



Retrieval shifts in spatial skill acquisition are collective rather than item-specific

David J. Frank¹ · Brooke N. Macnamara²

© The Psychonomic Society, Inc. 2019

Abstract

How do people improve their ability to intercept moving targets? Prior research and theories of skill acquisition suggest that individuals engage in item-specific retrieval shifts (Anglim & Wynton, 2015; Logan, 1988; Palmeri, 1997; Rickard, 1997, 2004; Touron, 2006; Wilkins & Rawson, 2010). However, this prior research examined performance on nonspatial, nondynamic tasks. In three experiments, we pitted four hypotheses against each other, to test skill acquisition for intercepting repeated trajectories in a spatial and dynamic task: the item-specific algorithmic speedup hypothesis, the item-specific retrieval shift hypothesis, the collective retrieval shift hypothesis, and the combined hypothesis (item-specific algorithmic speedup followed by a collective retrieval shift). We found evidence for the combined hypothesis. Specifically, under easy conditions, we found small improvements on repeated trajectories that were attributable to item-specific algorithmic speedup. By contrast, under difficult conditions, we found strong evidence that the performance benefits for repeated trajectories were driven primarily by a collective shift from algorithmic to direct-retrieval strategies. This evidence for collective retrieval shift is in direct contrast to theories suggesting item-specific retrieval shifts. Theoretical and practical implications are discussed.

Keywords Skill acquisition · Strategy · Spatial · Retrieval · Interception

Humans perform a variety of dynamic spatial interception tasks of varying complexity: A baseball player moves under a fly ball to catch it, a goalie blocks a shot, a boater changes course to avoid a collision with another vessel, or a person swats a fly. Failure in these tasks can be as trivial as losing a pickup game, or as critical as ending up in a serious accident. A large body of research has focused on accomplished interceptors (e.g., Bootsma & Oudejans, 1993; Bootsma & Peper, 1992; Gray, 2002; Gray & Regan, 1998, 2006; Lee, 1976; McBeath, Shaffer, & Kaiser, 1995; McLeod & Dienes, 1996; Michaels & Oudejans, 1992; Regan, 1997). However, little research has examined how the cognitive skill of anticipating the interception point improves with time.

Nonspatial cognitive skill acquisition

With practice, performance rapidly improves on most tasks (e.g., Bryan & Harter, 1897; Fitts & Posner, 1967).

Performance improvements typically occur in two ways: (1) increases in the proficiency of processing (e.g., algorithmic speedup), and/or (2) switching from a less efficient to a more efficient process—that is, a strategy shift (Anglim & Wynton, 2015; Haider & Frensch, 1996; Logan, 1988; Strayer & Kramer, 1990). For example, when learning to read, children increase the proficiency with which they sound out words. However, they eventually switch from a sounding-out algorithm to a direct memory-retrieval, “sight-reading” strategy (Siegler, 1988a). A similar path—from improved algorithmic proficiency to eventual direct retrieval—occurs in mental arithmetic (Delaney, Reder, Staszewski, & Ritter, 1998; Reder & Ritter, 1992; Rickard, 1997, 2004; Siegler, 1988a, 1988b; Tenison & Anderson, 2016; Wilkins & Rawson, 2010) and visual search tasks (Ackerman & Cianciolo, 2002; Ackerman & Woltz, 1994; Wilkins & Rawson, 2010). It is through this shift from an algorithmic to a direct-retrieval strategy that much of our automatic processing works (Logan, 1988). Importantly, these shifts from algorithm to retrieval tend to be item-specific (Anglim & Wynton, 2015; Logan, 1988; Palmeri, 1997; Rickard, 1997, 2004; Touron, 2006; Wilkins & Rawson, 2010). That is, people readily switch between solving novel items using an algorithm and retrieving well-learned items from memory (e.g., Rickard, 1997), as opposed to collectively using one strategy (either algorithm or

✉ David J. Frank
david.frank@tamuc.edu

¹ Texas A&M University - Commerce, Commerce, TX, USA

² Case Western Reserve University, Cleveland, OH, USA

retrieval) exclusively for all items on a given test, with no regard for how well the individual items are learned.

Identifying retrieval shifts

The shift from algorithm to retrieval is typically measured via systematic changes in response times and/or eye movements. For example, in alphabet arithmetic verification tasks, participants verify whether a letter plus a number of subsequent alphabet positions equals another letter (e.g., $E + 5 = J$, $C + 2 = E$). The larger the number (addend) in these problems, the longer it takes people to use the counting algorithm to solve them—an addend effect (e.g., Logan, 1988). So long as the addend effect remains, any reductions in overall response times are thought to result from algorithmic speedup—from learning to count through the alphabet faster. Addend effects eventually disappear, as participants switch from the algorithm (counting through the alphabet) to retrieving the answer directly from memory. That is, they memorize that $E + 5 = J$ and $C + 2 = E$ are correct, and now the two problems are verified using memory in roughly the same amount of time (and much quicker than either can be solved using the algorithm).

Nondynamic tasks do not impose response deadlines. Therefore, researchers typically examine response times in these tasks. For dynamic interception tasks, the moving stimuli impose a response deadline by default. If a participant takes too long to respond, he or she will invariably miss the target. Unlike in alphabet arithmetic tasks, a slow algorithm will result in errors in a dynamic task. We took advantage of this effect by imposing two different response deadlines. If participants are able to accurately respond within a particular deadline, then one can infer that the mental process (algorithm or retrieval) and subsequent motor movements were able to be accurately completed within that time. By contrast, if accuracy is exceedingly poor at a particular response deadline, then one can infer that the physical and mental processes could not be accurately completed within the allotted time with the amount of information processed by the deadline. What is critical is whether the time needed to make accurate responses differs on the basis of whether an item is repeated or novel. If accuracy is relatively high for repeated but not for novel items, then we can infer that the mental process (algorithm or retrieval) has either sped up or changed (i.e., from algorithm to retrieval) as a result of repetition.

Dynamic interception

Interception research has focused almost exclusively on algorithmic strategies (e.g., Bootsma & Oudejans, 1993; Gray, 2002; Gray & Regan, 1998, 2006; Regan, 1997). Recently, we (Frank

& Macnamara, 2017) examined whether participants could spontaneously switch from algorithm to retrieval in a dynamic, spatial task. In these studies, participants performed a computerized interception task (similar to the game Pong; Atari, 1972). The participants' goal was to determine the point of interception and to move a paddle into position to intercept the ball. Critically, participants had to do so before the ball reached one of two lines on a screen (an easy, late response deadline or a difficult, early deadline; see Fig. 1). These response deadlines required participants to anticipate the future location of the ball by using early trajectory information (i.e., launch point and direction) and prevented participants from merely tracking the ball visually throughout the entire trajectory.

We argued that participants shifted from an algorithmic mental extrapolation strategy (visualizing the trajectory of the ball) to a direct-retrieval strategy (retrieving the interception point from memory on the basis of early trajectory information). We based our argument on three key findings. First, when participants practiced intercepting the same four trajectories across blocks, they improved rapidly. However, they were significantly less accurate at intercepting novel trajectories introduced later in the task. Had participants been improving their general ability to mentally extrapolate, they should have performed similarly on the novel trajectories. Relatedly, participants who self-reported that they had used a direct-retrieval strategy performed better than those who indicated they had used a mental extrapolation strategy. This suggests that not only did some participants shift from algorithm to retrieval, but that this retrieval shift was related to superior performance, particularly at the early, difficult response deadline. Given that retrieval was associated with superior performance when quick responding was required, we concluded that the retrieval strategy is likely faster and/or more accurate than algorithmic extrapolation. Finally, participants in an all-novel control group were shown novel trajectories on each trial. These participants not only performed drastically worse than those who responded to repeated trajectories, but also

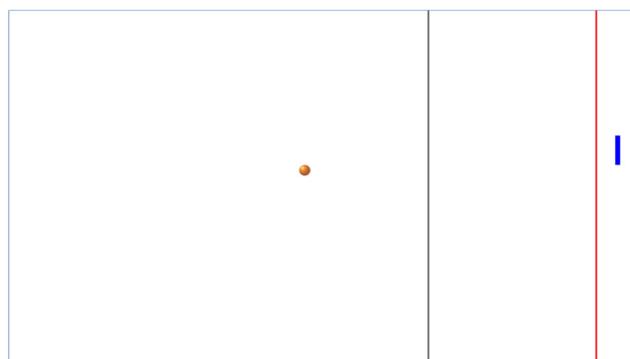


Fig. 1 Screenshot of the Pong task, showing the ball in midtrajectory, the paddle (rectangle on the right side of screen), the learning-phase deadline (line near the paddle), and the earlier, difficult-phase deadline (dark gray line closer to the ball).

showed only minimal improvement across trials. This is again consistent with a direct-retrieval explanation, and suggests minimal evidence for general improvements in algorithmic speed or accuracy in a dynamic spatial task.

