



Testing enhances motor practice

Tobias Tempel¹ · Christian Frings²

Published online: 24 April 2019
© The Psychonomic Society, Inc. 2019

Abstract

We investigated how retrieval of a set of newly learned motor sequences influences subsequent learning of another set of motor sequences. In four experiments, retrieval reduced an acceleration of movement execution over subsequent study trials. This relative slowing-down was associated with better recall performance in a final memory test. Explicit retrievability of motor sequences benefited from longer study-trial response times (RTs), suggesting that retrieval caused more attentive encoding. The use of motor sequences requiring overt action during encoding allowed for this demonstration of a twofold forward effect of testing on encoding quality and on recall. Experiment 1 adopted a paradigm used in previous studies with verbal materials. Experiment 2 changed the test format to be less susceptible to interference. Experiments 3 and 4 additionally switched from a between-participants design to a within-participants design. These modifications did not affect the occurrence of the twofold forward effect of testing but enabled detecting a correlation between recall and study-trial performance that had been precluded by the strongly interference-dependent test format of the original paradigm. Our findings demonstrate an immediate learning benefit of testing. It enhances encoding in subsequent study trials.

Keywords Body movement · Retrieval · Testing effect · Interim-test effect · Motor sequence learning

Introduction

The acquisition of motor sequences is only seldom examined from a memory perspective; instead, a learning perspective predominates. Many of the variables that have been examined in memory research (which mostly relies on words or images as item materials) that pertain to encoding or retrieval have never been considered as affecting motor memory as well. With regard to encoding, motor-sequence-learning research mostly focuses on incidental processes. Indeed, numerous studies included conditions of intentional learning of motor sequences as well but many of the influences on intentional encoding that were identified in memory experiments have been neglected in motor action so far. Moreover, the kind of intention in question differs typically. Studies using the serial-reaction-time task (Nissen & Bullemer, 1987), for example, routinely compare incidental and intentional conditions, but the respective intention there is to

watch whether certain regularities might be detected, not to memorize a defined set of distinct items with the goal of recalling them in a later memory test (e.g., Unsworth & Engle, 2005). Furthermore, direct memory tests, asking for retrieval of practiced motor sequences, usually only serve as a supplement to retention and transfer tests assessing motor and/or sequence learning by requiring the execution of actions that match previously practiced actions in response to the same or different stimuli as during practice (e.g., Hoffmann & Koch, 1997). Effects of memory testing on subsequent accessibility in motor memory have only recently received attention in a relatively low number of studies.

By adapting paradigms from memory research to the use of motor sequences as item material, we found that retrieval can shape motor memory in opposing ways by demonstrating retrieval-induced forgetting as well as test-potentiated learning (Tempel, Aslan, & Frings, 2016; Tempel & Frings, 2013, 2014a, 2014b, 2015, 2016a, 2017; Tempel, Loran, & Frings, 2015; Tempel & Kubik, 2017). Thus, retrieval effects extend to motor memory. In addition, we demonstrated costs and benefits of directed forgetting of motor sequences (Tempel & Frings, 2016b). In that study, directed forgetting of a first item list improved encoding of a second list of motor sequences. This was the first demonstration of behavioral evidence directly from study trials in support of theoretical assumptions that directed forgetting not only reduces

✉ Tobias Tempel
tobias.tempel@ph-ludwigsburg.de

¹ Ludwigsburg University of Education, Reuteallee 46,
71634 Ludwigsburg, Germany

² Department of Cognitive Psychology, University of Trier,
Trier, Germany

interference in a memory test, but already facilitates learning (Pastötter & Bäuml, 2010; Sahakyan & Delaney, 2003). We were able to show this effect because motor sequences necessitate movement execution. This execution involved key presses that were recorded. The recorded data on movement execution could then be analyzed. Thus, motor sequences enable more direct insights into encoding processes than more common materials in memory studies. Participants in a typical memory experiment learn items in a passive manner without overt action, by reading words, listening to sounds, or viewing images, whereas encoding motor sequences requires action. The executed responses can be measured and analyzed as an index of encoding quality. Therefore, adapting memory paradigms for the use of motor sequences as items not only enables novel insights on processes contributing to the acquisition of motor sequences but also produces new measures of encoding processes that are subject of theories in memory research in general.

In the present study, we examined retrieval effects on the encoding of motor sequences. Hence, we combined our previous approaches of investigating how retrieval shapes motor memory with analyzing a behavioral index of encoding quality that is offered by sequence execution during study trials. Szpunar, McDermott, and Roediger (2008) demonstrated that a group of participants learning several word lists each separated by retrieval of a just-studied list recalled more items of the last study list in the row compared to a group learning the lists separated by restudy of a just-studied list. This forward effect of testing (for reviews, see: Pastötter & Bäuml, 2014; Chan, Meissner, & Davis, 2018) generalizes to a variety of different materials, such as texts (Wissman, Rawson, & Pyc, 2011), faces and names (Weinstein, McDermott, & Szpunar, 2011), images (Pastötter, Weber, & Bäuml, 2013), and videos (Szpunar, Khan, & Schacter, 2013). The forward effect of testing has been assumed to result from reduced interference as well as from better learning as a consequence of retrieval. Szpunar et al. (2008) originally posited that retrieval causes mental context changes that segregate individual lists. This segregation accounts for reduced interference at the later memory test. Pastötter, Schicker, Niedernhuber, and Bäuml (2011) suggested that retrieval additionally causes a reset of encoding, that is, it abolishes memory load and inattentive encoding. The better encoding of items learned subsequent to retrieval of a previous list then contributes to the forward effect of testing in a later test. Using EEG, Pastötter et al. (2011) observed that the typical increase in alpha oscillation during the encoding of a sequence of item lists vanished when retrieval separated the individual lists. The strength of alpha oscillations was considered an index of memory load and correspondingly correlated with memory performance in an equivalent manner as in previous studies (e.g., Bäuml, Hanslmayr, Pastötter, & Klimesch, 2008).

Similar processes are discussed in studies with the list method of directed forgetting. The list method compares two groups of participants, both receiving two study lists and a recall test. After studying the first list, one group is instructed to forget the thus-far presented items, whereas the other group is informed that they had just learned the first half of items and would continue with the second half now. In a final test, participants in both groups then are asked to recall the items from both lists. Typically, a benefit of directed forgetting emerges as more items of the second list are recalled in the forget group than in the remember group. In addition, usually a cost effect emerges, that is, the forget group recalls significantly fewer items of the first list. Reduced interference from the to-be-forgotten list (Bjork, 1989) as well as better encoding of items of the second list contribute to the benefit effect. In a recent paper, Pastötter, Tempel, and Bäuml (2017) review behavioral and neurocognitive evidence on encoding benefits by directed forgetting. For example, Hanslmayr (2012) demonstrated that alpha amplitude during item encoding increases from L1 to L2 in the remember condition, but not in the forget condition. Thus, the forgetting instruction eliminated the typical increase of alpha oscillations during encoding of a sequence of lists, as did retrieval of a just-studied list in the study by Pastötter et al. (2011). This similarity between the forward effect of testing and the benefit effect of directed forgetting suggested to us that we might observe further similarities on a behavioral level when using motor sequences as study material. At present, there is much less evidence on encoding processes that contribute to the forward effect of testing as compared to evidence on encoding processes that contribute to the benefit effect of directed forgetting. There are findings that participants choose to allocate more study time to a next study list after receiving a test for a preceding list when they are free to do so (Yang, Potts, & Shanks, 2018). However, from this observation it remains unclear whether encoding would also benefit without that choice of study-time allocation.

