performance was actually shown in Figures 4 and 5, respectively. Contrary to the vast Results and Discussion door selections were self-paced. pants experienced the same order of rooms as before, and their caused the participant to be transported to the next room. Partici-pants in both conditions then completed the same final test. Participants in both conditions as a function of encoding/test phase. All par-ticipants in the Study condition repeated the encoding three additional times. Participants in the Test-Move conditions as a function of encoding/test phase. All par-ticipants then completed three initial tests over the route, while participants in the Study condition repeated the encoding three additional times. Participants in both conditions then completed the same final test. Error bars represent ±1 SEM. Asterisk indicates a significant difference between conditions at p < .05. Mean errors were always zero on encoding trials, but are artificially elevated in the figure to distinguish data lines from the x-axis.

Results and Discussion

Mean total errors on the final test in Experiments 1a and 1b are shown in Figures 4 and 5, respectively. Contrary to the vast research on testing effects, final test performance was actually worse in the Test-Move condition than in the Study condition, Experiment 1a, t(28) = 2.55, p = .017, d = .92; Experiment 1b, t(30) = 2.30, p = .029, d = .81. Instead of observing the usual benefits of retrieval, Experiments 1a and 1b showed that retrieval impaired learning of the route relative to restudying.

What might explain this reversed testing effect? It is possible that the nature of the route learning task—which differs markedly from the verbal learning tasks that have been used in most research on testing effects—does not lend itself to retrieval-enhanced learning. In particular, the procedural component of route learning (e.g., Gillner & Mallot, 1998; Golledge, 1991) may be sensitive to the effects of erroneous movements (e.g., Maxwell et al., 2001). After moving to the wrong door, even if this error is corrected, participants in the Test-Move conditions may have retained a procedural memory of the erroneous movement that interfered with memory for the correct movement. Such interference would be absent in the Study (i.e., errorless) condition, in which participants always moved toward the correct door without making errors. Experiment 2 was designed to explore the role of movement errors in learning route knowledge through retrieval.

Experiments 2a and 2b

Experiments 2a and 2b explored whether prohibiting movement errors during retrieval would result in positive effects of retrieval on route knowledge acquisition. The design of Experiment 1b, compared to Experiment 1a, appeared less vulnerable to ceiling effects. Therefore, all subsequent experiments used the same number of encoding and test trials as Experiment 1b. Experiments 2a and 2b included a new condition, referred to as Test-Click, in which participants selected a door by clicking on it with the mouse. Door selection was followed by feedback (i.e., highlighting the correct door in green), after which par-ticipants moved toward the correct door. In this way, participants still tried to retrieve the correct door, but did not commit any erroneous movements during retrieval. Experiment 2a com-pared Study and Test-Click performance, and Experiment 2b compared Study, Test-Click, and Test-Move. Examples of each condition are shown in the supplementary video.

Method

Participants. One hundred students (40 in Experiment 2a and 60 in Experiment 2b) at Iowa State University participated in exchange for course credit. Participant gender was balanced across condition within each experiment.

Stimuli, design, and procedure. The virtual environment was identical to the previous experiments. Participants in the Study condition completed the encoding task three times prior to the final test, and participants in the Test-Move and Test-Click conditions completed the encoding task two times followed by one test trial prior to the final test.

![Figure 4](image1.png)

Figure 4. Mean total errors in Experiment 1a for participants in the Study and Test-Move conditions as a function of encoding/test phase. All participants completed an encoding task. Participants in the Test-Move condition then completed three initial tests over the route, while participants in the Study condition repeated the encoding three additional times. Participants in both conditions then completed the same final test. Error bars represent ±1 SEM. Asterisk indicates a significant difference between conditions at p < .05. Mean errors were always zero on encoding trials, but are artificially elevated in the figure to distinguish data lines from the x-axis.