Our previous (Frank & Macnamara, 2017) results further suggest that retrieval shift may be collective. Specifically, after seeing all repeated trajectories, some participants indicated that they had reverted back to algorithmic strategies on all trials once novel trajectories were introduced, even though they continued to see well-practiced repeated trajectories as well. This result differs from findings from mathematical and visual search tasks, in which retrieval shifts tend to be item-specific rather than collective (Rickard, 2004; Touron, 2006; Wilkins & Rawson, 2010; but see Haider & Frensch, 2002). Additionally, whereas our previous results were inconsistent with a general algorithmic speedup for repeated trials, it is possible that *item-specific* algorithmic speedup, not retrieval, was responsible for the improvements on repeated trajectories. The present experiments built upon this previous research while directly testing whether retrieval shifts are item-specific or collective.

Comparing hypotheses

In the present study, we examined these possibilities by testing four hypotheses regarding how participants improve their interception for repeated trials: the item-specific algorithmic speedup hypothesis, the item-specific retrieval shift hypothesis, the collective retrieval shift hypothesis, and the combined hypothesis (item-specific algorithmic speedup followed by a collective retrieval shift). None of Frank and Macnamara's (2017) results supported a collective algorithmic speedup hypothesis; therefore, we did not consider this hypothesis in the present study. We also tested whether the strategy adopted by participants depends on the frequency of repeated trials.

To test these hypotheses, we conducted three experiments using an interception task similar to the one in Frank and Macnamara (2017). In each experiment, we assigned half the participants to a repeated and half to a mixed group. The participants in the repeated groups saw the same four or five trajectories throughout the experiment. Two of these repeated trajectories were the "critical repeated trajectories" that the mixed group also saw. The remaining repeated trajectories seen by the repeated group were filler trials. The participants in the mixed group saw the same two critical repeated trajectories that the repeated group saw, but saw them only once per eight novel trajectories (see Table 2 below). During the learning phase, a late response deadline was imposed; during the difficult phase, this deadline was replaced by a difficult, early one. On the basis of each hypothesis, we could predict particular outcomes.

Item-specific versus collective strategy hypotheses

Performance might improve due to *item-specific* strategies. That is, once an item has been encountered multiple times, performance on that item might improve because participants can more quickly trace familiar trajectories (i.e., item-specific algorithmic speedup) or because they recognize the familiar trajectories and retrieve the correct position of the paddle from memory for that trajectory (item-specific retrieval). If performance improves due to either type of *item-specific* strategy, the participants in the repeated and mixed groups should perform similarly on the two critical repeated trajectories throughout the task (see Fig. 2, upper left panel). This is because both groups would have practiced those specific items the same amount and become faster on those items due to either item-specific algorithmic speedup or item-specific retrieval.

In contrast to the item-specific strategies found in nondynamic tasks, performance might improve due to a *collective* shift to retrieval strategy. That is, participants might only shift to retrieving the correct position of the paddle from memory for a given trajectory if they can do so for *all* trajectories in the task. If performance improves due to a *collective* retrieval strategy, participants in the repeated group would outperform those in the mixed group on the two critical repeated trajectories (see Fig. 2, upper right panel). This is because all items for the repeated group would be repeated, allowing for this collective strategy shift to retrieval (i.e., using a retrieval strategy for all items), whereas the mixed groups would experience a mix of repeated and novel items, which would prevent their being able to use retrieval for all items.

Additional information can be gathered by examining novel versus repeated trials within the mixed group. If performance improves due to *item-specific* strategies (either item-specific algorithmic speedup or item-specific retrieval), participants in the mixed group should perform better on repeated than on novel items (see Fig. 2, lower left panel). This is because they would have practiced the repeated items and either improved the speed of their algorithm (item-specific algorithmic speedup) or been able to retrieve the correct position of the paddle from memory for that item (item-specific retrieval); with novel items, there would be no opportunities to practice and improve one's performance.

In contrast, if performance improves due to a *collective* retrieval strategy, participants in the mixed group should perform similarly on repeated and novel items (see Fig. 2, lower right panel). This is because the mix of novel and repeated items would prevent a collective shift to retrieval. In other words, because novel items could not be retrieved from memory and were mixed in with repeated trajectories, participants would abandon using retrieval altogether.

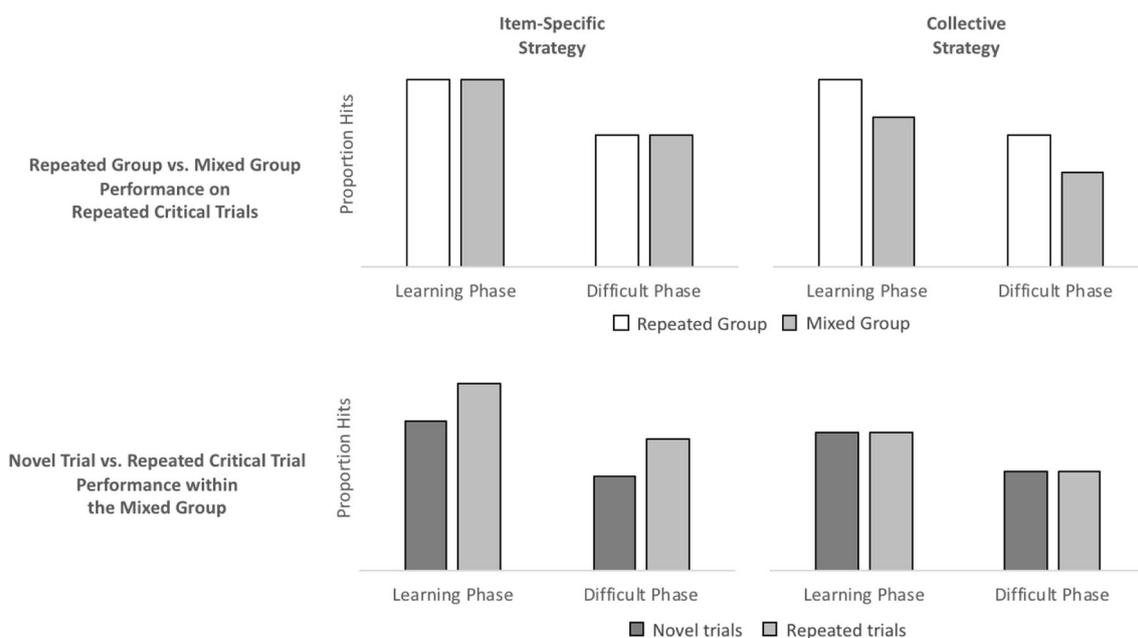


Fig. 2 Predicted patterns of results for repeated critical trials by group (upper row) and for novel trials versus critical repeated trials within the mixed group (lower row), if participants used an item-specific strategy (left graphs) versus a collective strategy (right graphs). The medium gray

bars represent the mixed group's performance on the repeated critical trials in both rows. Thus, the medium gray bars that are directly above/below one another are identical.

Item-specific algorithmic speedup versus item-specific retrieval

Performance might improve due to item-specific *algorithmic speedup*. That is, once an item has been practiced, performance on that particular item might improve because participants can more quickly trace familiar trajectories. If performance improves due to item-specific *algorithmic speedup*, neither the repeated nor the mixed group should experience switching costs or mixing costs—costs to performance from needing to frequently switch strategies (Vandierendonck, Liefoghe, & Verbruggen, 2010; see Fig. 3, left panel). This is because the same strategy (an algorithm) would be used for all items, whether repeated or novel.

In contrast, performance might improve due to item-specific *retrieval*. That is, once participants recognize a familiar trajectory, performance might improve because they can recall where the ball will cross the line and can move the paddle into place quickly. If performance improves due to a shift to item-specific retrieval, there should be switching and mixing costs for the mixed group (see Fig. 3, right panel). This is because the participants in this group would be using an algorithmic strategy for novel items and retrieval for repeated items. The repeated group should not experience switching or mixing costs, because they would only see one type of trial (all repeated) and thus would not need to switch strategies based on trial type.