Here, we examined whether retrieval of a just-studied set of sequential finger movements (SFMs) would subsequently further intentional learning of another set of SFMs, scrutinizing response times (RTs) at executing SFMs in study trials as an index of encoding quality. Given the electrophysiological evidence for an influence of retrieval on subsequent encoding, we assumed the behavioral index of motor-sequence execution used in the present experiments to reflect a forward effect of testing on subsequent encoding.

Overview

The investigation of retrieval effects necessitates learning to be intentional, with instructions announcing memory tests, in order to equate expectations in the contrasted experimental conditions with and without tests separating study lists.

Thus, participants were instructed to memorize motor sequences, as in our previous studies on retrieval-induced forgetting, test-potentiated learning, and directed forgetting. Items were SFMs, that is, *item* does not denominate a single key-press response but a response pattern consisting of four subsequent finger movements (i.e., key presses) in a defined sequence. These items were assigned to several study lists. The lists (comprising between three and seven SFMs) were studied one after the other. We manipulated tasks separating study lists. Participants were asked to retrieve the items of the just studied list (testing) or received an additional study cycle (restudy) or worked on an unrelated distractor task (arithmetic problems). We conducted four experiments. In Experiment 1 we tried to replicate the original experiments of Szpunar et al. (2008) with SFMs as material. In Experiment 2 we changed the paradigm so as to minimize interference effects between study lists and thus disentangled the potential effects of encoding from interference. In Experiments 3 and 4 we developed the design further and foremost manipulated all independent variables within participants. To foreshadow the results, we observed three main results here. First, in all four experiments items of the test list (which was presented after the retrieval vs. restudy manipulation) were better recalled in the retrieval-group/condition (i.e., we observed the forward effect of testing in motor memory). Second, in all four experiments the speed of encoding in the study phase after retrieval was slower as compared to the group/condition after restudy (suggesting better encoding after retrieval). Third, in three out of four experiments speed of encoding correlated with recall performance, suggesting that retrieval leads to better encoding in subsequent study trials.

Experiment 1

Participants learned four lists of five SFMs each. The design followed that of Szpunar et al. (2008) very closely. Two groups were compared: The individual lists were either separated by retrieval of the items of the just presented list or by one additional restudy cycle (after the items had been presented several times per list). Finally, both groups were required to recall the items of the fourth list. We expected the retrieval group to recall more SFMs of the fourth list than the restudy group. In addition, we expected RTs at executing the SFMs of the fourth list to differ between groups, indicating better encoding quality in the retrieval group.

Method

Participants Eighty-eight undergraduate students at the University of Trier (44 per group) either received course credit for their participation or were paid 8 €. Based on our experience with motor sequences as material in memory experiments we

expected a medium-sized to large effect of Cohen's d (Szpunar et al., 2008, reported large effect sizes) and calculated the sample size based on $\alpha = .05$ and $1 - \beta = .8$ (power analysis was run with G-Power; Faul, Erdfelder, Lang, & Buchner, 2007).

Design The study had a 4 (study list) \times 2 (list-separating task: retrieval, restudy) mixed design with repeated measures on the first factor.

Material The experiment was conducted using Dell Optiplex 755 PCs with Eizo FlexScan S1901 monitors and standard German QWERTZ keyboards. The software PXLab (Irtel, 2007) served for running the experiment.

Items were SFMs. Each item was a four-finger movement to be performed with fingers of the right hand. Altogether, 20 SFMs were studied in four subsequently presented lists of five SFMs each (L1 to L4). Four sets of five items each were constructed and assigned to L1 to L4 in two sequences (ABCD, CDAB), counterbalanced between participants. The items of sets A and B were highly similar to each other, as were the items of sets C and D, in order to maximize attentional demands when proceeding from L3 to L4, that is for each item in one set there was a highly similar item in the other set only differing with regard to the last finger (see Appendix A).

There were two different kinds of trials: study trials and retrieval trials. During study trials, an animation of the four-finger movement appeared on the screen (cf. Fig. 1, upper section). The participants placed their right index finger, middle finger, and ring finger on the marked keys ‘,’ ‘.’, and ‘-’. A display of the right hand demonstrated which fingers should be moved by showing four consecutively flashing fingers (first finger was colored yellow, second finger was colored blue, third finger was yellow again, fourth finger was colored blue again; 200 ms per flash). After the display of the hand disappeared participants could perform the movement. If the performed sequence was incorrect a feedback appeared, displaying: “Fehler!” (English: “Error!”). In retrieval trials, participants entered SFMs in response to an exclamation point appearing on the screen (cf. Fig. 1, lower section). As soon as the exclamation point appeared, input could begin. After pressing four keys, it disappeared from the screen. It reappeared after 1 s, signaling input of the next SFMs.

Procedure The experiment consisted of ten phases (study of L1, retrieval or restudy of L1, study of L2, retrieval or restudy of L2, study of L3, retrieval or restudy of L3, study of L4, distractor, test for L4, test for L1-to-L4). Instructions were given on the screen. Participants were told that they were going to learn four lists of five SFMs each, repeated in several cycles per list. In addition, they were informed that each of the four parts of the learning phase would be followed by one of three tasks: retrieval of the just studied list, restudy in one further cycle, or solving arithmetic problems.

