



# Training, retention, and transfer of data entry perceptual and motor processes over short and long retention intervals

Alice F. Healy<sup>1,2</sup> · James A. Kole<sup>3</sup> · Vivian I. Schneider<sup>1</sup> · Immanuel Barshi<sup>4</sup>

Published online: 18 June 2019  
© The Psychonomic Society, Inc. 2019

## Abstract

In two experiments, subjects trained in a standard data entry task, which involved typing numbers (e.g., 2147) using their right hands. At an initial test (20 min or 6 months after training), subjects completed the standard task, followed by a left-hand variant (typing with their left hands) that involved the same perceptual, but different motoric, processes as the standard task. At a second test (2 days or 8 months after training), subjects completed the standard task, followed by a code variant (translating letters into digits, then typing the digits with their right hands) that involved different perceptual, but the same motoric, processes as the standard task. At test, for each of the three tasks, half the trials were trained numbers (old) and half were new. Repetition priming (faster execution times to old than new numbers) was found for each task, with extended delays only slightly decreasing the magnitude of the effect. Repetition priming for the standard task reflects retention of trained numbers, for the left-hand variant reflects transfer of perceptual processes, and for the code variant reflects transfer of motoric processes. There was, thus, evidence for both specificity and generalizability of training data entry perceptual and motoric processes even over very long retention intervals.

**Keywords** Transfer of training · Retention · Specificity and generalizability of training · Perceptual and motoric processes

## Introduction

Although it is hoped that learners are flexible and adaptive so they can generalize the knowledge and skills they acquire to novel situations, research has shown that instead learning in many cases is highly specific to the training context, with little or no transfer of training when the test and training situations differ. These differences may be relatively minor, such as changes to context (e.g., Bjork & Richardson-Klavehn, 1989; Eich, 1985; Godden & Baddeley, 1975; Kole, Healy, Fierman, & Bourne, 2010), changes to stimuli (e.g., Bourne, Healy, Pauli, Parker, & Birbaumer, 2005; Rickard, Healy, &

Bourne, 1994), or changes to required responses (e.g., Healy, Wohldmann, Sutton, & Bourne, 2006; Yamaguchi & Proctor, 2009), or they may be relatively major, such as changes to task procedures (e.g., Healy, Wohldmann, Parker, & Bourne, 2005).

A number of theories have been proposed to account for such demonstrations of specificity. Thorndike's (1906) theory of identical elements (see also Rickard & Bourne, 1996, and Singley & Anderson, 1989) suggests that as the similarity between training and test elements increases, better performance during test will be observed. The elements in this theory are often the physical/perceptual features of a task, such as cues and responses (Carragher & Schliemann, 2002; Pan & Rickard, 2018). By Tulving and Thomson's (1973) theory of encoding specificity, information is encoded within a context, and the provision of contextual cues that were present at the time of training may facilitate recall at the time of test. The procedural reinstatement principle suggests that as the similarity between the procedures required during training and those required at test increases, better performance during test will be observed (Healy et al., 1992; Healy, Wohldmann, & Bourne, 2005; Lohse & Healy, 2012; for a related alternative see Proteau & Carnahan, 2001). Thus, each of these theories explicitly predicts specificity in that greater overlap between

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

<sup>1</sup> Department of Psychology and Neuroscience, University of Colorado Boulder, Boulder, CO 80309-0345, USA

<sup>2</sup> Philadelphia, PA, USA

<sup>3</sup> School of Psychological Sciences, University of Northern Colorado, Greeley, CO, USA

<sup>4</sup> NASA, Ames Research Center, Moffett Field, CA, USA

training and test, whether based on cues, responses, context, or procedures, should improve performance at test. However, collectively these theories also allow for the possibility of transfer, to the extent that there is partial overlap between training and test in cues, responses, context, or procedures. Although not explicitly specified in these theories, an implication of them is that, given partial overlap in task parameters, retention should always be greater than transfer as retention implies consistency for all four task parameters (cues, responses, context, procedures), whereas transfer implies a change in one or more task parameters.

### Retention intervals

In addition to being flexible, it is hoped that learning is durable so that individuals may retain the knowledge and skills they acquire over long periods of time, even when there is little to no use or application of the acquired knowledge and skills during the delay. For example, for long-duration space missions astronauts must train on Earth, then apply the acquired skills to tasks in space after delays as long as months or even years, and college students might be required to apply the knowledge they gained in the classroom years after graduating and securing a job. The effect of retention interval on retention has been widely documented, starting with the work of Ebbinghaus (1885/1913), who demonstrated that retention of declarative knowledge in the form of non-word trigrams declines rapidly over the first day, then asymptotes such that those items that were retrievable after the first day remained so for the next 30 days (see Murre & Dros, 2015, for a replication of Ebbinghaus' classic study). Other research has examined even longer retention intervals, from 1 year (Fendrich, Healy, & Bourne, 1993) to several years (8 years, Bahrck & Phelps, 1988; 3–16 years, Ellis, Semb, & Cole, 1998; 9 years, Squire, 1989), to several decades (Bahrck, Bahrck, & Wittlinger, 1975). Squire tested memory for television programs, and Bahrck and colleagues for face-name pairs and Spanish vocabulary. Even with these differences in retention intervals and materials, a forgetting function similar to Ebbinghaus' was found, with most forgetting occurring early within the retention interval and asymptotic levels of retention reached after 8 years (Squire, 1989). In general, forgetting may be characterized by a power function, with differences in the time required to reach asymptotic levels of retention due to differences in the degree of initial learning as well as in the forgetting rate (Wixted & Carpenter, 2007). However, for pragmatic reasons, most experimental research examining long-term retention of knowledge employs retention intervals that are orders of magnitude shorter (Squire, 1989).

Whereas many studies have examined long-term retention, fewer have examined long-term transfer or have compared directly long-term retention to transfer, and to our knowledge, no study has compared retention to transfer over

time. Fendrich et al. (1993; Experiment 3) tested memory for single-operand multiplication problems over a 1-month retention interval. Subjects in this study practiced one set of problems, and were given three types of problems at test following the 1-month delay: same as during training, switch problems for which the order of operands were switched, and entirely new problems. At test, performance was better for those problems that were practiced and tested in the same format than for those problems that involved an operand switch, demonstrating long-term retention. However, performance on the operand switch problems was better than that for unpracticed problems (even when the answers were the same), demonstrating some long-term transfer. Ellis et al. (1998) tested long-term retention (up to 16 years) of knowledge from a course on child development, and included both retention (fact recall) questions and transfer (concept application) questions. Subjects performed better on fact recall questions than on concept application questions. Thus, in accord with the implications from previous theories of specificity, both Fendrich et al. and Ellis et al. found that retention was greater than transfer over extended retention intervals; however, neither examined whether retention interval had differential effects on retention and transfer, which would require multiple retention/transfer tests over time.

A second limitation of previous research comparing retention to transfer is that only declarative knowledge was tested, whereas the present study examines procedural knowledge. By the procedural reinstatement principle (Healy, 2007; Healy & Bourne, 1995; Healy, Wohldmann, & Bourne, 2005; Healy et al., 2013), declarative knowledge demonstrates greater forgetting but more robust transfer, whereas procedural knowledge demonstrates greater retention but more limited transfer. Thus, strong retention is expected in the present investigation of procedural knowledge.

### The data entry task

In the present study, we focus on the specificity and generalizability of trained perceptual and motoric procedures. That is, when individuals are trained with one set of requirements on a perceptual-motor skill, can they retain what they have learned and generalize it to new requirements, or does changing either the perceptual requirements or the motoric requirements of the skill eliminate or reduce the benefits of training especially after long retention intervals? The skill we examine here involves data entry. In the data entry paradigm, subjects typically see four-digit numbers as numerals (e.g., 2147), read them silently, and then type them into the computer without seeing their responses and without any feedback. The sequences usually remain on the screen until the subjects hit the concluding keystroke (such as the return key). This task has been used to study skill retention and transfer and has been the test bed for multiple training principles (see Healy, Kole, Wohldmann,

Buck-Gengler, & Bourne, 2011, for a review). The task is a sequential task with both cognitive and motor requirements that can be examined separately through the different components of response time. For example, response *initiation time* (which is the time to type the first digit) largely reflects cognitive processes, such as planning, whereas response *execution time* (which is the average time to type the second, third, and fourth digits after typing the first) has been shown to reflect primarily the perceptual and motor aspects of the task (e.g., Chapman, Healy, & Kole, 2016; Fendrich, Healy, & Bourne, 1991).