Switch and mixing costs are typically assessed using response times; however, this is not feasible when a response deadline requires participants to respond by a certain time point. In a task

with a response deadline, mixing and switch costs could deprive participants of valuable time needed to execute retrieval or algorithmic strategies prior to the response deadline, thus resulting in increased performance errors. If strategy switching were occurring, the participants in the mixed group would (1) perform worse on novel trajectories than would participants who only saw novel trajectories—where only a single strategy would be viable, and strategy switches would therefore not be possible—and (2) perform better on repeated trajectories (approaching performance of the repeated group's performance) than on novel trajectories. See Fig. 3, right panel.

Item-specific algorithmic speedup hypothesis

The item-specific algorithmic speedup hypothesis assumes that participants would use a general algorithmic process (i.e., tracing the ball forward) and apply this algorithm to both novel and repeated trajectories. Specifically, participants' performance on repeated trajectories would improve only because they got faster at predicting the ball's interception point by mentally tracing the ball's trajectory for these practiced trajectories. The algorithmic process would become more efficient but otherwise would be unchanged by repetition. This hypothesis would be supported if the participants in the repeated and mixed groups performed similarly on the two critical repeated trajectories. This is because both groups would have practiced mentally tracing the ball's trajectories for these items the same number of times (see Fig. 2, upper left panel; Table 1).

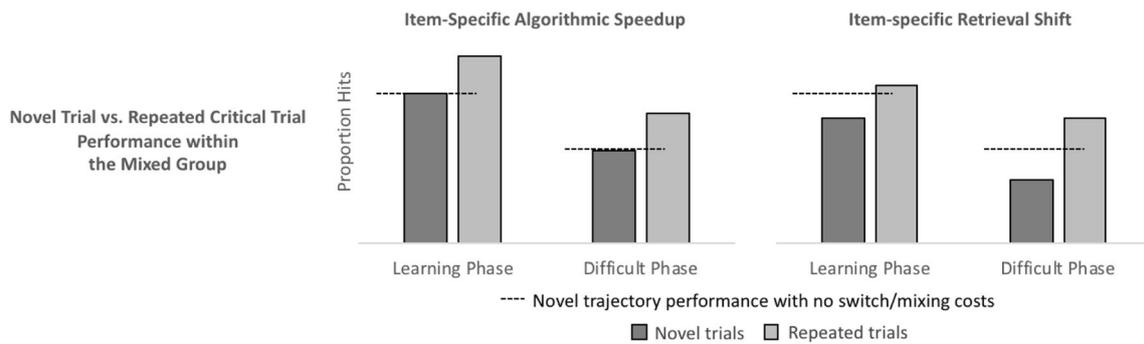


Fig. 3 Predicted patterns of results for novel trials versus critical repeated trials within the mixed group if participants engaged an item-specific algorithmic speedup strategy (left) versus an item-specific retrieval shift

(right). The dashed lines represent the performance levels on novel trajectories without switch/mixing costs.

Additionally, for this hypothesis to be supported, the participants in the mixed group would need to perform better on repeated than on novel trajectories (see Fig. 2, lower left panel). This is because they would have practiced mentally tracing the trajectory for repeated items, but would have no opportunities for this practice on novel trajectories. Critically, the participants in the mixed group should *not* experience switching or mixing costs—costs to performance from needing to frequently switch strategies (Vandierendonck et al., 2010). That is, if item-specific algorithmic speedup is responsible for improved performance on the repeated trials, then the participants in the mixed group would not switch strategies according to the type of trajectory—an algorithm would be used for all trajectories (see Fig. 3, left panel; Table 1).

repeated and mixed groups would perform similarly on the two critical repeated trajectories (see Fig. 2, upper left panel) and participants in the mixed group would improve more on repeated than on novel trajectories (see Fig. 2, lower left panel). The critical difference would be that if the participants in the mixed group were switching strategies based on trial type—algorithm for novel trials and retrieval for repeated trials—they should experience switching and mixing costs (Vandierendonck et al., 2010). That is, performance should be worse following a switch in strategy and when there was a need/option to switch strategies by trial type, relative to when only a single strategy was used (see Fig. 3, right panel; Table 1).

Item-specific retrieval shift hypothesis

The item-specific retrieval shift hypothesis assumes that participants’ performance on repeated trajectories improves because they shift from algorithm to retrieval on practiced trajectories once the trajectories are learned, while continuing to use an algorithm on novel trajectories. In this case, we should observe a pattern of results nearly identical to the one under the item-specific algorithmic speedup hypothesis—participants in the

Collective retrieval shift hypothesis

The collective retrieval shift hypothesis assumes that participants’ performance on repeated trajectories improves because they shift more or less collectively from algorithm to retrieval, producing superior performance on repeated trajectories. That is, participants will use an algorithm when trials include any novel trajectories or before trajectories are learned, but will shift to retrieval only if *all* trajectories have been learned. Our previous (Frank & Macnamara, 2017) results suggested that retrieval shift was collective: Participants who saw all repeated trajectories

Table 1 Comparison of predictions across hypotheses

	Item-Specific Algorithmic Speedup Hypothesis	Item-Specific Retrieval Hypothesis	Collective Retrieval Hypothesis	Combined Hypothesis
Between-group comparison of critical repeated trajectories	repeated group = mixed group	repeated group = mixed group	repeated group > mixed group	<i>Learning phase:</i> repeated group = mixed group; <i>Difficult phase:</i> repeated group > mixed group
Within-group comparison of repeated vs. novel trajectories	repeated > novel	repeated > novel	repeated = novel	<i>Learning phase:</i> repeated > novel; <i>Difficult phase:</i> repeated = novel
Mixing costs in mixed group?	no	yes	no	no

“=” denotes similar performance. “>” denotes better performance by the condition to the left of the symbol.

reported using retrieval until novel trajectories were introduced late in the task, in which case many participants indicated that they reverted back to algorithmic strategies on all trials, even though they continued to see the same repeated trajectories they had been retrieving, as well.

In the present experiments, the collective retrieval shift hypothesis would be supported if the participants in the repeated group outperformed those in the mixed group on the two critical repeated trajectories (see Fig. 2, upper right panel). This is because the participants in the repeated group would use a retrieval strategy for these trials, whereas the participants in the mixed group would use an algorithmic strategy for these trials. Additionally, for this hypothesis to be supported, the participants in the mixed group would need to perform the same on repeated as on novel trajectories (see Fig. 2, lower right panel). This is because when blocks include novel trajectories, as was the case for the mixed group, it should discourage the use of a retrieval strategy, despite providing ample practice on those trajectories that were repeated throughout the task (see Fig. 2, lower right panel; Table 1).

Combined hypothesis

The combined hypothesis assumes that participants benefit from item-specific algorithmic speedup under relatively easy conditions, but shift to an even more efficient collective retrieval strategy when it is viable. That is, when participants can respond close to the point of contact (learning phase), these conditions encourage algorithmic responding. The assumption is that with enough time and trajectory information, an algorithm is less error-prone than retrieval. In contrast, when participants must respond quickly (difficult phase), algorithmic responding is less effective, thereby encouraging a shift to direct retrieval if all trajectories are retrievable (as in the repeated group). The assumption is that algorithms, even those that have become faster, are slower than retrieval. Thus, with limited time to respond, retrieval is less error-prone than an algorithm.

The combined hypothesis would be supported if participants' performance is predicted by item-specific algorithmic speedup in the learning phase, but predicted by a collective shift to retrieval in the difficult phase. That is, during the learning phase, the repeated and mixed groups would perform similarly on the two critical repeated trials, and the participants in the mixed group would perform better on repeated than on novel trajectories. However, during the difficult phase, the repeated group would outperform the mixed group on the two critical repeated trajectories, and the participants in the mixed group would perform similarly on the repeated and novel trajectories (see Fig. 4; Table 1).

Present study

We conducted the present experiments to examine the strategies used on repeated trials in a spatial skill acquisition task, and to evaluate whether the strategies varied by context or difficulty. We varied the context by manipulating the frequency of repetitions between subjects—that is, the repeated group saw only repeated trajectories, and the mixed group saw both repeated and novel trajectories. We varied difficulty by initially imposing an easy, prolonged deadline, under which participants could view almost the entire trajectory before moving to intercept (i.e., learning phase), and then a difficult, early deadline, under which participants needed to attempt to intercept quickly (i.e., difficult phase). The data sets, data analysis scripts, and tasks for this study are available at <https://osf.io/4pjwh/> and <https://osf.io/sjb79/>.