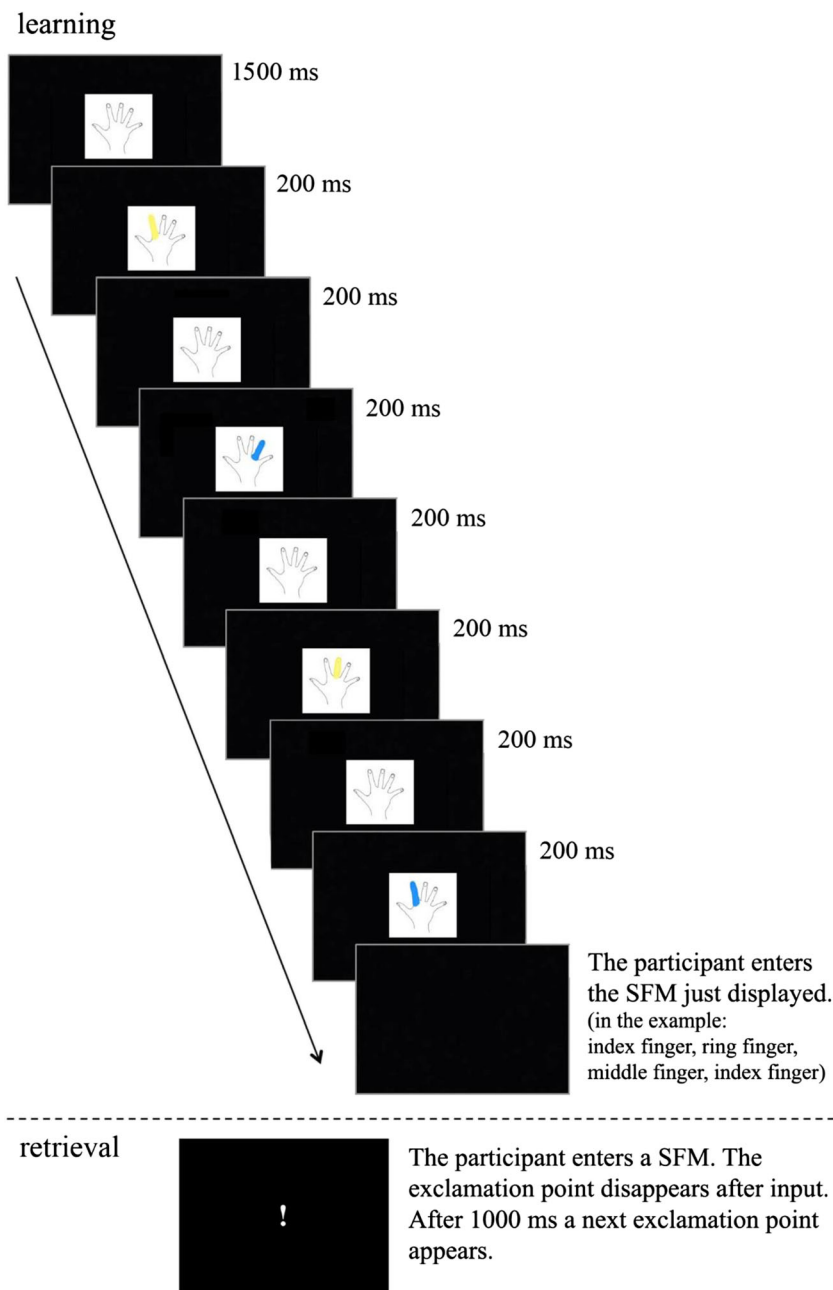


Fig. 1 The upper section depicts a study trial. It starts with a drawing of the right hand. After 1,500ms the first finger illuminates yellow for 200 ms followed by 200 ms of the uncolored drawing. Then the second finger flashes blue, the third finger yellow, and the fourth finger blue again in the same fashion. Subsequently, the hand display disappears

and the participant can enter the SFM just illustrated. The lower section depicts a retrieval trial. Participants are supposed to enter a SFM as soon as an exclamation point appears on the screen. *SFM* sequential finger movement

After having read the initial instructions for the learning phase, the participant clicked an on-screen button to start with study of L1. The five items of one list were repeated in 15 cycles presenting the items in a random order in each study phase, that is, after all five SFMs had been studied (one cycle) they were studied again (next cycle, repeated 14 times). Restudy repeated the five items in one additional cycle with the identical format as during previous study. All retrieval phases comprised five retrieval trials. Participants were

encouraged to guess if they were not able to recall all five items of a list with certainty. In the distractor phase, several arithmetic problems were presented simultaneously on the screen for 60 s while participants were supposed to write down solutions on a sheet of paper. After the distractor phase, both groups received a recall test for L4, that is, participants were instructed to retrieve and execute the five SFMs of the just learned list. Subsequently, both groups received a further recall test asking for retrieval of all 20 previously learned SFMs.

Results

For analyses of RTs in the study phases, trials with incorrect input as well as RTs exceeding 3 SD from the mean were excluded. RTs were measured as time between the end of the sequence animation and completed movement execution (i.e., the last key press of the respective sequence). Double input of SFMs in the test phase for L4 or, respectively, the subsequent test phase for L1-to-L4 was considered as incorrect recall, that is, when a SFM was entered twice it was counted as recalled only once. To-be-recalled SFMs were scored all-or-none, that is, a SFM was scored as recalled if all four responses were correct. We included control factors representing the counterbalancing of items in all ANOVAs. Dependent variables were individual arithmetic means of RTs or, respectively, the number of recalled SFMs. Analyses followed the same principles and exclusion criteria in Experiments 2–4 as well.

Study phases

Whereas the groups did not differ with regard to input speed for L1, L2, or L3, $|t_s| < 1$, input of L4 items was significantly slower in the retrieval group than in the restudy group, $t(86) = 2.18$, $p = .032$, $d = 0.47$. Additionally, a 4 (study phase) \times 2 (group) \times 2 (item sequence) ANOVA with repeated measures on the first factor examined RTs in study trials. A significant main effect of study phase demonstrated a general acceleration from L1 to L4, $F(3, 252) = 60.89$, $p < .001$, $\eta_p^2 = .42$, whereas the main effect of group was not significant, $F < 1$. The linear trends were computed as to code the main effect of study phase entering the means for L1 to L4 in consecutive order. Both groups did show linear trends (i.e., the slopes of the regression lines as depicted in Fig. 2). Separate analyses per group showed that the linear trend was weaker in the retrieval group, $F(1, 42) = 16.50$, $p < .001$, $\eta_p^2 = .28$, than in the restudy group, $F(1, 42) = 62.19$, $p < .001$, $\eta_p^2 = .60$. The slopes differed significantly between groups, $F(1, 84) = 6.20$, $p = .015$, $\eta_p^2 = .07$ showing that the acceleration differed between the retrieval and restudy group.¹

Test phase The retrieval group correctly recalled significantly more SFMs in the test for L4 than the restudy group, $t(86) = 3.19$, $p = .002$, $d = 0.68$ (see Table 1), and showed significantly fewer intrusions from previous lists ($M = 27\%$) than the restudy group ($M = 43\%$), $t(86) = 2.58$, $p = .011$, $d = 0.56$. In the test for L1-to-L4, separate t-tests for items of the individual lists showed that groups did not differ regarding recall of L1, L2, L3, or L4 items, $|t(86)| < 1.12$, $p > .269$. We assume that the limited item

pool contributed to guessing in a way that might have blurred any differences between groups in the test for L1-to-L4.

Study-test link The longer RTs in study trials for L4 in the retrieval group might indicate more attentive encoding, perhaps, predicting the better recall of L4 items in this group. Therefore, we analyzed the link between encoding and memory accessibility by examining the correlation between study-trial RTs and the number of recalled items. However, there was no significant correlation between study-trial RTs for L4 and the recall of L4 items, $r = -.01$, $p = .901$.

Discussion

Testing reduced the acceleration of motor sequence execution during study. This relative slowing-down (as compared to the restudy group) did not reflect a decrease in encoding quality. On the contrary, it came along with superior memory for SFMs of the last study list. The slower execution of these SFMs entailed their better recall. Typically, accelerations are considered as indices of sequence learning (e.g., in the serial reaction time task; Nissen & Bullemer, 1987). However, they probably reflect limited practice effects that are insufficient for impacting explicit sequence knowledge. It has been shown that such practice effects do not require attention to be focused on the sequential structure, as they equally occur under dual-task conditions (Curran & Keele, 1993). In addition, mere sequence practice by means of repeated execution does not necessarily translate into explicit retrievability or transfer (e.g., Destrebecqz & Cleeremans, 2001; Shanks & Perruchet, 2002), both of which are promoted by intentional sequence learning (e.g., Dominey, Lelekov, Ventre-Dominey, & Jeannerod, 1998; Jiménez, Vaquero, & Lupiáñez, 2006). Here, instructions generally required intentional encoding. Thus, the relatively slower execution does not indicate the presence of intention only in the retrieval group but it demonstrates a higher degree of attention as compared to the restudy group. Participants took more time for carefully studying the to-be-retained SFMs. This finding corresponds to a recent study by Yang, Potts, and Shanks (2017) showing that participants chose to allocate more time to studying word lists after receiving tests on previously studied word lists as compared to not receiving tests when they were free to spend as much time as they liked for studying. More time on study trials there also was accompanied with better recall in a final test, as was slower movement execution in the present experiment. A crucial difference was, of course, that participants here were not free to allocate study time, yet, relatively slower movement execution was beneficial for later recall.