## Present study

Here we review two experiments addressing the training, retention, and transfer of data entry perceptual and motor processes, one over relatively short retention intervals and the other over very long retention intervals. In both experiments, at training, subjects were given the standard version of the data entry task, in which they were shown four-digit numbers presented as numerals in a box in the middle of the computer screen and they entered them using the number row of the keyboard with their right hand. After training, subjects were given two tests. In the *change-hand* test, subjects performed the standard task along with a left-hand variant. The left-hand variant involved different motoric processes because the numbers were entered with the left hand rather than with the right hand, but the perceptual aspects of the task did not change. In the *change-stimuli* test, subjects performed the standard task along with a code variant. The code variant involved different perceptual processes because participants saw letters in the box but entered the corresponding digits (e.g., if they saw badg, they typed 2147) with their right hand, so the motoric aspects of the task did not change.

The four-digit numbers entered during the two tests either were the same as during training (old) or were seen for the first time during the test (new). Repetition priming (old faster than new) at test for the standard task reflects retention of the old numbers and specificity of training. Repetition priming for the left-hand task also requires perceptual transfer of the trained task (across changes in motor responses), and repetition priming for the code task also requires motoric transfer of the trained task (across changes in perceived stimuli). On the basis of the theories and results reviewed earlier explicitly documenting specificity of training, our working experimental hypothesis was that we would find repetition priming for the standard task, reflecting specificity, but either no repetition priming or (given an overlap in task parameters) at least less repetition priming for the two task variants (left-hand and code), both of which require transfer of training. Thus, the index of transfer that we used was the existence of repetition priming on the task variants, so that, ironically, we are, in effect, examining an index of the transfer of specificity of

learning. It should be noted that this is a novel transfer index, and that many different transfer indices have been used in past research (see, e.g., Wohldmann & Healy, 2010).

In a study also examining transfer from the right arm to the left arm, Boutin et al. (2012) found that practice led to rapid improvements in motor skill (which consisted of a sequence of extension-flexion movements of the arm), but these improvements were not transferred from the dominant to the nondominant arm, whether practice was limited or prolonged. The findings by Boutin et al. are consistent with our working experimental hypothesis. However, an earlier study by Park and Shea (2005) provided evidence of transfer from practice with the right limb to testing with the left limb in a sequence-learning task, with the amount of transfer dependent on the amount of practice (surprisingly less transfer after more extensive practice). These results provide some doubt as to the validity of the working hypothesis.

In addition, in an early data entry study by Fendrich et al. (1991; Experiment 2), subjects used two different keypad layouts, one similar to that on a calculator (with the first three digits at the bottom) and the other similar to that on a telephone (with the first three digits at the top). Half of the subjects were trained on one layout and tested on the other layout. For those subjects, two different types of old lists of 10 four-digit numbers were considered and compared to new lists of numbers, one (*old digit*) in which the sequence of digits displayed in each number in a list was the same as during training (but the sequence of required key presses for each number in the list was different) and the other (*old motor*) in which the sequence of required key presses was the same as during training (but the sequence of digits displayed was different). Fendrich et al. found an advantage for both types of old numbers relative to the new numbers for two of the response-time measures they examined (although not for the execution time measure). These results suggest generalizability of training across changes in either perceptual or motoric aspects of the task. On the basis of these prior results from Fendrich et al., we thus also considered the opposing alternative experimental hypothesis that generalizability of training would be obtained so that repetition priming would be found for the left-hand and code tasks, as well as for the standard task. This alternative hypothesis is also consistent with Schmidt's (1975) influential schema theory of motor learning, according to which practice promotes the use of schemata, which can be viewed as general rules that relate the external requirements of a task to the internal movements needed to perform the task. Generalizability could also be explained in terms of the subjects' representation of the digit sequences and/or motor patterns. In particular, compatible with generalizability would be a representation of the abstract structure of the skill (rather than its surface structure; see, e.g., Dominey, Lelekov, Ventre-Dominey, & Jeannerod, 1998), an effector-independent representation (Boutin et al., 2012;

Park & Shea, 2005; Verwey & Clegg, 2005), or a representation that does not necessitate a contextual match between the tasks given in practice and transfer (Yamaguchi, Chen, & Proctor, 2015).

In summary, the present study examines specificity and generalizability of a perceptual-motor skill, specifically examining perceptual and motoric transfer. Unlike previous studies, multiple retention/transfer tests were administered, which allows for a comparison of retention and transfer effects over time, and different delays (shorter in Experiment 1, longer in Experiment 2) were employed between training and test, which allows for an examination of how the length of the retention interval impacts retention and transfer.

## Experiment 1

### Method

**Subjects** Twenty-four undergraduate students in a course on General Psychology at the University of Colorado participated for class credit. Fourteen of the subjects were female, and ten were male. Twenty-three of them were right-handed, and one was left-handed. The left-handed subject was instructed to use his right hand for all the conditions except for the left-hand condition, where he was instructed to use his left hand, just as all the other subjects were. Three additional subjects were tested but their data were not used because they did not follow instructions. With respect to our stopping rule for data collection, 24 was our original target number of subjects based both on practical considerations and on previous research finding repetition priming effects for execution time in data entry (see, e.g., Chapman et al., 2016, who included 24 subjects in each between-subjects training condition).

**Materials** There were 2 days of training. All training was done with the standard task. Both days of training included three blocks of 100 trials, with each trial consisting of a different four-digit number. Allowable four-digit numbers were those in which no digit was repeated and the digit 0 was not used. A block was constructed by randomly selecting, without replacement, four-digit numbers from the total set of 3,024 unique numbers meeting these criteria. The same 100 trials were used in all three blocks of both training days, in a different random order. Thus, each stimulus number was presented a total of six times during the training days.

The test trials included 100 trials of the standard variant followed by 100 trials of one of the two novel variants (left-hand or code). Half of the trials in each test variant of each test day were old trials, that is, they had been presented once in each block of training during each training day. In each test, all old numbers were shown exactly once, in either the standard task or the task variant

(although for the code task the old numbers were shown as letters), with the set of 50 old numbers assigned to the standard task alternating across tests. Thus, each old stimulus trial was shown six times in training and again once during each test. The other 50 trials in each task variant of each test day were new trials that had not been shown during the training days. Every new trial was seen only once; unlike the old trials, the new trials were unique and never repeated across tests. The old and new test trials were randomly intermixed. This method allowed for a retention test of the trained stimuli in the standard variant and transfer tests of the trained stimuli in the two untrained variants. The change-hand test consisted of the standard variant followed by the left-hand variant (seeing digits and typing them with the left hand). The change-stimuli test consisted of the standard variant followed by the code variant (seeing letters and translating them into digits, with  $a = 1$ ,  $b = 2$ , etc., and typing the digits with the right hand), always in that order. [There were two other tests, both also starting with the standard task – one including a word variant involving digits presented as words (e.g., two one four seven) rather than as numerals and the other including a three-digit variant involving three-digit numbers, the first three digits of each four-digit number (e.g., 214) – but those tests are not discussed here because they were not included in Experiment 2. Also, the relationship between the standard and novel task variants with respect to perceptual and motoric processes is more complex in these cases (specifically, for the word variant the stimuli differ in length as well as in type and for the three-digit variant both the number of stimuli and the number of responses differ). The change-hand and change-stimuli tests were the second and fourth tests, respectively (the word and three-digit variants were included in the first and third tests, respectively). Furthermore, before training there was a short pretest and after the tests there was an identical short posttest, both involving all five task variants; these tests are also not discussed here because of their abbreviated nature.]