![Figure 5](image2.png)

Figure 5. Mean total errors in Experiment 1b for participants in the Study and Test-Move conditions as a function of encoding/test phase. All participants completed two repetitions of encoding. Participants in the Test-Move condition then completed one initial test over the route, while participants in the Study condition repeated the encoding one additional time. Participants in both conditions then completed the same final test. Error bars represent ±1 SEM. Asterisk indicates a significant difference between conditions at p < .05. Mean errors were always zero on encoding trials, but are artificially elevated in the figure to distinguish data lines from the x-axis.
Encoding was identical to the previous experiments. The only difference between the Test-Move and Test-Click conditions occurred during the initial test. In the Test-Move condition, participants moved to the selected door prior to receiving feedback (just like Experiments 1a and 1b). In the Test-Click condition, participants clicked on the selected door using a mouse, which was followed by visual feedback of the correct door being highlighted in green. If participants clicked on the incorrect door, they were required to click on the correct door before moving toward it. Thus, Test-Move and Test-Click participants relied upon retrieval to choose the correct door, but only Test-Move participants were allowed to make erroneous movements through the environment prior to feedback.

Results and Discussion

Mean total errors on the final test in Experiments 2a and 2b are shown in Figures 6 and 7, respectively. A significant testing effect occurred in Experiment 2a, whereby Study participants committed more final test errors than did Test-Click participants, $t(38) = 2.22, p = .03, d = .70$. The same comparison in Experiment 2b revealed a marginally significant testing effect, $t(38) = 1.89, p = .06, d = .60$. In line with the findings from Experiment 1, the Test-Move condition did not enhance final test performance more than the Study condition in Experiment 2b. Furthermore, participants in the Test-Move condition of Experiment 2b committed more errors than those in the Test-Click condition, $t(38) = 2.23, p = .03, d = .71$, suggesting, as in Experiment 1, that movement errors associated with retrieval impair final test performance.

Figure 7. Mean total errors in Experiment 2b for participants in the Study, Test-Move, and Test-Click conditions as a function of encoding/test phase. All participants completed two repetitions of encoding. Participants in the Test-Move and Test-Click conditions then completed one initial test over the route, while participants in the Study condition repeated encoding one more time. Participants in all conditions then completed the same final test. Error bars represent $\pm 1$ SEM. Asterisk indicates a significant difference between conditions at $p < .05$. Mean errors were always zero on encoding trials, but are artificially elevated in the figure to distinguish data lines from the $x$-axis.

Mean errors on the final test in Experiments 2a and 2b are also shown in Figures 6 and 7, respectively. A significant testing effect revealed a marginally significant testing effect, $t(38) = 2.22, p = .06, d = .60$. In line with the findings from Experiment 1, the Test-Move condition did not enhance final test performance more than the Study condition in Experiment 2b. Furthermore, the room sequence was randomly ordered during encoding, with half of the rooms assigned to Test-Click and half to Study. The room sequence was also randomly ordered during initial testing/restudying to prevent participants from spontaneously retrieving the next room in the sequence.

Experiment 3

Experiment 3 was a replication of Experiment 2a under conditions in which learning condition (Test-Click vs. Study) was manipulated within-participants. After encoding, a randomly selected half of the rooms were assigned to Test-Click and half to Study. Furthermore, the room sequence was randomly ordered during initial testing/restudying to prevent participants from spontaneously retrieving the next room in the sequence.

Method

Participants. Nineteen students at Iowa State University participated in exchange for course credit.

Stimuli, design and procedure. Participants completed the encoding task twice with rooms arranged in the same sequence. Participants then experienced the rooms in a random sequence. Participants were shown the correct door in 15 of the rooms (Study) and were tested and given feedback in the other 15 rooms (Test-Click). As in Experiments 2a and 2b, door selection in the Test-Click condition required clicking with the mouse prior to...
moving through the room. After a 15-min distractor task, participants completed a final test, with rooms appearing in the same sequence as in original encoding.

Results and Discussion

Replicating the findings of Experiments 2a and 2b, more errors occurred on the final test for rooms learned through Study than through Test-Click, \( t(18) = 2.22, p = .04, d = .66 \) (see Figure 8). This indicates that testing can enhance route learning under conditions in which movement errors do not occur during learning.

General Discussion

The current results shed important new light on our understanding of how retrieval affects learning. When participants navigated the route by moving to the door that they believed was correct (Experiments 1a and 1b), participants in the Test-Move condition learned the route less effectively than those in the Study condition, who continued to navigate the route without error. When participants were only allowed to move to the correct door after it was revealed to them (Test-Click conditions of Experiments 2a, 2b, and 3), they learned the route more effectively through testing than through restudying. The effectiveness of retrieval as a strategy for route learning, therefore, appears to depend upon whether participants make movement errors during navigation.