The results match the theoretical assumption that retrieval causes more attentive encoding in subsequent study trials (Pastötter et al., 2011). Yet, RTs in study trials did not predict recall of L4 items but it could be premature to take this particular null result as evidence against an impact on encoding. It might

¹ The somewhat slower execution of L3 than L2 in both groups in Experiment 1 reflects that, while the SFMs of L2 were very similar to the SFMs of L1 and the SFMs of L4 were very similar to L3, there was less similarity between the items of L2 and L3

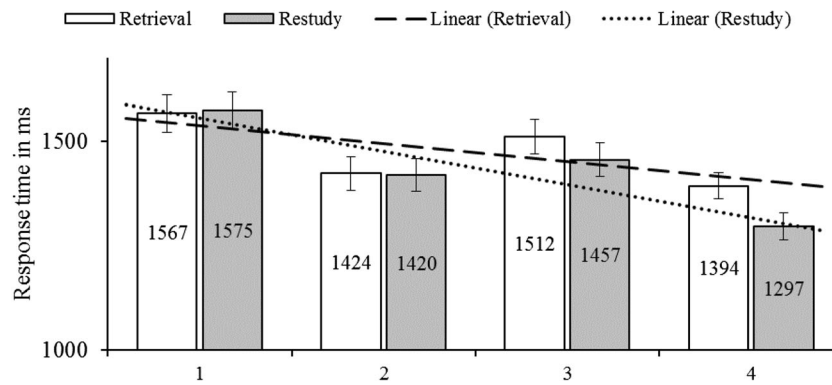


Fig. 2 Mean response times per study trial as a function of study list (1–4) and list-separating tasks in Experiment 1, including linear trends within groups. Error bars depict standard error of the mean

indicate that the retrieval group was able to recall more L4 items because previous retrieval reduced interference between individual lists but not because of better encoding. The format of the final recall tests was highly susceptible to interference produced by retrieval attempts because it required recall of L4 items only. Interference by L1, L2, and L3, therefore, might have been so strong that encoding processes did not substantially determine recall performance. In a second experiment, we aimed at reducing the relative impact of (lower) interference between lists in order to allow for a potentially stronger contribution of encoding processes. Thus, our aim was to disentangle effects of interference that is triggered by recalling items in the final test phase from effects of encoding that occur during study.

Experiment 2

Participants learned only two lists of SFMs but the number of items per list was increased to seven. Again, two groups were compared. The retrieval group received a recall test for L1 before studying L2 whereas the restudy group received one additional study cycle of L1. The format of the final test phase was changed to a recall test for all items irrespective of which list they belonged to. Not restricting the test to only one (the last) list eliminated the need for participants to distinguish items from individual lists during retrieval attempts. Thus, recall performance was less affected by interference between lists.

Table 1. Proportional recall of items in dependence of preceding list-separating tasks

	Retrieval		Restudy		Arithmetic	
	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>
Experiment 1	.61	.04	.41	.05	-	-
Experiment 2	.69	.02	.61	.02	-	-
Experiment 3	.44	.03	.31	.03	.28	.03
Experiment 4	.59	.03	.47	.03	-	-

We expected to replicate the results of Experiment 1: slower execution of L2 study trials and higher recall of L2 items in the retrieval group than in the restudy group. In addition, we assumed the modifications to enable a relatively stronger contribution of encoding processes to recall performance compared to interference between lists and, therefore, expected a correlation between RTs in study trials and recall performance.

Method

Participants One hundred and twenty undergraduate students at the University of Trier (60 per group) either received course credit for their participation or were paid 4 €.

Design The study had a 2 (study list) × 2 (list-separating task: retrieval, restudy) mixed design with repeated measures on the first factor.

Material and procedure The material again consisted of four-finger movements learned in the same fashion as in Experiment 1 but the number of lists was reduced to two whereas the number of items per list was increased to seven (see Appendix B). The experiment consisted of five phases (study of L1, retrieval or restudy of L1, study of L2, distractor, final recall test). The seven items of one list were repeated in ten cycles presenting the items in a random order. Restudy repeated the seven items in one additional cycle. After the distractor phase, both groups received a final recall test for the items of L1 and L2, that is, participants were instructed to retrieve and execute all previously learned SFMs. In all other respects, the procedure matched Experiment 1.

Results

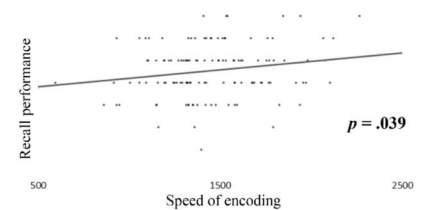
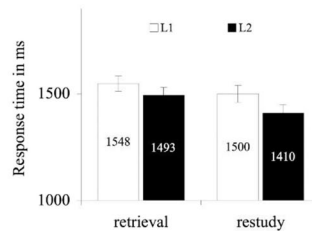
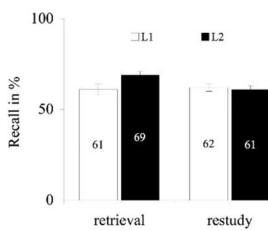
Study phases Input speed of L1 items was similar in both groups, $|t| < 1$. Input of L2 items was slightly slower in the retrieval group than in the restudy group, though the effect did not reach conventional levels of statistical significance, $t(118) = 1.54, p = .064$, one-tailed, $d = 0.28$ (see Fig. 3, upper

Recall performance

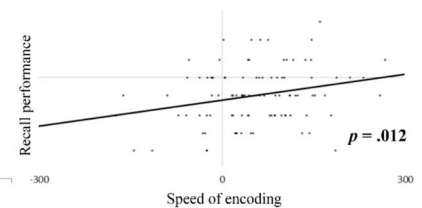
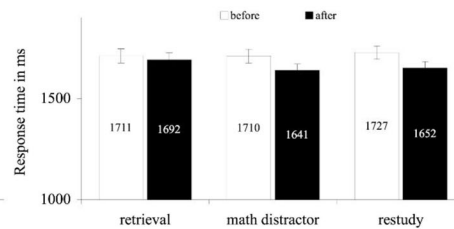
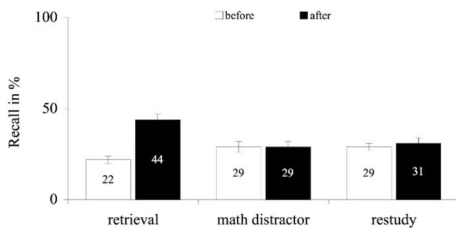
Speed of encoding