**Design** The design for training and for the tests included only within-subject independent variables. For training, the variables were training day (Day 1, Day 2) and block (Block 1, Block 2, Block 3). For the tests, the variables were task (standard, left hand) or (standard, code) and trial type (new, old). For brevity, in this report, we emphasize a single dependent variable, execution time (the average time to type the second, third, and fourth digits) for correct trials (i.e., for trials in which all four digits were typed correctly). This is a relatively pure measure of the perceptual and motor aspects of the task, which are shared among the three tasks (standard, left-hand, and code), so that interpreting transfer effects involving this measure is straightforward. We also report here, with less



emphasis, the dependent variable of initiation time (time to type the first digit) for correct trials, with the caveat that this measure reflects cognitive processes that differ from task to task (e.g., the coding task, but not the standard or left-hand task, requires cognitive processes involved in translating letters to digits), so any results involving initiation time would be difficult to interpret and little or no transfer involving this measure would be expected for the task variants.

It should be noted that the order of the two tests (change-hand and change-stimuli) was not counterbalanced. Such a lack of counterbalancing would be a problem if the two tests or the two novel task variants (left-hand and code) were being compared, but such comparisons were not made and are not of interest in the present study. Also, it should be kept in mind that the order of the two tasks within each test (the standard and the novel) was also not counterbalanced (the standard always preceded the novel), and those two tasks are compared in the present analyses. Such an order confounding needs to be considered whenever this comparison is made, but other confounding differences between the two tasks also occurred and are a necessary feature of the design. Specifically, the standard task, which is treated as a baseline, is the only task used in training, occurs more often at test (i.e., in each test), and is simpler than the novel variants. These confounding differences are not a problem, however, for testing the primary and alternative experimental hypotheses, which concern the lack or presence of repetition priming in each novel task variant. Repetition priming in the left-hand task would indicate transfer of training perceptual processes and repetition priming in the code task would indicate transfer of training motoric processes, and such evidence for transfer could not be explained by either the lack of counterbalancing or the presence of confounding differences between the standard task and the novel task variants.

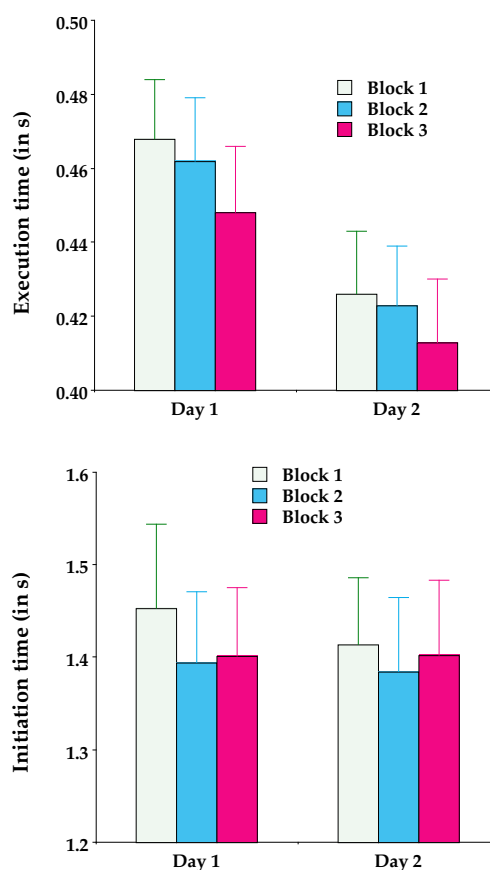
**Timing** There were three experimental sessions that were spread over 5 days, with each session lasting less than 1 h. The first experimental session included training on the standard task. For training, there were three blocks of trials, with 100 trials per block, the same 100 numbers in all six blocks. The second experimental session was 2 days after the first, and also included training as during the first experimental session (three blocks of trials, 100 trials per block, with the same 100 as during the first training session). However, 20 min after the end of training was the first test (change-hand test), which consisted of 100 trials of the standard task and then 100 trials of the left-hand task. The third experimental session was 2 days after the second; subjects completed a second test (change-stimuli test), which included 100 trials of the standard task and then 100 trials of the code task. The three experimental sessions were usually Monday, Wednesday, and Friday

with three exceptions, in which subjects delayed a session by up to 5 days.

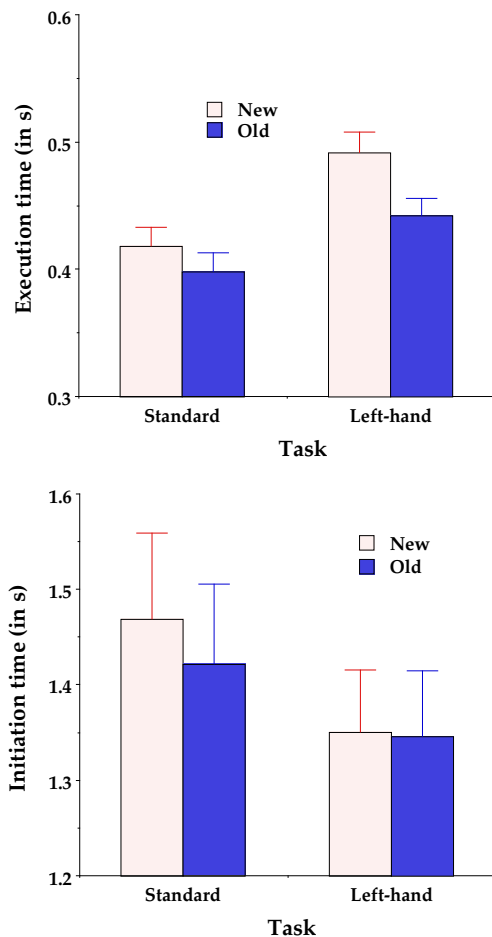
## Results and discussion

**Training** The execution time results concerning training are summarized in the top panel of Fig. 1 in terms of mean execution time for correct trials (the overall proportion correct was .921) as a function of day and block of training. There were significant main effects of both day of training,  $F(1, 23) = 54.98$ ,  $MSE = .003$ ,  $\eta_p^2 = .705$ ,  $p < .001$ , and block of training,  $F(2, 46) = 4.40$ ,  $MSE = .002$ ,  $\eta_p^2 = .161$ ,  $p = .018$ , documenting the large improvement in execution speed as a result of training.

For initiation time (see the bottom panel of Fig. 1), there were no significant main effects of either day of training,  $F(1, 23) < 1$ , or block of training,  $F(2, 46) = 3.08$ ,  $MSE = .008$ ,  $\eta_p^2 = .118$ ,  $p = .056$ , although the latter effect approached significance, documenting a modest improvement in initiation speed from the first block to the second and third blocks as a result of training.



**Fig. 1** Mean correct execution time (**top panel**) and initiation time (**bottom panel**) during training as a function of day and block in Experiment 1. Here and in the subsequent figures error bars represent between-subjects standard errors of the mean

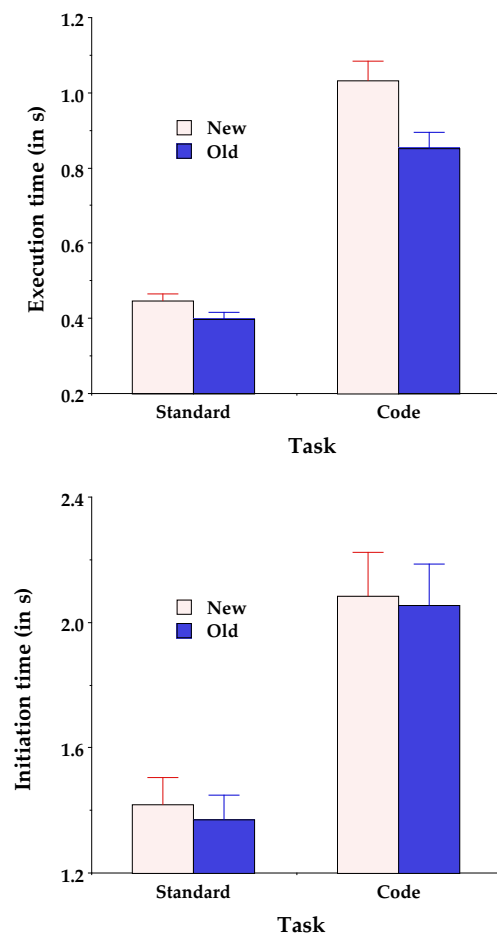


**Fig. 2** Mean correct execution time (**top panel**) and initiation time (**bottom panel**) during the change-hand test as a function of task and trial type in Experiment 1