Experiment 1

We initially conducted two nearly identical experiments. On the basis of the large effect sizes ($d = 0.76$ – 1.01) found in Frank and Macnamara (2017) with a similar task comparing accuracy on repeated versus novel trajectories, both within and between subjects, we determined that 68 participants were needed in order to obtain power = .80 to observe a similar effect size for each of these experiments. However, we found substantially smaller effect sizes in these experiments than we had expected, leading to suggestions that we were underpowered. At the request of an anonymous reviewer, we present the experiments as a single, combined experiment. The only difference between these two experiments is that participants completed 16 blocks in the learning phase in Experiment 1a, whereas participants completed 20 blocks in the learning phase in Experiment 1b. The results of Experiments 1a and 1b separately can be found in the supplemental materials and at <https://osf.io/sjb79/>.

Method

Participants

A total of 132 students enrolled in psychology courses at Case Western Reserve University participated in exchange for partial course credit or extra credit.

Materials and procedures

Pong task The Pong task was a modified version of the task used in Frank and Macnamara (2017). This task was created via E-Prime 2 (Schneider, Eschman, & Zuccolotto, 2002) and was performed on computers with $1,920 \times 1,080$ pixel monitors, at a refresh rate of 59 Hz. On each trial, an image of a ball

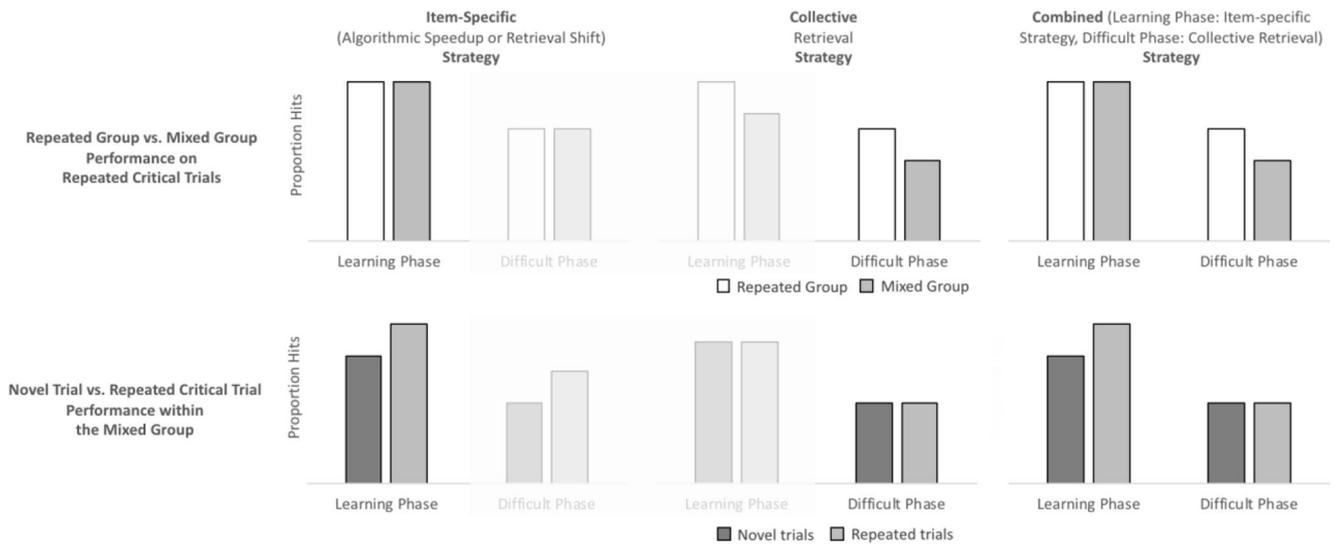


Fig. 4 Predicted patterns of results for repeated critical trials by group (upper right) and for critical novel versus repeated trials for the mixed group (lower right) if participants use a combined strategy. Specifically, if participants use an item-specific algorithmic strategy during the learning phase (left column, nonfaded bars), but switch to a collective retrieval strategy during the difficult phase, if such a strategy is viable (i.e., in the

repeated group; center column, nonfaded bars), the predicted pattern of results is depicted in the right column as a combined strategy. The medium gray bars represent the mixed group’s performance on the repeated critical trials in both rows. Thus, the medium gray bars that are directly above/below one another are identical.

moved from the left to the right side of the screen, bouncing once in the process (see Fig. 1). The ball moved to the right at a constant 565 pixels per second. Simulated gravity decelerated the ball by 11 pixels per second. The elasticity of the ball was set to 95%. Launch position varied along the *y*-axis, and launch angle was manipulated by changing the initial vertical velocity of the ball. The ball moved as if in a vacuum (i.e., no effects of air resistance) and without spin or friction. Each trajectory lasted 3.4 s.

A blue paddle was located 80 pixels from the right edge of the screen. Participants moved the mouse in order to vertically position the paddle. Each trial began with the participant clicking a “Start” button in the center of the *y*-axis on the right side of the screen. This ensured that the paddle was in the same position at the start of each trial. The ball was then launched, and the participant attempted to position the paddle to intercept it. When the paddle was in position, the participant clicked the left mouse button to “freeze” the paddle. The participant had to click the mouse before the ball crossed a vertical line—a response deadline—on the screen; otherwise, the paddle would disappear and fail to intercept the ball. After the mouse button had been clicked, the paddle would turn black and would no longer move until the next trial. If the ball hit the paddle, the ball bounced back to the left; otherwise, it continued off screen with a 1.4-s delay, to allow for visual processing of this feedback.

Two difficulty levels were used. During the learning phase, participants clicked the paddle into position before the ball reached a vertical red line, positioned 140 pixels from the right side of the screen, giving participants 3.25 s to respond before

the ball reached the red line. Later in the task, participants advanced to the difficult phase, in which they were tasked with clicking the paddle into position before the ball reached a vertical gray line, located 650 pixels from the right side of the screen, decreasing the time to respond to roughly 2.25 s (see Fig. 1). The paddle remained in the same position on the right side of the screen for both deadlines, and the goal was to intercept the ball with the paddle on the right side of the screen in both cases. The only difference between the two deadlines was how early in the ball’s trajectory the participant was required to position the paddle and commit to this position. Participants were initially told to click the paddle into place before the ball reached the red line, but they received the following warning:

Later in the task, the paddle will remain next to the red line, but we will increase the difficulty by requiring you to click the paddle into position before the ball reaches the GRAY line. Therefore, you should try to click the paddle into position as early as possible during the easier part of the experiment so that you are ready when the experiment gets harder.

Repeated group: Learning phase During the learning phase, participants completed blocks of four trials. Each block used the same four trajectories: A, B, C, and D, shown in Fig. 5. (Trajectories B and C were the critical repeated trajectories that the mixed group also saw.) Trajectory order was

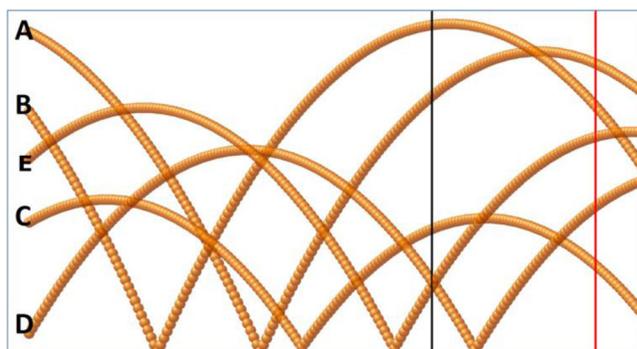


Fig. 5 Repeated trajectories. Trajectories A, B, C, and D were used in Experiments 1 and 2, and trajectory E was used in Experiment 2 only. Trajectories B and C are the critical repeated trajectories.

randomized within blocks, with the exception that the same trajectory that ended one block could not begin the next block. (See Table 2 for an example.)

Repeated group: Difficult phase Following the learning phase, participants saw the message, “You will now have to position the paddle and click it into place before the ball reaches the GRAY line.” Afterward, participants completed 20 blocks using the same four repeated trajectories; however, only the first ten blocks were analyzed, to equate the number of repetition trials with the mixed group.

Mixed group: Learning phase During the learning phase, participants completed blocks of eight trials. Each block was

Table 2 Example stimuli for Experiments 1 and 2

		Experiment 1		Experiment 2	
		Repeated	Mixed	Repeated	Mixed
Block 1	Mini-Block	B	Novel	B	Novel
		D	B	E	B
		A	Novel	A	Novel
		C	Novel	D	Novel
	Mini-Block		Novel	A	Novel
			Novel	E	Novel
		C	C	C	
		Novel	D	Novel	
Block 2	Mini-Block	D	Novel	E	Novel
		B	Novel	A	Novel
		C	B	B	B
		A	Novel	D	Novel
	Mini-Block		Novel	E	Novel
			C	D	C
		Novel	A	Novel	
		Novel	C	Novel	

Novel = novel trajectories. Letters indicate trajectories A, B, C, D, and E. Bold letters indicate the critical trials.

separated into two equal mini-blocks. Mini-blocks contained three novel trajectories and either trajectory B (odd mini-blocks) or trajectory C (even mini-blocks). This procedure provided the mixed group with twice as much practice overall, but the same number of repetitions of trajectories B and C as in the repeated group (see Table 2).