Correlation between recall performance and speed of encoding

Experiment 2



Experiment 3



Experiment 4

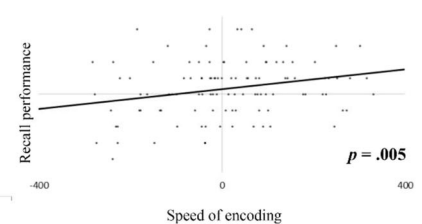
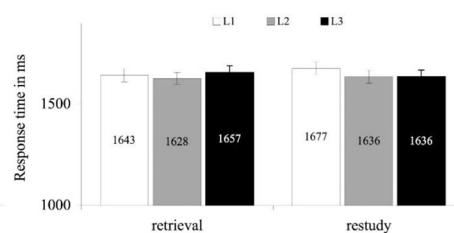
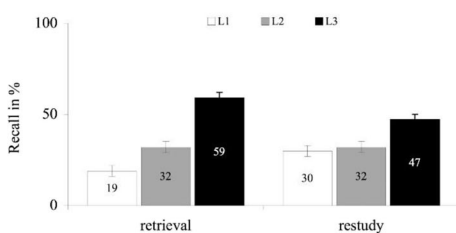


Fig. 3 Overview of the main results across Experiments 2–4. The left section shows recall performance as a function of study list and list-separating tasks in Experiments 2, 3, and 4. The middle section shows mean response times per study trial as a function of study list and list-

separating tasks in Experiments 2, 3, and 4. The right section shows scatterplots representing the correlation between speed of encoding and recall. See the result sections for further details. Error bars depict standard error of the mean

section). An across-experiments 2 (Experiment 1, Experiment 2) \times 2 (retrieval, restudy) ANOVA examining input speed in the last study list (L2/L4) showed that the main effect of group, $F(1, 204) = 5.96, p = .016, \eta_p^2 = .03$, was not significantly moderated by Experiment, $F < 1$.

Test phase The retrieval group correctly recalled significantly more SFMs from L2 than the restudy group, $t(118) = 2.45, p = .016, d = 0.45$. Groups did not significantly differ regarding recall of L1 items, $|t| < 1 (M = 62\%)$.

Study-test link We analyzed the link between encoding and memory accessibility by examining the correlation between study-trial RTs and the number of recalled items. Whereas there was no significant correlation between study-trial RTs for L1 and the recall of L1 items, $r = -.04, p = .680$, the correlation between study-trial RTs for L2 and the recall of L2 items was significant, $r = .19, p = .039$, indicating that longer encoding time predicted better recall. The two correlations differed significantly from each other, $z = 1.75, p = .04$.

Discussion

The main findings of Experiment 1 were replicated. A twofold forward effect of testing occurred. The retrieval group recalled more L2 items and tended to execute them more slowly in the study phase. A comparison with Experiment 1 showed that the effect on movement execution was stable over experiments. In addition, execution speed predicted recall of L2 items supporting the assumption that retrieval of L1 caused more attentive encoding that accounted for better recall in the retrieval group.

Experiment 3

The procedural modifications in Experiment 2 resulted in the demonstration of a link between RTs in study trials with recall performance. However, this demonstration came at a cost of only a relatively weak difference in mean RTs between the retrieval and restudy group. Paralleling the reduced

acceleration of execution speed in the retrieval group of Experiment 1, this weak difference also matches the finding that mean RTs between groups significantly differed only with regard to the last of four lists there. Hence, reducing the number of lists to two probably precluded the emergence of a comparably strong effect on RTs. In Experiment 3, we followed a different approach, increasing the number of lists again but manipulating list-separating tasks within participants.

This design also allowed testing an alternative explanation of the retrieval group's recall advantage to the assumption of a forward effect of testing. Perhaps, participants simply were able to give more correct responses because they were already familiar with the test format whereas the restudy group received retrieval trials for the first time only in the final test phase. A within-participants manipulation of list-separating tasks, then, would not result in a benefit for items learned after retrieval of a just studied list because all participants were equally familiar with retrieval trials.

Participants studied four lists of four SFMs each. The first three lists were followed by three different tasks: retrieval of the just learned list, restudy of the list, or solving arithmetic problems. The assignment of these tasks was counterbalanced between participants. L4 was always followed by solving arithmetic problems and, subsequently, by a final recall test for the items of all four lists. We expected retrieval again to reduce an acceleration of execution speed in study trials. In addition, we expected significantly higher recall of SFMs learned subsequently to retrieval of a previous list as compared to SFMs learned after restudy of a previous list or after solving arithmetic problems.

Method

Participants Ninety-six undergraduate students at the University of Trier either received course credit for their participation or were paid 8 €.

Design The study had a one-factorial (list-separating task: retrieval, restudy, arithmetic problems) design with repeated measures.

Material and procedure The material consisted of four lists of four four-finger movements each, learned in the same fashion as in the previous experiments. The participants placed their right index finger, middle finger, ring finger, and pinkie on the marked keys 'M', ',', '.', and '-'. Because there were now four fingers being used, for most subsequences no key press was repeated within each four-item subsequence. However, as in Experiments 1 and 2, when one key press was repeated, the repetition never occurred consecutively (see Appendix C). The items of one list were repeated in 15 cycles presenting the items in a random order. Participants learned the four lists separated by different tasks: retrieval of the just studied list, restudy of the

just studied list for one additional study cycle, or solving arithmetic problems for 60 s. The assignment of tasks was counterbalanced between participants. In addition, the assignment of items to the four lists was counterbalanced between participants. As a consequence, all used SFMs were assigned equally often to each list number as well as to list status (regarding list-separating tasks) across participants. After L4, all participants solved arithmetic problems for 60 s. In the subsequent final recall test, they were instructed to retrieve and execute all previously learned SFMs. In every other respect, the procedure matched the previous experiments.

Results

We analyzed the impact of list-separating tasks from two perspectives. Our main focus was on examining the impact on subsequent study trials. For this purpose, we analyzed a factor of *preceding task* that classified a study list according to the list-separating task that preceded it. Additionally, we examined potential impacts of *subsequent task* by a factor that classified a study list according to the task performed after studying that list. The three levels of both factors were retrieval, restudy, and solving arithmetic problems.

Study phases A 3 (preceding task) \times 6 (task sequence) \times 4 (item sequence) ANOVA with repeated measures on the first factor examined RTs in study trials. The main effect of preceding task was significant, $F(2, 144) = 4.22, p = .016, \eta_p^2 = .06$. Planned contrasts showed that execution of SFMs was significantly slower after retrieval of a preceding list than after restudy of a preceding list or solving arithmetic problems, $F(1, 72) = 6.85, p = .011, \eta_p^2 = .09$, whereas SFM execution after restudy or solving arithmetic problems did not differ significantly, $F < 1$ (see Fig. 3, middle section). Additionally, a 3 (subsequent task) \times 6 (task sequence) \times 4 (item sequence) ANOVA with repeated measures on the first factor examined RTs in study trials. The main effect of subsequent task was not significant, $F < 1$. Moreover, there were significant accelerations in input speed from before to after restudy of a list, $t(95) = 4.34, p < .001, d = 0.45$, as well as from before to after solving arithmetic problems, $t(95) = 3.97, p < .001, d = 0.41$, but there was no acceleration from before to after retrieval, $|t| < 1$.