**Change-hand test** The execution time results for the change-hand test are summarized in the top panel of Fig. 2 as a function of task and trial type. There were significant main effects of task,  $F(1, 23) = 35.52, MSE = .007, \eta_p^2 = .607, p < .001$ , and of trial type,  $F(1, 23) = 77.26, MSE = .001, \eta_p^2 = .771, p < .001$ , and a significant interaction between task and trial type,  $F(1, 23) = 8.34, MSE = .002, \eta_p^2 = .266, p = .008$ . In separate analyses of each task, there was significant repetition priming (old faster than new), but repetition priming was larger for the left-hand task,  $F(1, 23) = 91.67, MSE = .001, \eta_p^2 = .799, p < .001$ , than for the standard task,  $F(1, 23) = 7.36, MSE = .002, \eta_p^2 = .242, p = .012$ . The significant repetition priming in the standard task demonstrates the subjects’ retention of the 100 old four-digit sequences. It also demonstrates specificity of training because the subjects were only able to achieve the fast response speed they had acquired on the previously entered old sequences. The significant repetition priming for the left-hand task also demonstrates generalizability of the learned perceptual processes across changes in the motoric aspects of the task (changing the response hand from the right to the left).

For initiation time (see the bottom panel of Fig. 2), there were significant main effects of task,  $F(1, 23) = 5.20, MSE = .043, \eta_p^2 = .184, p = .032$ , and of trial type,  $F(1, 23) = 5.52, MSE = .003, \eta_p^2 = .194, p = .027$ , and a significant interaction between task and trial type,  $F(1, 23) = 4.52, MSE = .002, \eta_p^2 = .164, p = .044$ . In separate analyses of each task, repetition priming was significant for the standard task,  $F(1, 23) = 11.01, MSE = .002, \eta_p^2 = .324, p = .003$ , but not for the left-hand task,  $F(1, 23) < 1$ . The significant repetition priming in the standard task again demonstrates the subjects’ retention of the 100 old four-digit sequences and specificity of training. The lack of significant repetition priming for the left-hand task suggests no generalizability of the initial cognitive processes like planning across changes from the standard to the left-hand task.

**Change-stimuli test** The execution time results for the change-stimuli test are summarized in the top panel of Fig. 3 as a function of task and trial type. There were significant main effects of task,  $F(1, 23) = 157.19, MSE = .124, \eta_p^2 = .872, p < .001$ , and of trial type,  $F(1, 23) = 147.26, MSE = .006, \eta_p^2 = .865, p < .001$ , and a significant interaction between task and



**Fig. 3** Mean correct execution time (**top panel**) and initiation time (**bottom panel**) during the change-stimuli test as a function of task and trial type in Experiment 1

trial type,  $F(1, 23) = 83.44$ ,  $MSE = .052$ ,  $\eta_p^2 = .784$ ,  $p < .001$ . In separate analyses of each task, there was significant repetition priming, but repetition priming was larger for the code task,  $F(1, 23) = 124.14$ ,  $MSE = .010$ ,  $\eta_p^2 = .844$ ,  $p < .001$ , than for the standard task,  $F(1, 23) = 33.84$ ,  $MSE = .002$ ,  $\eta_p^2 = .595$ ,  $p < .001$ . Again, the significant repetition priming for the standard task demonstrates both retention of the trained sequences and specificity of training. The significant repetition priming for the code task also demonstrates generalizability of the learned motoric processes across changes in the perceptual aspects of the task (changing the stimuli from digits to letters).

For initiation time (see the bottom panel of Fig. 3), there was a significant main effect of task,  $F(1, 23) = 35.44$ ,  $MSE = .309$ ,  $\eta_p^2 = .606$ ,  $p < .001$ , but not of trial type,  $F(1, 23) = 2.28$ ,  $MSE = .016$ ,  $\eta_p^2 = .090$ ,  $p = .145$ , or of the interaction between task and trial type,  $F(1, 23) < 1$ . In separate analyses of each task, there was significant repetition priming for the standard task,  $F(1, 23) = 11.16$ ,  $MSE = .003$ ,  $\eta_p^2 = .327$ ,  $p = .003$ , but not for the code task,  $F(1, 23) < 1$ . Again, the significant repetition priming for the standard task demonstrates both retention of the trained sequences and specificity of training. The lack of repetition priming for the code task suggests no generalizability of the initial cognitive processes like planning across changes from the standard to the code task.

## Experiment 2

Experiment 1 showed evidence both for specificity of training, because of the significant repetition priming (typing advantage for old relative to new numbers), and for generalizability of training, because the repetition priming effect for execution time was obtained even when either the motoric aspects of the task were changed from training to test (in the left-hand task) or the perceptual aspects of the task were changed from training to test (in the code task). We wondered how long lasting these specificity and generalizability effects are. Would they survive much longer retention intervals between training and testing, such as those experienced by astronauts visiting the International Space Station for a six-month mission? Experiment 2 was aimed to address this question using essentially the same design as in Experiment 1 but with much longer delays between training and testing.

## Method

**Subjects** Twenty-six undergraduate students at the University of Colorado participated for payment of \$10 per experimental session plus a \$100 bonus at the final session for a total of \$160. They also received some NASA souvenirs as additional bonuses at some of the sessions. Those bonuses included: stickers, a keychain, a patch, a pen, and a mug. Eleven of the subjects were female, and 15 were male. As was true in

Experiment 1, all but one of the subjects were right handed. However, in Experiment 2, at the beginning of the computer program, the subjects were asked to indicate which hand was their dominant one. Then at the beginning of each condition the computer instructed the subjects to use the hand that they had declared was their dominant hand (right hand for all but one subject) for all of the conditions except for the non-dominant condition, for which they were instructed to use their non-dominant hand (left hand for all but one subject). As there was only one left-hand dominant subject that condition will continue to be called “left hand” although the one left-handed subject used his right hand for that condition and his left hand for all the other conditions. Four additional subjects started the experiment, but their data were not used because three of them attended only the first session and one of them attended only the first two sessions. With respect to our stopping rule for data collection, as for Experiment 1, 24 was our original target number of subjects. We had expected that more than four subjects would drop out of the experiment due to the large number of required sessions over a 480-day interval, which is why we tested two subjects beyond our stopping rule.

**Materials** The materials were the same as in Experiment 1, except, as noted earlier, only two tests were included, the two reported here (first the change-hand test including the left-hand task and second the change-stimuli test including the code task).