Mixed group: Difficult phase After the learning phase, participants saw the same message as those in the repeated group, before performing ten blocks of the difficult phase. The difficult-phase block composition was the same as in the learning phase. The entire procedure lasted approximately 45 min.

Results

The alpha level was set to .05. The dependent variable was accuracy—the percentage of trials in which the participant successfully hit the ball. Confirmatory analyses were preregistered via the Open Science Framework (<https://osf.io/tz5rb/>).

Repeated-trajectory accuracy

Learning phase We removed the first block of trials from the analysis because, at that point in the task, those trajectories were still novel (rather than repeats). When collapsing across the remaining repetitions of trajectories B and C for the learning phase, accuracy was lower in the repeated group ($M = .53$, $SD = .17$) than in the mixed group ($M = .59$, $SD = .17$), $t(134) = -2.32$, $p = .022$, $d = -0.40$). See Fig. 6, upper center panel. This finding, though unexpected, can be explained by item-specific algorithmic speedup. In this case, the algorithm appears to have been the superior strategy during the learning phase.

Difficult phase When collapsing across the first ten repetitions of trajectories B and C for the difficult phase, accuracy was higher in the repeated group ($M = .38$, $SD = .16$) than in the mixed group ($M = .32$, $SD = .16$), $t(134) = 2.22$, $p = .028$, $d = 0.38$. See Fig. 6, upper center panel. This finding suggests a collective retrieval shift.

Repeated versus novel-trajectory accuracy in the mixed group

During the learning phase, repeated trajectories ($M = .59$, $SD = .17$) were hit more often than novel trajectories ($M = .54$, $SD = .113$), $t(69) = 4.79$, $p < .001$, $d = 0.57$, within the mixed group. Difficult-phase accuracy was also higher for repeated trajectories ($M = .32$, $SD = .16$) than for novel trajectories ($M = .29$, $SD = .10$), $t(69) = 2.13$, $p = .037$, $d = 0.25$. See Fig. 6, lower center panel. The results from both the learning and difficult phases suggest that the mixed group used an item-specific strategy.

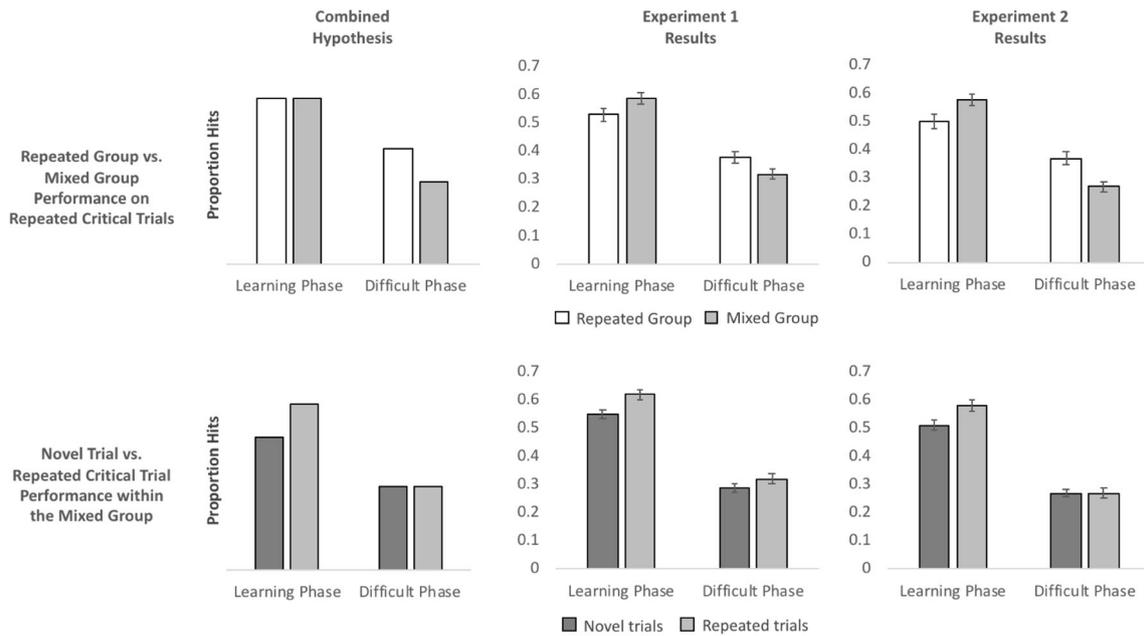


Fig. 6 Expected results according to the combined hypothesis (left column), along with the Experiment 1 results (center column) and Experiment 2 results (right column), for repeated critical trials by group (upper row) and for novel versus repeated critical trials for the mixed

group (lower row). The medium gray bars represent the mixed group’s performance on the repeated critical trials in both rows. Thus, the medium gray bars that are directly above/below one another are identical.

Mixed-group trajectory accuracy comparison

Our finding that accuracy was higher for repeated than for novel trajectories in the mixed group suggests that participants relied on either item-specific algorithmic speedup or item-specific retrieval. To investigate this, we examined accuracy in the mixed group’s difficult phase in relation to the accuracy from Frank and Macnamara’s (2017) Experiment 2, which included an all-novel-trajectory group, for whom the algorithm was the only possible option. The novel-trajectory performance in the difficult phase in the present experiment ($M = .29$) was nearly identical to (and, if anything, better than) the difficult-phase novel trajectory performance for the all-novel group in Frank and Macnamara’s Experiment 2 ($M = .25$). Hence, the participants in the present experiments likely were not mixing their strategies. If they had, they would have shown switch/mixing costs on novel trajectories (relative to Frank & Macnamara, 2017) and retrieval benefits during the difficult phase more akin to their repeated-group counterparts in the present experiment. Neither of these outcomes occurred.

Discussion

As is predicted by the combined hypothesis, we observed evidence for item-specific algorithmic speedup during the learning phase for both groups, as well as evidence for a collective shift to retrieval in the difficult phase for the repeated group (see Fig. 6, left and center columns). That is, during the learning phase, performance was highest for repeated

trajectories in the mixed group, who, by the end of the learning phase had the benefit of twice as much practice overall on the Pong task. During the difficult phase, participants in the repeated group outperformed those in the mixed group on repeated trajectories. This occurred despite equal exposure to the repeated trajectories B and C in both groups. Additionally, the participants in the mixed group hit repeated trajectories more often than novel trajectories but showed no evidence for switching or mixing costs. This suggests that the mixed group continued to benefit from item-specific algorithmic speedup in the difficult phase, though this strategy was not as beneficial as the collective shift to retrieval that was viable for the repeated group.

An important difference between the repeated and mixed groups in this experiment was that the participants in the repeated group saw trajectories B and C once each for every four trajectories. By contrast, the participants in the mixed group saw these trajectories only once each for every eight trajectories. Thus, more time elapsed between B and C repetitions for the mixed group. Although spaced repetitions have been shown to improve retention following a delay, they often slow initial learning (Ebbinghaus, 1885). We addressed this issue with Experiment 2.

Experiment 2

In Experiment 2 we increased the number of repeated trajectories in the repeated group from four to five. This allowed us

to show trajectories A, D, and E (Fig. 5) twice per block, but to show trajectories B and C once per block—thus equating the frequencies of these critical trajectories and the total number of trials for the repeated and mixed groups (see Table 2). We sought to replicate the effects of Experiment 1.

Method

Participants

Given the effect sizes obtained in Experiment 1, we optimized our sample size so as to obtain a power of .80 for an effect size of $d = 0.50$, yielding a required sample of 136. Thus, we terminated data collection at the end of the week in which the 68th participant was collected for each of the two research sites. Two participants were excluded from the data analysis because the experimenter observed noncompliant behavior during the session (i.e., checking their phone and/or not responding to the task). After this exclusion, there were exactly 68 participants from each of two universities (Case Western Reserve University and Texas A&M University–Commerce). However, there were two more participants in the mixed group (69) than in the repeated group (67). All participants were enrolled in psychology courses and participated in exchange for partial course credit or extra credit.

Procedures

The procedures were identical to those of Experiment 1b, except that we increased the number of repeated trajectories in the repeated group from four to five. This allowed us to show trajectories A, D, and E once per mini-block but to show trajectories B and C alternately between mini-blocks, as in the mixed group.