Test phase A 3 (preceding task) \times 6 (task sequence) \times 4 (item sequence) ANOVA with repeated measures on the first factor examined the number of recalled items. The main effect of preceding task was significant, $F(2, 144) = 11.95, p < .001, \eta_p^2 = .14$. Planned contrasts showed that recall of SFMs learned after retrieval of a preceding list was higher than recall of SFMs learned after restudy of a preceding list or after solving arithmetic problems, $F(1, 72) = 20.69, p < .001, \eta_p^2 = .22$, whereas recall of SFMs learned after restudy or solving arithmetic problems did not differ significantly, $F < 1$.

Additionally, a 3 (subsequent task) \times 6 (task sequence) \times 4 (item sequence) ANOVA with repeated measures on the first factor examined the number of recalled items. There was a marginal main effect of subsequent task, $F(2, 144) = 2.93$, $p = .057$, $\eta_p^2 = .04$. Fewer SFMs of the previously to-be-retrieved list were recalled ($M = 22\%$) than restudied SFMs ($M = 28\%$) or SFMs studied before solving arithmetic problems ($M = 28\%$). Furthermore, the before-after difference was only significant with regard to retrieval separating two lists, $t(95) = 5.98$, $p < .001$, $d = 0.86$ (i.e. more items that had been studied after retrieval of a previous list were recalled as compared to items from the to-be-retrieved list), but not regarding restudy or solving arithmetic problems, $|ts| < 1$.

Study-test link The double status of study lists being encoded before a list-separating task as well as after a list-separating task precluded distinct analyses examining the link between study-trial RTs and recall with regard to the different tasks. However, we did analyze the study-test link by examining the correlation between the general acceleration and the corresponding difference in recall of items learned before and after a list-separating task. The correlation was significant, $r = .26$, $p = .012$, indicating that stronger accelerations predicted lower recall of items, that is, the less time was spent on study trials the fewer of those items were recalled.

Discussion

Again, a twofold forward effect of testing occurred. Retrieval as a list-separating task slowed down movement execution in subsequent study trials and enhanced recall of SFMs from these trials. From before to after restudy of a just learned list as well as from before to after an unrelated distractor task, there was an acceleration in study-trial RTs that vanished when retrieval of the just learned list separated two lists. This acceleration was in fact a predictor of recall performance, with weaker accelerations predicting stronger benefits for items of the list learned after the list-separating task. Hence, the non-accelerated execution of SFMs learned subsequently to retrieval of a previous list involved their superior recall. However, the present design did not allow for calculating an index representing the forward effect of testing as a performance difference with regard to items learned after retrieval and items learned after restudy or solving arithmetic problems because we had counterbalanced the sequence of list-separating tasks. The design demonstrated that retrieval precluded an acceleration of study-trial RTs by comparing mean RTs before and after list-separating tasks. However, the general acceleration from L1 to L4 was contaminated with accelerations or, respectively, non-accelerations from one list to the next one on an individual level, thus, ruling out a correlation of an individual index of the forward effect of testing on study-trial RTs with an index of the forward effect of testing on recall.

Experiment 4

We changed the design once more. Experiment 4 consisted of two main parts. Each part comprised studying three lists of SFMs and ended with a recall test for the items of these three lists. Retrieval of a just studied list separated the three lists in one part whereas restudy separated lists in the other part. We expected a forward effect of testing with regard to a reduced acceleration of study-trial RTs in the retrieval part compared to the restudy part and with regard to recall performance.

Method

Participants One hundred and eight undergraduate students at the University of Trier either received course credit for their participation or were paid 8 €.

Design The study had a 3 (study list) \times 2 (list-separating task: retrieval, restudy) design with repeated measures on both factors.

Material and procedure The material consisted of altogether six lists of three four-finger movements each, learned in the same fashion as in the previous experiments. The participants placed their right index finger, middle finger, ring finger, and pinkie on the marked keys ‘M’, ‘,’’, ‘.’’, and ‘-’ (see Appendix D). The items of one list were repeated in ten cycles presenting the items in a random order. The experiment consisted of two main parts, each comprising study of three lists and a final recall tests for all items of these three lists. In the retrieval part, retrieval of a just learned list separated lists. In the restudy part, restudy of a just learned list separated lists. The sequence of the two parts was counterbalanced between participants. In addition, the assignment of items to lists was counterbalanced between participants. As a consequence, all used SFMs were assigned equally often to each list number as well as to list status (regarding list-separating tasks) across participants. In both parts, participants solved arithmetic problems for 60 s after studying the third list. In every other respect, the procedure matched the previous experiments.

Results

Study phases A first 3 (list) \times 2 (task sequence) \times 6 (item sequence) ANOVA with repeated measures on the first factor examined RTs in the restudy block. A significant main effect of list indicated an acceleration from L1 to L3, $F(2, 192) = 3.84$, $p = .023$, $\eta_p^2 = .04$, also reflected by a significant linear trend, $F(1, 96) = 4.79$, $p = .031$, $\eta_p^2 = .05$. A second 3 (list) \times 2 (task sequence) \times 6 (item sequence) ANOVA with repeated measures on the first factor examined RTs in the retrieval block. The main effect of list was not significant, $F(2, 192) = 1.20$, $p = .303$, neither was the respective linear-trend

analysis, $F < 1$. Additionally, a 3 (list per block) \times 2 (list-separating task) \times 2 (task sequence) \times 6 (item sequence) ANOVA with repeated measures on the first two factors showed that the significant linear trend indicating acceleration in the restudy block differed significantly from the absent acceleration in the retrieval block, $F(1, 95) = 4.64$, $p = .034$, $\eta_p^2 = .05$ (see Fig. 3, lower section).

Test phase A 3 (list per block) \times 2 (list-separating task) \times 2 (task sequence) \times 6 (item sequence) ANOVA with repeated measures on the first two factors showed that there was a general recall advantage for items from later lists, indicated by a significant main effect of list per block, $F(2, 192) = 68.03$, $p < .001$, $\eta_p^2 = .42$, also reflected by a significant linear trend, $F(1, 96) = 104.11$, $p < .001$, $\eta_p^2 = .52$, whereas the main effect of list-separating task was not significant, $F < 1$. However, the linear trend indicating a recall advantage for later lists differed significantly between blocks, $F(1, 96) = 20.14$, $p < .001$, $\eta_p^2 = .17$. Separate 3 (list) \times 2 (task sequence) \times 6 (item sequence) ANOVAs with repeated measures on the first two factors per block confirmed that this linear trend was stronger in the retrieval block, $F(1, 96) = 108.16$, $p < .001$, $\eta_p^2 = .53$, than in the restudy block, $F(1, 96) = 20.47$, $p < .001$, $\eta_p^2 = .18$. This particularly strong difference also reflects that the recall advantage of L3 items in the retrieval block (see Table 1) came at the expense of lower recall of L1 items in the retrieval block ($M = 19\%$) as compared to the restudy block ($M = 30\%$), whereas mean recall of L2 items was 31% in both blocks.

Study-test link

In contrast to Experiment 3, the design now allowed analysis of the link between RTs in study trials and recall performance with regard to the specific list-separating tasks. For this purpose, we examined the correlation between the recall difference of items learned after the retrieval of a preceding list and items learned after restudy of a preceding list (as an index of the forward effect of testing in memory performance) with the difference between RTs in the respective study trials. The correlation was significant, $r = .27$, $p = .005$.