**Design** The design of Experiment 2 was the same as that of Experiment 1.

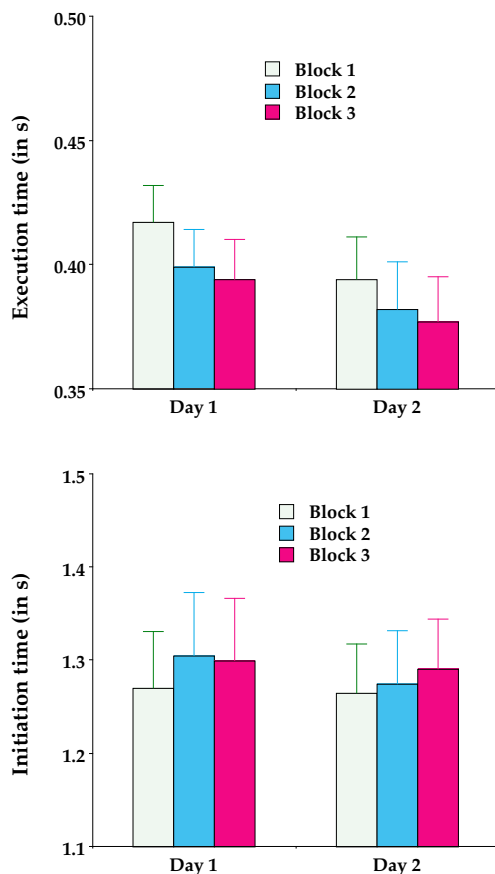
**Timing** There were six experimental sessions in Experiment 2 instead of three (in Experiment 2 the pretest and posttest occurred on separate days and the first test occurred on a separate day from training), and the sessions were spread over approximately 480 days instead of over just 5 days. Also, there were changes in the delays between experimental periods because the schedule in Experiment 2 was meant to mimic that used by astronauts on a 6-month mission to the International Space Station (astronauts are in fact participating in companion experiments). Specifically, in Experiment 1, there were 48 h between the two training days, but there were 3 months in Experiment 2 (mean = 99.56 days, standard deviation = 5.16 days). In Experiment 1, the first test occurred 20 min after the end of training on the second day, but occurred 6 months after the end of training on the second day in Experiment 2 (mean = 182.08 days, standard deviation = 4.60 days). The second test occurred 2 days after training in Experiment 1, but occurred 8 months after training in Experiment 2 (mean = 243.69 days, standard deviation = 5.64 days). The length of time between sessions in this experiment varied a bit depending on when subjects signed up. For each session, time-slots were posted

for a 2-week period. The exact time that subjects signed up for each session depended on their own schedules and time preferences. Thus, if one subject signed up early in the 2 weeks for Session 1 and late in the 2 weeks for Session 2, the time between that person’s sessions would be farther apart than that for a subject who signed up late in Session 1 and early in Session 2. Also, on a few occasions subjects were not tested within the 2-week periods.

### Results and discussion

**Training** The execution time results concerning training are summarized in the top panel of Fig. 4 in terms of mean execution time as a function of day and block of training for correct trials (the overall proportion correct was .926). Again, there were significant main effects of both day of training,  $F(1, 24) = 6.52, MSE = .006, \eta_p^2 = .214, p = .018$ , and block of training,  $F(2, 48) = 5.24, MSE = .003, \eta_p^2 = .179, p = .009$ , documenting once more the large improvement in execution speed as a result of training.

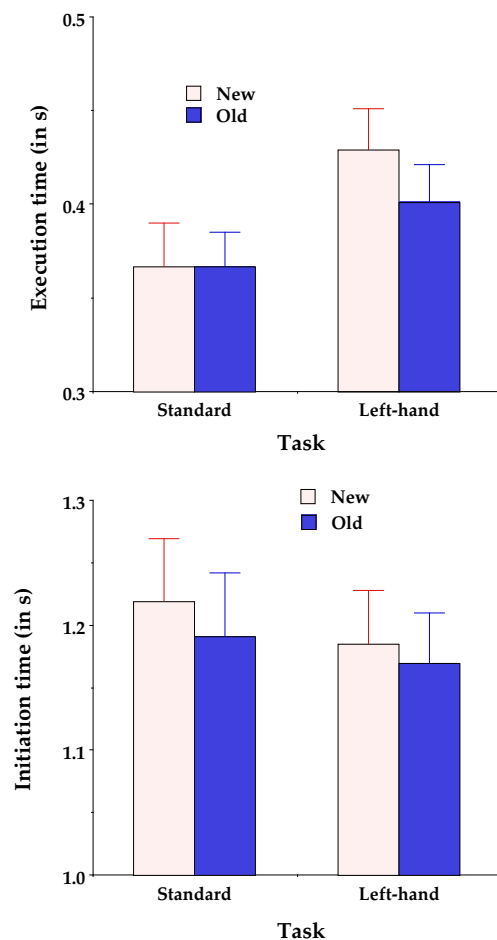
For initiation time (see the bottom panel of Fig. 4), there were no significant main effects of either day of training,  $F(1,$



**Fig. 4** Mean correct execution time (top panel) and initiation time (bottom panel) during training as a function of day and block in Experiment 2

$24) < 1$ , or block of training,  $F(2, 48) = 1.19, MSE = .009, \eta_p^2 = .047, p = .312$ , documenting no improvement in initiation speed as a result of training (in fact the numerical trend for block of training is in the opposite direction, showing a decline rather than an enhancement of performance from Block 1 to Block 3).

**Change-hand test** The execution time results for the change-hand test are summarized in the top panel of Fig. 5 as a function of task and trial type. There were significant main effects of task,  $F(1, 25) = 29.04, MSE = .006, \eta_p^2 = .537, p < .001$ , and of trial type,  $F(1, 25) = 21.88, MSE = .001, \eta_p^2 = .467, p < .001$ , and a significant interaction between task and trial type,  $F(1, 25) = 20.94, MSE = .001, \eta_p^2 = .456, p < .001$ . In separate analyses of each task, there was significant repetition priming (old faster than new) for the left-hand task,  $F(1, 25) = 45.41, MSE = .001, \eta_p^2 = .645, p < .001$ , but not for the standard task,  $F(1, 25) < 1$ . Perhaps the elimination of the repetition priming effect for the standard task is due to performance that approaches the floor in that case. Although there was, thus, no evidence from the standard task for retention of the trained



**Fig. 5** Mean correct execution time (top panel) and initiation time (bottom panel) during the change-hand test as a function of task and trial type in Experiment 2

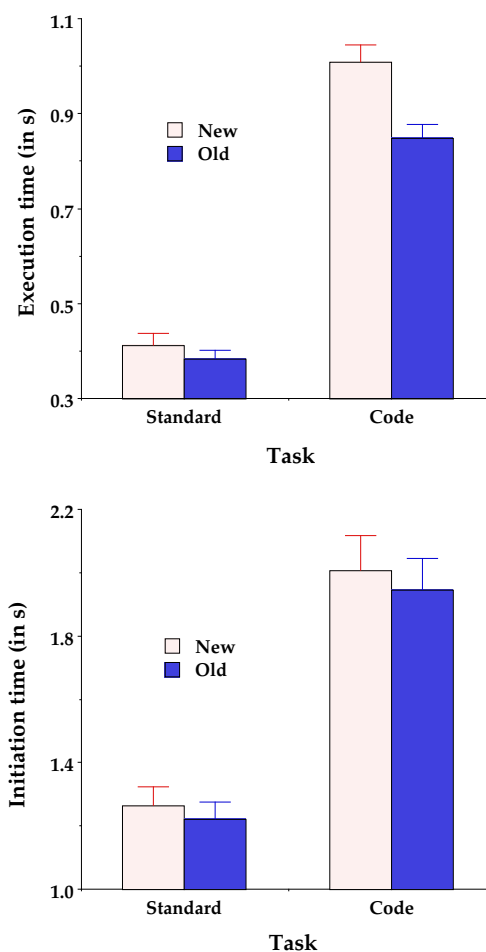


sequences or of specificity of training, the significant repetition priming for the left-hand task does demonstrate both retention and specificity of training and also demonstrates generalizability of the learned perceptual processes across changes in the motoric aspects of the task.

For initiation time (see the bottom panel of Fig. 5), there was not a significant main effect of task,  $F(1, 25) = 1.74$ ,  $MSE = .012$ ,  $\eta_p^2 = .065$ ,  $p = .200$ , although there was a significant main effect of trial type,  $F(1, 25) = 7.80$ ,  $MSE = .002$ ,  $\eta_p^2 = .238$ ,  $p = .010$ , but no significant interaction between task and trial type,  $F(1, 25) < 1$ . In separate analyses of each task, there was significant repetition priming (old faster than new) for the standard task,  $F(1, 25) = 10.06$ ,  $MSE = .001$ ,  $\eta_p^2 = .287$ ,  $p = .004$ , but not for the left-hand task,  $F(1, 25) = 1.29$ ,  $MSE = .002$ ,  $\eta_p^2 = .049$ ,  $p = .267$ . There was, thus, for this measure, unlike for execution time, evidence with the standard task for retention of the trained sequences and for specificity of training. In contrast, the lack of repetition priming with the left-hand task for this measure, unlike for execution time, suggests no generalizability of the initial cognitive processes like planning across changes from the standard to the left-hand task.