Results

The primary analyses were identical to those reported in Experiment 1 and were preregistered via the Open Science Framework (<https://osf.io/bu68m>).

Repeated-trajectory accuracy

Learning phase We removed the first block of trials from the analysis because those trajectories were novel rather than repeated. When collapsing across the remaining 19 repetitions of trajectories B and C for the learning phase, accuracy was again lower in the repeated group ($M = .50$, $SD = .20$) than in the mixed group ($M = .58$, $SD = .17$), $t(134) = -2.40$, $p = .018$, $d = -0.41$. See Fig. 6, upper right panel. This occurred despite equal practice for both groups. This is the same pattern of results found in Experiment 1 and suggests an item-specific strategy.

Difficult phase When we collapsed across all ten repetitions of trajectories B and C, accuracy was higher in the repeated group ($M = .37$, $SD = .19$) than in the mixed group ($M = .27$, $SD = .14$), $t(122.51) = 3.34$, $p = .001$, $d = 0.58$.¹ See Fig. 6, upper right panel. This is the same pattern of results found in Experiment 1 and suggests a collective shift to retrieval.

Repeated versus novel-trajectory accuracy in the mixed group

During the learning phase, repeated trajectories ($M = .58$, $SD = .17$) were hit more often than novel trajectories ($M = .51$, $SD = .14$), $t(68) = 5.03$, $p < .001$, $d = 0.64$, within the mixed group. In contrast to the learning phase, difficult-phase accuracy did not differ between repeated trajectories ($M = .27$, $SD = .14$) and novel trajectories ($M = .27$, $SD = .10$), $t(68) = 0.41$, $p = .683$. See Fig. 6, lower right panel. The results from both the learning and difficult phases suggest that the mixed group used an item-specific strategy.

Mixed-group trajectory accuracy comparison

As with Experiment 1, we compared accuracy in the mixed group's difficult phase with the accuracy from Frank and Macnamara's (2017) Experiment 2, which included an all-novel-trajectory group, for whom the algorithm was the only possible option. The novel-trajectory performance in the difficult phase in the present experiment ($M = .27$) was nearly identical to that in both Experiment 1 ($M = .29$) of the present article and the all-novel group in Frank and Macnamara's Experiment 2 ($M = .25$). Hence, the participants in the present experiments likely were not mixing their strategies. If they had, they would have shown switch/mixing costs on novel trajectories (relative to Frank & Macnamara, 2017) and retrieval benefits during the difficult phase more akin to those in the repeated group in the present experiment. Neither of these outcomes occurred.

Discussion

Experiment 2 equated the mixed and repeated groups in terms of both overall task practice and frequency of trajectory repetition for the critical trajectories (B and C). The results were largely the same as in Experiment 1: When participants could respond close to the point of contact (learning phase), conditions that encouraged algorithmic responding produced better performance. In contrast, when participants were required to respond early (difficult phase), conditions that encouraged a

¹ We used Satterwaite-corrected degrees of freedom and p values because the test for equivalence of variances was violated, $F(66, 68) = 1.77$, $p = .021$. The results were similar when using the pooled variance.

direct-retrieval strategy produced superior performance. Thus, again our findings supported the combined hypothesis (see Fig. 6).

As in Experiment 1, we found that accuracy on repeated trajectories was higher, even under conditions that encouraged algorithm use. Unlike in Experiment 1, the participants in Experiment 2 had equal overall task practice during the learning phase. This suggests that participants may have benefited from some degree of item-specific algorithmic speedup. However, this benefit was again restricted primarily to the learning phase, when participants had more time to respond. This suggests that the participants in the mixed group did not shift to a direct-retrieval strategy at the item level, but instead continued using an algorithmic strategy for all trajectories. If they had switched to a direct-retrieval strategy for repeated trajectories (i.e., item-specific retrieval), they should have benefited from this strategy during the difficult phase, and so performed as well as the repeated group, which they did not. Likewise, if the mixed-group participants were using retrieval for repeated trajectories during the learning phase, they should *not* have outperformed the repeated-group participants during that phase. These findings also suggest that the participants in the repeated group shifted to a direct-retrieval strategy during the learning phase, even though doing so impaired their performance. This behavior is reasonable, given that the participants were encouraged to respond as early as possible in preparation for the difficult phase.

General discussion

The item-specific algorithmic speedup hypothesis predicts that performance improves, not due to retrieval shift, but to item-specific increases in algorithmic efficiency for repeated trajectories. By contrast, the item-specific retrieval shift hypothesis predicts that people shift from algorithm to retrieval at the item level, regardless of how often retrieval is viable. The collective retrieval shift hypothesis predicts that people more readily shift from an algorithmic strategy to a direct-retrieval strategy only when the task conditions allow for retrieval use on all trials. Finally, the combined hypothesis predicts that people's strategy use depends on the context and difficulty of the task.

The data clearly support the combined hypothesis—people benefit from item-specific algorithmic speedup when the task lends itself to algorithmic strategies—that is, when only an easy, late deadline is imposed, through which almost the entire trajectory can be tracked prior to response. However, people shift *collectively* to retrieval when all items are familiar and when a difficult, early deadline is imposed, so that less trajectory information is available prior to a response. Specifically, when participants were required to anticipate the future location of the ball early in the trajectory (difficult phase), they performed markedly better on critical trials if all of the trajectories

were repeated and thus could be solved using retrieval (repeated group). In contrast, the mixed group, who saw both novel and repeated trials, did not perform as well as the repeated group on the critical repeated trials, even when they saw the critical trials as frequently as did the participants in the repeated group.

Our data suggest that the participants in the mixed group used an algorithmic strategy for both repeated and nonrepeated trajectories, and during both the learning and difficult phases. Item-specific algorithmic speedup produced better accuracy for repeated trajectories when using an easier, late response deadline (learning phase). By contrast, the algorithmic speedup was not substantial enough to be of benefit when the difficult, early response deadline was imposed (i.e., during the difficult phase). The participants in the repeated group appeared to switch to a direct-retrieval strategy in an attempt to respond more quickly in preparation for the impending difficult phase. The participants in the repeated group were outperformed by the participants in the mixed group on repeated trajectories during the learning phase, but they outperformed the participants in the mixed group on repeated trajectories during the difficult phase. Thus, we found evidence of a slight benefit for item-specific algorithmic speedup during easier, late deadlines, and evidence of a benefit for retrieval during more difficult, early deadlines—that is, the combination hypothesis.

Implications for existing theories

Instance theory Logan's (1988) *instance theory* (sometimes referred to as a "race model") states that each time an item is encountered and correctly solved, a new neural trace is laid down. When the item is encountered again, an attempt to solve the item via algorithm begins. Meanwhile, a simultaneous attempt to locate the answer in memory takes place. These two processes—algorithm and retrieval—race against each other. Whenever one of the processes produces the answer, the other process terminates. As an item is encountered more times, more neural traces are laid down, and the probability of locating the item in memory before the algorithm completes increases. Our finding of a collective retrieval shift does not fit with item-specific instance theory. Our results suggest that instance theory does not apply to at least some forms of spatial skill acquisition.

Strategy reversions are not predicted by instance theory, since the race between algorithm and retrieval should be unaffected by novel items. However, Frank and Macnamara (2017) found that participants in a repeated group sometimes reported reverting back to using an algorithm on learned trajectories after novel trials were introduced. This is in contrast to research using static verbal and mathematical tasks, in which strategy reversions are rare and do not typically increase following the introduction of novel trials (Rickard, 1997; Rickard & Bourne, 1996). Thus, instance theory does not hold up to evidence from a dynamic spatial task.

Finally, in contrast to instance theory, our findings suggest a choice (conscious or otherwise) to engage in either an algorithm or a retrieval attempt. This argument has previously been made by Touron and colleagues (Frank, Touron, & Hertzog, 2013; Touron, 2006; Touron & Hertzog, 2004) and Bajic and Rickard (2009, 2011), with respect to static stimuli. Touron and colleagues demonstrated that older adults are particularly reluctant to use retrieval-based strategies, even after the requisite information has been learned. Bajic and Rickard (2009, 2011) demonstrated that once participants initiate an algorithm, they rarely terminate the algorithm prior to its completion (as would be expected if retrieval won the race in instance theory). Thus, the present experiments add to mounting evidence against instance theory in favor of choice-based models (to be discussed shortly).