Discussion

Experiment 4 replicated the twofold forward effect of testing once again. Retrieval as a list-separating task slowed down movement execution in subsequent study trials and enhanced recall of SFMs from these trials. In addition, we now could calculate individual difference scores between items learned after retrieval of previously learned lists and items learned after restudy of previously learned lists with regard to recall performance as well as with regard to study-trial RTs. The correlation between these two measures suggests that more attentive encoding accounted for the benefit in recall. Beside

an influence of retrieval on subsequently learned items, there was also an influence of retrieval on later recall of the to-be-retrieved list. Fewer L1 items were recalled in the retrieval block. This lower recall might reflect that the strengthened L3 items blocked access to L1 items.

General discussion

Four experiments found converging evidence that direct memory tests enhance motor practice. Whereas an acceleration of movement execution over repeated study trials certainly can be regarded (and often is) as reflecting a kind of learning, it must not be mistaken as indicating explicit accessibility in memory. Retrievability depends on attentive study. Here, RTs reflected both practicing motor-sequence execution (acceleration) and intentional encoding that benefited from preceding tests reducing the acceleration over repeated study cycles. These two facets of motor practice could be distinguished because we followed an approach from memory research on retrieval effects.

The present study demonstrates a twofold forward effect of testing, affecting a measure of encoding quality (study-trial RTs) as well as recall performance in a final test. The results nicely match prior electrophysiological findings (Pastötter et al., 2011). These two lines of evidence together suggest that retrieval causes more attentive encoding, similar to directed forgetting of a previously studied item list (Pastötter & Bäuml, 2010; Tempel & Frings, 2016b). This can also be understood as test potentiation. Although studies on test-potentiated learning involve restudy of tested items, the idea of test potentiation implies that retrieval enhances subsequent encoding as well (e.g., Arnold, & McDermott, 2013). In particular, the study by Tempel and Kubik (2017) showed that testing enhanced restudy of motor sequences. In contrast, the forward effect of testing refers to study of novel items, of course. Here, we show that study of novel motor sequences was also enhanced by testing.

Experiments 2, 3, and 4 additionally showed that indices of study-trial RTs reflecting encoding quality predicted recall performance. There was a general link of longer RTs with better recall as well as between a specific index of the forward effect of testing on study-trial RTs with an index of the forward effect of testing on recall performance. Taken together, these results suggest that retrieval of a just learned list enhances encoding. Retrieval slowed participants down (i.e. reduced the acceleration) because encoding of the to-be-learned movements became better. The relatively longer RTs compared to study trials after restudy of a just presented list or after solving arithmetic problems reflect more attentive processing that entailed better accessibility at the later recall test. The deceleration might also be linked to test expectancy. Weinstein, Gilmore, Szpunar, and McDermott (2014)

demonstrated that receiving tests after individual study lists heightened expectancy of another test after the next study list. Perhaps, test expectancy was heightened by retrieval as a list-separating task in the present study as well.

Although Experiment 1 did find a forward effect of testing on study-trial RTs, there was no link with recall performance. We had followed the original design by Szpunar et al. (2008) most closely in that experiment. Probably, recall was strongly influenced by interference between lists whereas the design precluded a substantial contribution of encoding processes to impact recall. Therefore, we switched from a test asking for recall of items from only one (i.e. the last) list to a format not restricting recall but asking for recall of all previously learned items. This abandonment of a test format that is strongly susceptible to between-list interference allowed detecting that study-trial performance in fact predicted memory accessibility. Moreover, the reduced number of lists in Experiment 2 and the change to a within-participants design in Experiments 3 and 4 also strengthened the relative influence of encoding on recall performance, which resulted in the observed correlations. This shows how the choice of experimental design determines what you are able to detect. The use of motor sequences opened up a window into encoding processes because movement execution was measured and could be analyzed as a dependent variable impacted by the forward effect of testing. Yet, the difference in encoding apparently did not contribute much to recall performance in Experiment 1 compared to the subsequent experiments. This is probably also true for other kinds of materials. However, items that do not involve overt action during learning (e.g., reading words or sentences, listening to sound, or watching images) are not capable to reveal such effects. Response measures during encoding should be used more extensively in memory research. Moreover, the occurrence of the same effects with a newly developed within-participants manipulation of list-separating task as in a between-participants design rules out the alternative explanation that retrieval entailed a benefit for subsequently learned SFMs because participants already were familiar with test trials whereas participants that had received restudy were not.

The observed impact of testing on subsequent study trials was not merely an effect of task switching. Note that the different task affordances of study and retrieval trials might have contributed to a somewhat slower movement execution after switching from retrieval back to study trials. However, the affordance of executing motor sequences actually did not distinguish these two tasks but remained the same. Moreover, retrieval slowed down movement execution in subsequent study trials also compared to the unrelated distractor task of solving arithmetic problems that really involved a substantial change in task affordances. Taken together, it seems safe to conclude that the slowed down execution did not reflect switch costs of mere task switching (e.g., Monsell, 2003).

In memory research, it is very common to investigate incidental as well as intentional learning. Beside an unsurprising general advantage of intentional learning on explicit accessibility, some factors influence intentional and incidental learning to a similar degree, such as, classic levels-of-processing tasks (Roediger & Gallo, 2002), whereas others do not. For example, self-reference has been found to improve memory for incidentally encoded items more strongly than intentionally encoded items (Symons & Johnson, 1997). In addition, there are effects that typically are only examined with regard to intentional learning, such as, retrieval effects. Retrieval is an important component of practice. In recent years, memory researchers have emphasized the educational relevance of retrieval effects and advocated the beneficial consequences of testing on retention (e.g., Karpicke & Blunt, 2011; Roediger, Agarwal, McDaniel, & McDermott, 2011). Such benefits pertain to using retrieval as a tool during practice. The relevant processes of practice thus are controlled cognitive processes. Learning results not merely as a product of repetition but is an outcome of memorization, that is, actively storing information in an organized manner. By instructing participants to memorize motor sequences, we were able to demonstrate that benefits of retrieval extend to motor practice, discovering that longer study RTs indicated an immediate benefit of retrieval. The relative slowing-down of movement execution that reflected enhanced memorization contrasts sharply with the general acceleration over repeated study cycles that was equally present but reflects a standard sequence-learning effect. The benefit on accessibility occurred not due to learning to execute the SFMs smoothly and quickly but due to enhanced memorizing. This controlled storing requires care and time. Although in all experiments execution of motor sequences became quicker over the course of the learning phases (i.e. participants demonstrated learning regarding movement execution), the core manipulation of contrasting retrieval practice with other list-separating tasks slowed down movement execution. The observed correlations with recallability in the memory test suggest that slower sequence execution reflected intentional encoding processes.

Intentional motor practice is especially common in the domains of sports and music. For example, dancers rehearsing a new choreography must memorize the individual dance steps and their sequence. When they change choreographies, new movement patterns are acquired. A piano player practices defined sets of sequential finger movements when learning to play a new piece of music. Memorizing the sequence is a crucial prerequisite before musical expression can show a musician's individual interpretation of a composition. The executed actions immediately allow observing learning progress. Yet, smoother execution after several repetitions ought not to be mistaken for an index of superior explicit access at later retrieval attempts, both of which are crucial for mastering a

performance. Absent-mindedness may be beneficial for implicit motor knowledge whereas explicit memorization requires focused attention on the features of to-be-learned movements.