**Change-stimuli test** The execution time results for the change-stimuli test are summarized in the top panel of Fig. 6 as a function of task and trial type. There were again significant main effects of task,  $F(1, 25) = 260.55$ ,  $MSE = .085$ ,  $\eta_p^2 = .912$ ,  $p < .001$ , and of trial type,  $F(1, 25) = 102.28$ ,  $MSE = .007$ ,  $\eta_p^2 = .804$ ,  $p < .001$ , and a significant interaction between task and trial type,  $F(1, 25) = 51.08$ ,  $MSE = .007$ ,  $\eta_p^2 = .671$ ,  $p < .001$ . In separate analyses of each task, there was significant repetition priming for both tasks, but repetition priming was larger for the code task,  $F(1, 25) = 84.76$ ,  $MSE = .012$ ,  $\eta_p^2 = .772$ ,  $p < .001$ , than for the standard task,  $F(1, 25) = 22.97$ ,  $MSE = .002$ ,  $\eta_p^2 = .479$ ,  $p < .001$ . Again, the significant repetition priming for the standard task demonstrates both retention of the trained sequences and specificity of training. The significant repetition priming for the code task also demonstrates generalizability of the learned motoric processes across changes in the perceptual aspects of the task.

For initiation time (see the bottom panel of Fig. 6), there was a significant main effect of task,  $F(1, 25) = 74.96$ ,  $MSE = .187$ ,  $\eta_p^2 = .750$ ,  $p < .001$ , but not of trial type,  $F(1, 25) = 2.57$ ,  $MSE = .027$ ,  $\eta_p^2 = .093$ ,  $p = .122$ , nor of the interaction between task and trial type,  $F(1, 25) < 1$ . In separate analyses of each task, there was significant repetition priming for the standard task,  $F(1, 25) = 9.25$ ,  $MSE = .002$ ,  $\eta_p^2 = .270$ ,  $p = .006$ , but not for the code task,  $F(1, 25) < 1$ . Again, the significant repetition priming for the standard task demonstrates both retention of the trained sequences and specificity of training, whereas the lack of significant repetition priming for the code task suggests no generalizability of the initial cognitive processes like planning across changes from the standard to the code task.



**Fig. 6** Mean correct execution time (top panel) and initiation time (bottom panel) during the change-stimuli test as a function of task and trial type in Experiment 2

### Comparison of effect sizes from Experiments 1 and 2

An examination of the effect sizes for the execution time repetition priming effects under the standard task and the task variants (left hand, code) for Experiments 1 and 2 allows for a comparison of retention (repetition priming in the standard task) to transfer (repetition priming in the task variants) as well as an assessment of the effects of time interval on specificity and transfer. For the change-hand test, interestingly, in Experiment 1, there was a greater repetition priming effect for the left-hand task ( $\eta_p^2 = .799$ , 90% CI = .641, .859) than for the standard task ( $\eta_p^2 = .242$ , 90% CI = .032, .443); this pattern was replicated in Experiment 2 with somewhat reduced effect sizes (left-hand task,  $\eta_p^2 = .645$ , 90% CI = .420, .748; standard task,  $\eta_p^2 = .000$ , 90% CI [NA]) relative to Experiment 1. The fact that the effect size for transfer was greater than that for retention is surprising, as previously reviewed theories of specificity would suggest that retention should always be greater than transfer. In the present case, the perceptual and motor requirements are identical for the standard task at test as at training, but these requirements only

partially overlap for the left-hand task (same perceptual, but different motor requirements). Furthermore, the extended time interval in Experiment 2 resulted in only small decrements in effect size for the repetition priming effects, showing similarly durable retention and transfer.

Examining the execution time effect sizes for the standard and code tasks in the change-stimuli test of Experiment 1, there was greater repetition priming for the code task ( $\eta_p^2 = .844$ ; 90% CI = .715, .890) than for the standard task ( $\eta_p^2 = .595$ , 90% CI = .342, .716), with a similar pattern found for Experiment 2 (code task,  $\eta_p^2 = .772$ , 90% CI = .607, .839; standard task,  $\eta_p^2 = .479$ , 90% CI = .221, .627). Once again, the extended time interval in Experiment 2 only modestly decreased the effect sizes for both repetition priming effects.

One possibility that could explain the finding of greater transfer than retention for execution time is that execution time was longer for the left-hand and code tasks relative to the standard task, and the longer times might have allowed more room to observe differences between old and new trials. Alternatively, the greater repetition priming effects for the novel than for the standard variants could be due in part to higher variability of performance on the code task, rather than to longer average times, although higher variability on the left-hand task was not evident (see the *MSE* values).

In any event, greater retention than transfer was found for initiation time, as expected given the different initial planning processes between the standard task and the task variants. Furthermore, the pattern of effect sizes for the execution time repetition priming effects under the standard task and the task variants is a non-crossover interaction, and interpretation of such interactions is difficult because they are tied to the scale of measurement and are, thus, removable (see, e.g., Loftus, 1978; Wagenmakers, Krypotos, Criss, & Iverson, 2012). The crucial finding, though, with respect to testing the working and alternative experimental hypotheses of the present study is not that the repetition priming effect under the task variants is larger than that under the standard variant but rather that the repetition priming effect under both task variants is substantial and, thus, reflects transfer of both perceptual and motoric processes, thereby clearly supporting the alternative experimental hypothesis.

## General discussion

As found in many earlier studies with different paradigms (e.g., Healy, Wohldmann, Parker, & Bourne, 2005; Healy et al., 2006; see Johnson & Proctor, 2017, for a recent review), there was evidence in the present two experiments of a significant degree of specificity of training because the repetition priming effect found here for the standard task (for both the execution time and initiation time measures) shows that subjects were faster at responding to numbers that had occurred

six times during training than they were at responding to new numbers not shown earlier. There was no evidence for generalizability of training for initiation time, which reflects primarily initial cognitive processes involved in planning, and those processes should differ across the task variants. However, contrary to our working hypothesis based on previous findings and models of specificity (e.g., procedural reinstatement; Healy et al., 1992), but in agreement with the opposing alternative hypothesis based on the findings of Fendrich et al. (1991) and consistent with Schmidt's (1975) schema theory, there was strong evidence for generalizability of training in the present study because the repetition priming effect for execution time was found even when the training and testing tasks differed in the motoric aspects of the responses (in the left-hand task) and when the training and testing tasks differed in the perceptual aspects of the stimuli (in the code task). In fact, the repetition priming effects for execution time were consistently larger for the novel task variants (left-hand and code) than for the standard task (although the caveat mentioned earlier about removable, non-crossover interactions needs to be kept in mind). This difference in the magnitude of the repetition priming effects might be attributed in part either to the longer response times with the novel variants or to higher variability in response times with the code task than with the standard task. In any case, it seems difficult to reconcile this finding with the earlier models of specificity, including the identical elements models (e.g., Thorndike, 1906) and the procedural reinstatement principle (e.g., Healy et al., 1992), because, although not explicitly specified in these theories, they seem to imply that, given partial overlap in task parameters, retaining a skill required for a particular task should be easier than transferring that skill to a different task. The present results, thus, provide food for thought as to how to incorporate findings of generalizability into models of specificity. Perhaps a partial solution to this dilemma lies in the fact that subparts of the digit sequences or motor patterns (e.g., pairs of digits or pairs of key presses) might be shared across the sequences, so specificity might apply at a level lower than that of an entire sequence. However, it is not clear how this subpart explanation could account for the transfer found from the right to the left hand or from the digits to the letters. A more promising possibility concerns the type of representation of the digit sequences and/or motor patterns. As mentioned in the *Introduction*, the generalizability found in the present study could be attributed to a representation of the abstract structure of the skill (as opposed to the surface structure; Dominey et al., 1998), or a representation that is effector independent (Boutin et al., 2012; Park & Shea, 2005; Verwey & Clegg, 2005) and does not require a contextual match between the practice and transfer tasks (Yamaguchi et al., 2015).