These results are consistent with research on “errorless learning,” which has demonstrated enhanced performance of procedural tasks under conditions in which erroneous movements are minimized (e.g., Maxwell et al., 2001). Although to some extent all learning requires both explicit and implicit memory, it has been proposed that the advantage of errorless learning is the result of implicit memory processes which do not distinguish self-generated errors from correct responses (Page, Wilson, Shiel, Carter, & Norris, 2006). The procedural nature of route knowledge, consisting of actions associated with decision points (Golledge, 1991), may lend itself to this type of processing, rendering it particularly vulnerable to overt movement errors that may be difficult to distinguish from correct movements.

Such processing may be quite different from what takes place during the verbal learning types of tasks that have comprised most of the literature on testing effects. These tasks fall into the category of what is traditionally considered explicit memory. In these tasks—which allow a conscious comparison between self-generated errors and correct responses—significant testing effects are observed even when participants make numerous errors during initial retrieval (Kornell et al., 2009). Although route learning likely involves at least some explicit knowledge, relative to these discrete verbal tasks it is likely to contain an added implicit procedural component (Gillner & Mallot, 1998; Golledge, 1991; Hartley et al., 2003) that is less likely to differentiate between an erroneous movement and the correct movement. As a result, errors made during retrieval could be more detrimental to route learning than to discrete verbal learning.

Previous studies have found that spatial learning methods that involve actively deciding where to travel are no more effective for route learning than passively following directions (Chrastil & Warren, 2012; Wilson, Foreman, Gillett, & Stanton, 1997; Wilson & Péruch, 2002). This appears inconsistent with the findings of Experiments 2 and 3, in which retrieval—which required participants to actively decide where to move—significantly improved route learning compared to passive restudying. However, in the retrieval conditions of the current studies, participants attempted to retrace a previously experienced route, whereas participants in the active conditions of previous studies were instructed to freely explore the space with few constraints or goals other than to simply learn it. Furthermore, navigation errors in the current studies were corrected with feedback, whereas previous studies have not distinguished between correct and incorrect exploration. Lastly, past research on active versus passive exploration has typically tested survey knowledge (e.g., pointing to an unseen landmark) rather than the more procedural type of route knowledge acquisition used in the current study. Therefore, the current results add a new finding to the distinction between active and passive navigation under conditions in which participants try to retrieve knowledge of a previously encoded route, and receive feedback to correct their errors.

Participants in the current experiments learned a sequence of movements that guided them toward their goal. Such sequential movement in response to visual landmarks is the core feature of route knowledge (Gillner & Mallot, 1998; Golledge, 1991). Although this task involves procedural learning, it is important to note that the task itself is not exclusively procedural in nature. It seems quite possible that participants could have incorporated strategies based on explicit verbal processing while navigating their way through the virtual rooms. For example, a participant could consciously recollect “bench: middle” to signify that the middle door is correct in the room containing the bench. The nature and extent of potential verbal strategies during navigation,

![Figure 8](image-url)
and how they might influence or interact with procedural aspects of the task, is presently not well-understood and could represent an interesting area for future research. Verbal strategies could come into play any time during navigation, and overall task performance probably reflects some mixture of both declarative and procedural learning. Given the heavy reliance on declarative verbal learning in most of the research on the testing effect, the current experiments contribute important new data on the nature of retrieval-enhanced learning by identifying the conditions under which retrieval is beneficial—and not beneficial—for types of learning that are not exclusively verbal in nature. Future research could benefit by further exploration of the contribution of verbal processing in procedural tasks, perhaps by using a verbal secondary task to discourage verbal encoding and isolate the aspects of the task that are more influenced by procedural processing.

Web-based driving directions and GPS navigation systems are common tools for navigating one’s environment. However, the current results suggest that testing the navigator’s memory can enhance route learning compared to repeatedly following the routine instructions/guidance, particularly under conditions in which movement errors are avoided. Therefore, a navigator who only follows the directions provided by a GPS system may be missing out on the benefits conferred by retrieval. A navigation system that tests route learning but prevents movement errors seems well within the reaches of current technology, and could provide the benefits of retrieval without the deleterious effects of movement errors. For example, questions when approaching key decision points to prevent movement errors. A comparison between this type of modified route guidance system and traditional systems would be a worthwhile endeavor for future research.

References


Received July 22, 2012
Revision received November 7, 2014
Accepted November 12, 2014