Exemplar-based random walk Palmeri's (1997) exemplar-based random-walk model (EBRW) extends and modifies Logan's (1988) instance theory. Both theories argue that multiple exemplars race to be retrieved from memory, and that the more exemplars (instances) that have been stored in memory, the faster the retrieval. Unlike instance theory, EBRW argues that items that share a response (similar interception points, in the present study) are categorized together. The similarity of the current item to the exemplars in a given category will dictate the speed with which retrieval occurs. This means that when an item is repeated, its identical nature to its past occurrences should allow for very fast retrieval.

This brings up an alternative explanation for our results, that both the repeated and mixed groups used retrieval on repeated trajectories, but that the participants in the mixed group had decreased accuracy due to interference from novel trajectories that might appear similar to the repeated trajectories, but that afford a different response (different interception point). If the early portions of many trajectories were similar, this would slow response times and might lead to misretrievals (mistaking of one trajectory for a similar trajectory). This could explain why the participants in the mixed condition performed poorly at an early response deadline (difficult phase).

Although we cannot completely rule this explanation out, we have reasons to be skeptical. First, interference cannot explain why the participants in the mixed group generally outperformed the repeated-group participants on repeated trajectories during the learning phase. If the mixed group was using a slowed retrieval strategy (or misretrieving) during the learning phase, they should have performed worse, not better, than the repeated group. Second, if the participants in the mixed group confused a novel trajectory for a repeated trajectory, this should have led to increased errors on the novel, not the repeated, trajectories.

It is possible that retrieval took so long for the mixed group that the algorithm finished first during the learning phase and proved to be the more accurate strategy, whereas both the

algorithm and retrieval were too slow during the difficult phase. However, this explanation is at odds with Frank and Macnamara's (2017) finding that participants reported reverting back to the algorithm after novel trials had been introduced. Furthermore, in these two experiments (Frank & Macnamara, 2017), participants performed 40 repetitions of four trajectories before one novel trajectory was added to each block of trials for the next 14 blocks. It seems unlikely that such a small number of novel trajectories would create substantial response competition that could slow retrieval to the point that the algorithm would complete prior to retrieval.

We suspect that potential interference encourages participants to use the algorithm. It is possible that the participants in the mixed group did retrieve the answer but did not trust this retrieval, because it was often inaccurate when novel trajectories were mistaken for repetitions. Thus, they may have retrieved but ultimately relied on an algorithmic strategy as a backup. This is in line with Touron and colleagues' (Frank et al., 2013; Touron & Hertzog, 2004) finding that older adults' retrieval use is driven, in part, by a reluctance to trust their memory. However, retrieval reluctance in younger adults is not typically found. Regardless, the present data do not support the hypothesis that the participants in the mixed group were frequently using retrieval as their principal strategy.

Component power laws theory In contrast to race models, Rickard's (1997) component power laws theory (CMPL) argues that the algorithm and retrieval do not race against each other, but rather that a choice to retrieve is made on the basis of familiarity with the item. Additionally, in CMPL there are not multiple neural traces, but rather a single trace that is strengthened each time an item is encountered. Like instance theory, CMPL predicts an item-specific retrieval shift based on the number of times an item has been encountered. Unlike both instance theory and EBRW, CMPL leaves open the possibility that the algorithm might be chosen despite retrieval being possible (and potentially faster). This would allow for top-down strategic factors to influence the retrieval shift process (Frank et al., 2013; Touron & Hertzog, 2004; Wilkins & Rawson, 2011). This might have been applicable to our study. Here, a participant in the mixed group could strategically choose to use the algorithm on all trials in order to avoid potential retrieval failures, in which a novel item would be mistaken for a repeated item, or to avoid potential mixing costs. By contrast, the participants in the repeated groups did not face these issues and could instead have chosen the benefits of retrieval with little concern about retrieval failures. Note that mixing costs, in particular, are potentially more harmful in dynamic tasks in which response time is limited, and this may explain why Frank and Macnamara (2017) observed strategy reversions not typically seen in static tasks (e.g., Rickard, 1997; Rickard & Bourne, 1996).

As with both instance theory and EBRW, CMPL does not predict the strategy reversions that Frank and Macnamara's (2017) participants reported. However, CMPL's decision process would allow participants to adopt a wholesale algorithmic strategy, in an attempt to avoid switching and mixing costs that might otherwise impair performance on novel trajectories.

Theory conclusion Each of the theories above was developed using data from static verbal or mathematical skill acquisition tasks. For such tasks, retrieval shifts tend to be item-specific, rather than collective (Anglim & Wynton, 2015; Rickard, 2004; Touron, 2006; Wilkins & Rawson, 2010). By contrast, our data suggest that in our dynamic spatial task, retrieval shifts would be adopted only when all items are repeated. Furthermore, Frank and Macnamara (2017) suggested that retrieval strategies might also be collectively abandoned when novel stimuli are introduced. Together, the results from our previous and present studies suggest that the item-specific retrieval shift assumption does not hold in dynamic spatial tasks. Future theories of skill acquisition will need to be adjusted so as to account for differences between static verbal and mathematical tasks and dynamic spatial tasks.

Limitations

The present data suggest a general tendency toward retrieval-based responding when retrieval is possible on all trials under difficult conditions, and continued algorithmic processing when retrieval is not possible on all trials. However, there may be a good degree of heterogeneity in strategy use (see Frank & Macnamara, 2017). Additionally, participants might vary in the amount and type of trajectory information used in each strategy. Our present paradigm cannot examine this topic, which could be a fruitful direction for future investigations.

In our present study, we have assumed that the algorithms used on novel and repeated trajectories are either the same or similar, only sped up for practiced trajectories. It could be that these algorithms differ. However, comparisons with Frank and Macnamara's (2017) all-novel group's performance do not suggest this to be the case. That is, we found no evidence of mixing costs in the present study for the mixed group, which presumably would have occurred if participants were switching between different algorithmic strategies for novel and practiced trajectories.

Response times are often used to infer strategy use from performance data, particularly when distinguishing between algorithmic and retrieval-based strategies (e.g., Tenison & Anderson, 2016; Wynton & Anglim, 2017). Our paradigm employed a response deadline rather than response times. This was done so that participants had to mentally anticipate the future location of the ball, rather than track the ball throughout the entire trajectory. A direction for future research

would be to examine these hypotheses in a paradigm in which response times could be examined.

Conclusion

We found evidence for both item-specific algorithmic speedup and collective retrieval shifts in a spatial skill acquisition task, depending on the context: Algorithmic speedup appeared to be beneficial at a long response deadline, whereas direct retrieval appeared to be the superior strategy when an early response deadline was imposed. Furthermore, when retrieval was a viable strategy throughout the task, our results suggest that participants readily adopted this strategy. In contrast, when retrieval was a viable strategy only on occasion, our results suggest that participants maintained an algorithmic strategy.

When dynamic spatial stimuli are encountered repeatedly, people appear to readily adopt, and to greatly benefit from, a direct-retrieval strategy—but only if response time is limited. When novel stimuli are common and make up the majority of trials, a direct-retrieval strategy appears less likely to be used. This has real-world implications for sports and other applied settings, because it suggests that until the majority of possible trajectories can be learned, people are unlikely to use direct-retrieval strategies. In a game such as baseball or cricket, only a (relatively) small subset of trajectories will send the ball into the strike zone, and thus direct retrieval could be a viable option, given enough practice. By contrast, there are likely far more unique trajectories in tennis or soccer that will put the ball in play. Thus, a direct-retrieval strategy could require substantially more practice in order to be beneficial in tennis or soccer, as compared to baseball or cricket, if such a strategy could be adopted at all.

Author note We thank Jeffrey Antoine, Dylan Brown, Jonathan Daniels, Jessie Kierbow, Michael King, Natasha Rupani, Jason Smith, Srivatsan Uchani, Jiayuan Zhang, and Yuxuan Zhao for assistance with the data collection. All data are available on the Open Science Framework (<https://osf.io/4pjhwh/>). Experiments 1a and 2 were preregistered on the Open Science Framework (<https://osf.io/tz5rb/> and <https://osf.io/bu68m>, respectively). No conflicts of interest are reported, and all research was in compliance with the ethical standards of the American Psychological Association.