In any event, there was clearly transfer of perceptual processes in the left-hand task and of motoric processes in the code task. In previous work by Healy, Schneider, and Barshi

(2015), involving a navigation task and different measures of transfer, either specificity or generalizability was found, but not both, for a given condition in each of their six experiments. Nevertheless, in the present study there was evidence for both specificity and generalizability of training for both perceptual and motoric processes of data entry even over very long retention intervals.

The finding in the present study of transfer of training across changes in both perceptual and motoric aspects of the task is perhaps most similar to findings in the early data entry experiment by Fendrich et al. (1991) (summarized in the present *Introduction*), in which for one group of subjects there was a change in keypad layouts between training and testing. For those subjects in that study some lists of numbers at test had old digits and others had old motor sequences, and there was an advantage at test for both types of old lists relative to new lists, implying generalizability of training across both perceptual and motoric dimensions of the task.

The finding of transfer of perceptual processes across changes in motoric processes between using the right hand or the left hand is also consistent with the finding by Healy et al. (2015) that the trained navigation processes they examined showed transfer across changes in the motoric requirements between using a mouse or using a keypad to make a response. However, unlike the present study, which also found transfer of motoric processes across changes in perceptual processes between showing the stimuli as digits or as letters, Healy et al. (2015) did not find any consistent evidence of transfer of motoric processes across changes in perceptual processes – for example, between changes in either the presentation mode of the navigation instructions or in the display type of the navigation space. In the cases where transfer was found in the navigation study, it was usually asymmetric, with transfer evident to a test condition that shared procedures with the training condition but not to a test condition that had some unique procedures.

A novel aspect of the present findings is the very long retention intervals involved (6 and 8 months following training in Experiment 2), in contrast to the 1-week interval used by Fendrich et al. (1991) and the 5-min to 1-week intervals used by Healy et al. (2015). It seems truly amazing that subjects could remember the 100 old four-digit sequences across such long retention intervals, with little difference in findings between the very long retention intervals of Experiment 2 and the relatively short retention intervals of Experiment 1. It is even more impressive that subjects could demonstrate such durable memory even when they entered the sequences with a hand different from that used at training or even when they saw the sequences in a format different from that used at training. These findings of remarkable memory for four-digit numbers over retention intervals of 6 and 8 months surely stand in stark contrast to findings from

the distractor paradigm (e.g., Healy, 1974; Peterson & Peterson, 1959) that after a delay of about 7 s filled with a distractor task subjects can remember only about two of four letters they had been shown.

The fact that the sequences learned in the present study were four-digit numbers (i.e., sequences of four digits), rather than lists of 10 four-digit numbers (i.e., sequences of 40 digits), also makes the present results impressive relative to those of the Fendrich et al. (1991) study, because the four-digit numbers are highly confusable and there were 100 of them learned during training, in contrast to the 20 lists of numbers learned in the study by Fendrich et al. Furthermore, learning was purely incidental in the present study, in contrast to the study by Fendrich et al., where subjects made a 1–6 recognition response after entering each list of numbers on their retention test, yielding a component of explicit memory in their study.

Beyond the theoretical implications concerning models of specificity, the practical implications of the present results concerning training are encouraging especially when there will be long delays between training and testing, as in the case of the astronauts on a 6-month mission to the International Space Station. Despite the long retention intervals and changes in the perceptual or motoric aspects of the trained tasks, learners could be expected to benefit from the training they received earlier even when the learning is incidental and even when what is learned is highly confusable. Furthermore, the significant transfer found for execution time across changes in either the perceptual or motoric aspects of the task has implications for the issue of how much fidelity to the target task is necessary for simulators and other training devices to be effective. The present results suggest that changes in perceptual or motoric aspects of the task will not detract from the usefulness of such training devices, so they can be highly effective without perfect fidelity to the target task.

**Author Note** This research was supported in part by National Aeronautics and Space Administration Grant NNX14AB75A to the University of Colorado Boulder. Preparation of this paper occurred while Alice Healy was a visiting scholar in the laboratory of Prof. Michael Kahana at the University of Pennsylvania. We are grateful to Steven Pan and Jason Arndt for very helpful comments on an earlier version of this paper.

**Open Practice Statement** The data and materials for both experiments are available from the authors by request, and neither of the experiments was preregistered.

## References

- Bahrick, H. P., Bahrick, P. O., & Wittlinger, R. P. (1975). Fifty years of memory for names and faces: A cross-sectional approach. *Journal of Experimental Psychology: General*, *104*, 54–75. <https://doi.org/10.1037/0096-3445.104.1.54>

- Bahrnick, H. P., & Phelps, E. (1988). The maintenance of marginal knowledge. In U. Neisser & E. Winograd (Eds.), *Remembering reconsidered: Ecological and traditional approaches to the study of memory* (pp. 178–192). New York: Cambridge University Press. <https://doi.org/10.1017/CBO9780511664014.008>
- Bjork, R. A., & Richardson-Klavehn, A. (1989). On the puzzling relationship between environment context and human memory. In C. Izawa (Ed.), *Current issues in cognitive processes: The Tulane Flowerree Symposium on Cognition* (pp. 313–344). Hillsdale, NJ: Erlbaum.
- Bourne, L. E., Jr., Healy, A. F., Pauli, P., Parker, J. T., & Birbaumer, N. (2005). The influence of stimulus array on training of a speeded response. *American Journal of Psychology*, *118*, 385–411.
- Boutin, A., Badets, A., Salesse, R. N., Fries, U., Panzer, S., & Blandin, Y. (2012). Practice makes transfer of motor skills imperfect. *Psychological Research*, *76*, 611–625. <https://doi.org/10.1007/s00426-011-0355-2>
- Carraher, D., & Schliemann, A. D. (2002). The transfer dilemma. *Journal of the Learning Sciences*, *11*, 1–24. [https://doi.org/10.1207/S15327809JLS1101\\_1](https://doi.org/10.1207/S15327809JLS1101_1)
- Chapman, M. J., Healy, A. F., & Kole, J. A. (2016). Memory load as a cognitive antidote to performance decrements in data entry. *Memory*, *24*, 1182–1196. <https://doi.org/10.1080/09658211.2015.1086380>
- Dominey, P. F., Lelekov, T., Ventre-Dominey, J., & Jeannerod, M. (1998). Dissociable processes for learning the surface structure and abstract structure of sensorimotor sequences. *Journal of Cognitive Neuroscience*, *10*, 734–751. <https://doi.org/10.1162/089892998563130>
- Ebbinghaus, H. (1913). *Memory: A contribution to experimental psychology*. New York: Teachers College, Columbia University. (Original work published 1885). <https://doi.org/10.1037/10011-000>
- Eich, E. (1985). Context, memory, and integrated item/context imagery. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *11*, 764–770. <https://doi.org/10.1037/0278-7393.11.1.4.764>
- Ellis, J. A., Semb, G. B., & Cole, B. (1998). Very long-term memory for information taught in school. *Contemporary Educational Psychology*, *23*, 419–433. <https://doi.org/10.1006/ceps.1997.0976>
- Fendrich, D. W., Healy, A. F., & Bourne, L. E., Jr. (1991). Long-term repetition effects for motoric and perceptual procedures. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*, 137–151. <https://doi.org/10.1037/0278-7393.17.1.137>
- Fendrich, D. W., Healy, A. F., & Bourne, L. E., Jr. (1993). Mental arithmetic: Training and retention of multiplication skill. In C. Izawa (Ed.), *Cognitive psychology applied* (pp. 111–133). Hillsdale, New Jersey: Erlbaum.
- Godden, D. R., & Baddeley, A. D. (1975). Context-dependent memory in two natural environments: On land and underwater. *British Journal of Psychology*, *66*, 325–331. <https://doi.org/10.1111/j.2044-8295.1975.tb01468.x>
- Healy, A. F. (1974). Separating item from order information in short-term memory. *Journal of Verbal Learning and Verbal Behavior*, *13*, 644–655. [https://doi.org/10.1016/S0022-5371\(74\)80052-6](https://doi.org/10.1016/S0022-5371(74)80052-6)
- Healy, A. F. (2007). Transfer: Specificity and generality. In H. L. Roediger, III, Y. Dudai, & S. M. Fitzpatrick (Eds.), *Science of memory: Concepts* (pp. 271–275). New York: Oxford University Press.
- Healy, A. F., & Bourne, L. E., Jr. (Eds.). (1995). *Learning and memory of knowledge and skills: Durability and specificity*. Thousand Oaks, CA: Sage. <https://doi.org/10.4135/9781483326887>
- Healy, A. F., Fendrich, D. W., Crutcher, R. J., Wittman, W. T., Gesi, A. T., Ericsson, K. A., & Bourne, L. E., Jr. (1992). The long-term retention of skills. In A. F. Healy, S. M. Kosslyn, & R. M. Shiffrin (Eds.), *From learning processes to cognitive processes: Essays in honor of William K. Estes, Volume 2* (pp. 87–118). Hillsdale, NJ: Erlbaum.
- Healy, A. F., Kole, J. A., Wohldmann, E. L., Buck-Gengler, C. J., & Bourne, L. E., Jr. (2011). Data entry: A window to principles of training. In A. S. Benjamin (Ed.), *Successful remembering and successful forgetting: A Festschrift in honor of Robert A. Bjork* (pp. 277–296). New York: Psychology Press.
- Healy, A. F., Schneider, V. I., & Barshi, I. (2015). Specificity and transfer in learning how to follow navigation instructions. In J. G. W. Raaijmakers, A. H. Criss, R. L. Goldstone, R. M. Nosofsky, & M. Steyvers (Eds.), *Cognitive modeling in perception and memory: A Festschrift for Richard M. Shiffrin* (pp. 259–273). New York: Psychology Press.
- Healy, A. F., Wohldmann, E. L., & Bourne, L. E., Jr. (2005). The procedural reinstatement principle: Studies on training, retention, and transfer. In A. F. Healy (Ed.), *Experimental cognitive psychology and its applications* (pp. 59–71). Washington, DC: American Psychological Association. <https://doi.org/10.1037/10895-005>
- Healy, A. F., Wohldmann, E. L., Kole, J. A., Schneider, V. I., Shea, K. M., & Bourne, L. E., Jr. (2013). Training for efficient, durable, and flexible performance in the military. In W. Arthur, Jr., E. A. Day, W. Bennett, Jr., & A. M. Portrey (Eds.), *Individual and team skill decay: The science and implications for practice* (pp. 176–204). New York: Routledge.
- Healy, A. F., Wohldmann, E. L., Parker, J. T., & Bourne, L. E., Jr. (2005). Skill training, retention, and transfer: The effects of a concurrent secondary task. *Memory & Cognition*, *33*, 1457–1471. <https://doi.org/10.3758/BF03193378>
- Healy, A. F., Wohldmann, E. L., Sutton, E. M., & Bourne, L. E., Jr. (2006). Specificity effects in training and transfer of speeded responses. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*, 534–546. <https://doi.org/10.1037/0278-7393.32.3.534>
- Johnson, A., & Proctor, R. W. (2017). *Skill acquisition and training: Achieving expertise in simple and complex tasks*. New York, NY: Routledge.
- Kole, J. A., Healy, A. F., Fierman, D. M., & Bourne, L. E., Jr. (2010). Contextual memory and skill transfer in category search. *Memory & Cognition*, *38*, 67–82. <https://doi.org/10.3758/MC.38.1.67>
- Loftus, G. R. (1978). On interpretation of interactions. *Memory & Cognition*, *6*, 312–319. <https://doi.org/10.3758/BF03197461>
- Lohse, K. R., & Healy, A. F. (2012). Exploring the contributions of declarative and procedural information to training: A test of the procedural reinstatement principle. *Journal of Applied Research in Memory and Cognition*, *1*, 65–72. <https://doi.org/10.1016/j.jarmac.2012.02.002>
- Murre, J. M. J., & Dros, J. (2015). Replication and analysis of Ebbinghaus' forgetting curve. *PLoS One*, *10* (7): e0120644. <https://doi.org/10.1371/journal.pone.0120644>
- Pan, S. C., & Rickard, T. C. (2018). Transfer of test-enhanced learning: Meta-analytic review and synthesis. *Psychological Bulletin*, *144*, 710–756. <https://doi.org/10.1037/bul0000151>
- Park, J.-H., & Shea, C. H. (2005). Sequence learning: Response structure and effector transfer. *Quarterly Journal of Experimental Psychology*, *58A*, 387–419. <https://doi.org/10.1080/02724980343000918>
- Peterson, L. R., & Peterson, M. J. (1959). Short-term retention of individual verbal items. *Journal of Experimental Psychology*, *58*, 193–198. <https://doi.org/10.1037/h0049234>
- Proteau, L., & Carnahan, H. (2001). What causes specificity of practice in a manual aiming movement: Vision dominance or transformation errors? *Journal of Motor Behavior*, *33*, 226–234. <https://doi.org/10.1080/00222890109601908>
- Rickard, T. C., & Bourne, L. E., Jr. (1996). Some tests of an identical elements model of basic arithmetic skills. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*, 1281–1295. <https://doi.org/10.1037/0278-7393.22.5.1281>
- Rickard, T. C., Healy, A. F., & Bourne, L. E., Jr. (1994). On the cognitive structure of basic arithmetic skills: Operation, order, and symbol transfer effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*, 1139–1153. <https://doi.org/10.1037/0278-7393.20.5.1139>



- Schmidt, R. A. (1975). A schema theory of discrete motor skill learning. *Psychological Review*, *82*, 225–260. <https://doi.org/10.1037/h0076770>
- Singley, M. K., & Anderson, J. R. (1989). *The transfer of cognitive skill*. Cambridge, MA: Harvard University Press.
- Squire, L. R. (1989). On the course of forgetting in very long-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*, 241–245. <https://doi.org/10.1037/0278-7393.15.2.241>
- Thorndike, E. L. (1906). *The principles of teaching based on psychology*. New York: A. G. Seiler. <https://doi.org/10.1037/11487-000>
- Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review*, *80*, 352–373. <https://doi.org/10.1037/h0020071>
- Verwey, W. B., & Clegg, B. A. (2005). Effector dependent sequence learning in the serial RT task. *Psychological Research*, *69*, 242–251. <https://doi.org/10.1007/s00426-004-0181-x>
- Wagenmakers, E.-J., Krypotos, A.-M., Criss, A. H., & Iverson, G. (2012). On the interpretation of removable interactions: A survey of the field 33 years after Loftus. *Memory & Cognition*, *40*, 145–160. <https://doi.org/10.3758/s13421-011-0158-0>
- Wixted, J. T., & Carpenter, S. K. (2007). The Wickelgren Power Law and the Ebbinghaus Savings Function. *Psychological Science*, *18*, 133–134. <https://doi.org/10.1111/j.1467-9280.2007.01862.x>
- Wohldmann, E. L., & Healy, A. F. (2010). Exploring specificity of speeded aiming movements: Examining different measures of transfer. *Memory & Cognition*, *38*, 344–355. <https://doi.org/10.3758/MC.38.3.344>
- Yamaguchi, M., Chen, J., & Proctor, R. W. (2015). Transfer of learning in choice reactions: The roles of stimulus type, response mode, and set-level compatibility. *Memory & Cognition*, *43*, 825–836. <https://doi.org/10.3758/s13421-015-0518-2>
- Yamaguchi, M., & Proctor, R. W. (2009). Transfer of learning in choice-reactions: Contributions of specific and general components of manual responses. *Acta Psychologica*, *130*, 1–10. <https://doi.org/10.1016/j.actpsy.2008.09.008>

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