References

- Ackerman, P. L., & Cianciolo, A. T. (2002). Ability and task constraint determinants of complex task performance. *Journal of Experimental Psychology: Applied*, 8, 194–208. doi:<https://doi.org/10.1037/1076-898X.8.3.194>
- Ackerman, P. L., & Woltz, D. J. (1994). Determinants of learning and performance in an associative memory/substitution task: Task

- constraints, individual differences, volition, and motivation. *Journal of Educational Psychology*, 86, 487–515. doi:<https://doi.org/10.1037/0022-0663.86.4.487>
- Anglim, J., & Wynton, S. K. A. (2015). Hierarchical Bayesian models of subtask learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 41, 957–974. <https://doi.org/10.1037/xlm0000103>
- Atari. (1972). Pong. Sunnyvale, CA: Atari Inc.
- Bajic, D., & Rickard, T. C. (2009). The temporal dynamics of strategy execution in cognitive skill learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 35, 113–121. doi:<https://doi.org/10.1037/a0013647>
- Bajic, D., & Rickard, T. C. (2011). Toward a generalized theory of the shift to retrieval in cognitive skill learning. *Memory & Cognition*, 39, 1147–1161. doi:<https://doi.org/10.3758/s13421-011-0114-z>
- Bootsma, R. J., & Oudejans, R. R. D. (1993). Visual information about time to collision between two projectiles. *Journal of Experimental Psychology: Human Perception and Performance*, 19, 1041–1052. doi:<https://doi.org/10.1037/0096-1523.19.5.1041>
- Bootsma, R. J., & Peper, C. E. (1992). Predictive visual information sources for the regulation of action with special emphasis on catching and hitting. In L. Proteau & D. Elliott (Eds.), *Advances in psychology*: Vol. 85. Vision and motor control (pp. 285–314). Amsterdam, The Netherlands: North-Holland. doi:[https://doi.org/10.1016/S0166-4115\(08\)62019-1](https://doi.org/10.1016/S0166-4115(08)62019-1)
- Bryan, W. L., & Harter, N. (1897). Studies in the physiology and psychology of the telegraphic language. *Psychological Review*, 46, 27–53. doi:<https://doi.org/10.1037/h0073806>
- Delaney, P. F., Reder, L. M., Staszewski, J. J., & Ritter, F. E. (1998). The strategy-specific nature of improvement: The power law applies by strategy within task. *Journal of Memory and Language*, 9, 120–130. doi:<https://doi.org/10.1111/1467-9280.00001>
- Ebbinghaus, H. (1885). *Memory: A contribution to experimental psychology*. New York, NY: Teachers' College, Columbia University.
- Fitts, P. M., & Posner, M. I. (1967). *Human performance*. Belmont, CA: Brooks/Cole.
- Frank, D. J., & Macnamara, B. N. (2017). Does the acquisition of spatial skill involve a shift from algorithm to memory retrieval? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 43, 1845–1856. doi:<https://doi.org/10.1037/xlm0000429>
- Frank, D. J., Touron, D. R., & Hertzog, C. (2013). Age differences in strategy shift: Retrieval avoidance or general shift reluctance? *Psychology and Aging*, 28, 778–788. doi:<https://doi.org/10.1037/a0030473>
- Gray, R. (2002). Behavior of college baseball players in a virtual batting task. *Journal of Experimental Psychology: Human Perception and Performance*, 28, 1131–1148. doi:<https://doi.org/10.1037/0096-1523.28.5.1131>
- Gray, R., & Regan, D. (1998). Accuracy of estimating time to collision using binocular and monocular information. *Vision Research*, 38, 499–512. doi:<https://doi.org/10.1016/S0042-698900230-7>
- Gray, R., & Regan, D. M. (2006). Unconfounding the direction of motion in depth, time to passage and rotation rate of an approaching object. *Vision Research*, 46, 2388–2402. doi:<https://doi.org/10.1016/j.visres.2006.02.005>
- Haider, H., & Frensch, P. A. (1996). The role of information reduction in skill acquisition. *Cognitive Psychology*, 30, 304–337. doi:<https://doi.org/10.1006/cogp.1996.0009>
- Haider, H., & Frensch, P. A. (2002). Why aggregated learning follows the power function of practice when individual learning does not: Comments on Rickard (1997, 1999), Delaney et al. (1998), and Palmeri (1999). *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28, 392–406. doi:<https://doi.org/10.1037/0278-7393.28.2.392>
- Lee, D. N. (1976). A theory of visual control of braking based on information about time-to-collision. *Perception*, 5, 437–459. doi:<https://doi.org/10.1068/p050437>
- Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review*, 95, 492–527. doi:<https://doi.org/10.1037/0033-295X.95.4.492>
- McBeath, M. K., Shaffer, D. M., & Kaiser, M. K. (1995). How baseball outfielders determine where to run to catch fly balls. *Science*, 268, 569–573. doi:<https://doi.org/10.1126/science.7725104>
- McLeod, P., & Dienes, Z. (1996). Do fielders know where to go to catch the ball or only how to get there? *Journal of Experimental Psychology: Human Perception and Performance*, 22, 531–543. doi:<https://doi.org/10.1037//0096-1523.22.3.531>
- Michaels, C. F., & Oudejans, R. R. D. (1992). The optics and actions of catching fly balls: Zeroing out optical acceleration. *Ecological Psychology*, 4, 199–222. doi:https://doi.org/10.1207/s15326969eco0404_1
- Palmeri, T. J. (1997). Exemplar similarity and the development of automaticity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 23, 324–354. doi:<https://doi.org/10.1037/0278-7393.23.2.324>
- Reder, L. M., & Ritter, F. E. (1992). What determines initial feeling of knowing? Familiarity with question terms, not with the answer. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18, 435–451. doi:<https://doi.org/10.1037//0278-7393.18.3.435>
- Regan, D. (1997). Visual factors in hitting and catching. *Journal of Sports Sciences*, 15, 533–558. doi:<https://doi.org/10.1080/026404197366985>
- Rickard, T. C. (1997). Bending the power law: A CMPL theory of strategy shifts and the automatization of cognitive skills. *Journal of Experimental Psychology: General*, 126, 288–311. doi:<https://doi.org/10.1037//0096-3445.126.3.288>
- Rickard, T. C. (2004). Strategy execution in cognitive skill learning: An item-level test of candidate models. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30, 65–82. doi:<https://doi.org/10.1037/0278-7393.30.1.65>
- Rickard, T. C., & Boume, L. E., Jr. (1996). Tests of an identical elements model of basic arithmetic skills. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22, 1281–1295. doi:<https://doi.org/10.1037/0278-7393.22.5.1281>
- Schneider, W., Eschman, A., & Zuccolotto, A. (2002). E-Prime user's guide. Pittsburgh, PA: Psychology Software Tools, Inc.
- Siegler, R. S. (1988a). Individual differences in strategy choices: Good students, not-so-good students, and perfectionists. *Child Development*, 59, 833–851. doi:<https://doi.org/10.1111/1467-8624.ep8589586>
- Siegler, R. S. (1988b). Strategy choice procedures and the development of multiplication skill. *Journal of Experimental Psychology: General*, 117, 258–275. doi:<https://doi.org/10.1037/0096-3445.117.3.258>
- Strayer, D. L., & Kramer, A. F. (1990). An analysis of memory-based theories of automaticity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 16, 291–304. doi:<https://doi.org/10.1037/0278-7393.16.2.291>
- Tenison, C., & Anderson, J. R. (2016). Modeling the distinct phases of skill acquisition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 42, 749–767. doi:<https://doi.org/10.1037/xlm0000204>
- Touron, D. R. (2006). Are item-level strategy shifts abrupt and collective? Age differences in cognitive skill acquisition. *Psychonomic Bulletin & Review*, 13, 781–786. doi:<https://doi.org/10.3758/BF03193997>
- Touron, D. R., & Hertzog, C. (2004). Distinguishing age differences in knowledge, strategy use, and confidence during strategic skill acquisition. *Psychology and Aging*, 19, 452–466. doi:<https://doi.org/10.1037/0882-7974.19.3.452>

- Vandierendonck, A., Liefoghe, B., & Verbruggen, F. (2010). Task switching: Interplay of reconfiguration and interference control. *Psychological Bulletin*, *136*, 601–26. doi:<https://doi.org/10.1037/a0019791>
- Wilkins, N. J., & Rawson, K. A. (2010). Loss of cognitive skill across delays: Constraints for theories of cognitive skill acquisition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*, 1134–49. doi:<https://doi.org/10.1037/a0019998>
- Wilkins, N. J., & Rawson, K. A. (2011). Controlling retrieval during practice: Implications for memory-based theories of automaticity. *Journal of Memory and Language*, *65*, 208–221. doi:<https://doi.org/10.1016/j.jml.2011.03.006>
- Wynton, S. K. A., & Anglim, J. (2017). Abrupt strategy change underlies gradual performance change: Bayesian hierarchical models of component and aggregate strategy use. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *43*, 1630–1642. doi:<https://doi.org/10.1037/xlm0000404>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.