Testing effects have been investigated extensively but mostly with regard to verbal materials. We here demonstrate forward effects of testing in motor action, which proves the broad generalizability of testing effects across domains. Training, for example, in sports and music, should include retrieval of body movements in their exercises as a learning tool. Indeed, retrieval practice already is an integral element of many training procedures. For example, musicians regularly switch to rehearsing a piece of music by rote after playing it from notes (e.g., Chaffin, Lisboa, Logan, & Begosh, 2010). Athletes preparing for a competitive performance requiring the precise execution of a defined sequence of body movement, such as in dancing, in bobsleigh runs, or at a high bar, have to retrieve practiced motor sequences. Yet, the effects of retrieval in sports are only poorly understood so far. Investigations on different forms of practice have, for example, scrutinized different practice schedules (e.g., Landin & Herbert, 1997) or compared imagery and physical practice (e.g., Feltz & Landers, 1983), but the specific consequences of memory retrieval have been overlooked to date in this particular field. The present results suggest to systematically include testing as a training tool particularly for the purpose to enhance subsequent study. When a training session consists of several episodes each on a different sequence (e.g., different parts of one choreography), then recalling the content of an episode before proceeding with the next episode can enhance subsequent learning. Thus, it might be advantageous to end training episodes by tests not by restudy (e.g., not by continued repetition of an instructor's model). To preclude negative side effects on memory for earlier training episodes, that is, to preclude that the enhanced learning of later episodes blocks access, initial episodes might benefit from test-potentiated learning (Tempel & Kubik, 2017) through restudy of such earlier episodes. Testing will benefit explicit access, in particular. Recalling motor sequences is enhanced when testing is used as a learning tool during practice. Teachers, trainers, musicians, or athletes, need to become aware of the powerful effects of testing on learning.

Acknowledgements This research was supported by grant TE 891/3-3 of the German Research Council (DFG).

Compliance with ethical standards

Open practices statement The data for all experiments are available at www.researchgate.net/publication/332230269_TEMP_E1_to_E4.

Appendix A

Table 2. Items in Experiment 1

set	first finger	second finger	third finger	fourth finger
A	index finger	ring finger	middle finger	index finger
A	index finger	middle finger	ring finger	middle finger
A	ring finger	middle finger	index finger	ring finger
A	ring finger	index finger	middle finger	index finger
A	middle finger	index finger	middle finger	ring finger
B	index finger	ring finger	middle finger	ring finger
B	index finger	middle finger	ring finger	index finger
B	ring finger	middle finger	index finger	middle finger
B	ring finger	index finger	middle finger	ring finger
B	middle finger	index finger	middle finger	index finger
C	index finger	ring finger	index finger	middle finger
C	middle finger	ring finger	index finger	ring finger
C	middle finger	index finger	ring finger	index finger
C	ring finger	middle finger	ring finger	middle finger
C	ring finger	index finger	ring finger	middle finger
D	index finger	ring finger	index finger	ring finger
D	middle finger	ring finger	index finger	middle finger
D	middle finger	index finger	ring finger	middle finger
D	ring finger	middle finger	ring finger	index finger
D	ring finger	index finger	ring finger	index finger

Appendix B

Table 3. Items in Experiment 2

set	first finger	second finger	third finger	fourth finger
A	index finger	ring finger	middle finger	index finger
A	index finger	middle finger	ring finger	middle finger
A	ring finger	middle finger	index finger	ring finger
A	ring finger	index finger	middle finger	index finger
A	middle finger	index finger	middle finger	ring finger
A	middle finger	index finger	ring finger	index finger
A	middle finger	ring finger	index finger	middle finger
B	index finger	ring finger	middle finger	ring finger
B	index finger	middle finger	ring finger	index finger
B	ring finger	middle finger	index finger	middle finger
B	ring finger	index finger	middle finger	ring finger
B	middle finger	index finger	middle finger	index finger
B	middle finger	ring finger	index finger	ring finger
B	middle finger	index finger	ring finger	middle finger

Appendix C

Table 4. Items in Experiment 3

set	first finger	second finger	third finger	fourth finger
A	middle finger	index finger	pinkie	index finger
A	index finger	pinkie	middle finger	ring finger
A	ring finger	index finger	middle finger	pinkie
A	pinkie	middle finger	pinkie	index finger
B	pinkie	ring finger	index finger	middle finger
B	middle finger	index finger	pinkie	ring finger
B	index finger	pinkie	index finger	middle finger
B	ring finger	middle finger	ring finger	pinkie
C	middle finger	ring finger	middle finger	pinkie
C	pinkie	middle finger	ring finger	index finger
C	ring finger	pinkie	index finger	ring finger
C	index finger	ring finger	pinkie	middle finger
D	ring finger	middle finger	index finger	pinkie
D	index finger	pinkie	middle finger	index finger
D	pinkie	index finger	ring finger	pinkie
D	middle finger	pinkie	ring finger	index finger

Appendix D

Table 5. Items in Experiment 4

set	first finger	second finger	third finger	fourth finger
A	middle finger	index finger	ring finger	index finger
A	index finger	pinkie	middle finger	ring finger
A	ring finger	middle finger	index finger	pinkie
B	pinkie	ring finger	index finger	middle finger
B	ring finger	middle finger	ring finger	pinkie
B	middle finger	index finger	pinkie	index finger
C	pinkie	middle finger	ring finger	index finger
C	index finger	ring finger	pinkie	middle finger
C	ring finger	pinkie	index finger	pinkie
D	middle finger	index finger	pinkie	ring finger
D	index finger	pinkie	ring finger	middle finger
D	pinkie	middle finger	index finger	pinkie
E	ring finger	index finger	middle finger	pinkie
E	index finger	ring finger	index finger	middle finger
E	middle finger	pinkie	ring finger	index finger
F	ring finger	pinkie	index finger	ring finger
F	middle finger	ring finger	middle finger	pinkie
F	pinkie	index finger	ring finger	index finger

References

- Arnold, K. M., & McDermott, K. B. (2013). Test-potentiated learning: Distinguishing between direct and indirect effects of tests. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*, 940–945. DOI: <https://doi.org/10.1037/a0029199>
- Bäumel, K.-H., Hanslmayr, S., Pastötter, B., & Klimesch, W. (2008). Oscillatory correlates of intentional updating in episodic memory. *NeuroImage*, *41*, 596–604. DOI: <https://doi.org/10.1016/j.neuroimage.2008.02.053>
- Bjork, R. A. (1989). Retrieval inhibition as an adaptive mechanism in human memory. In H. L. Roediger III & F. I. M. Craik (Eds.), *Varieties of memory and consciousness: Essays in honour of Endel Tulving* (pp. 309–330). Hillsdale: Erlbaum.
- Chaffin, R., Lisboa, T., Logan, T., & Begosh, K. T. (2010). Preparing for memorized cello performance: The role of performance cues. *Psychology of Music*, *38*, 3–30. DOI: <https://doi.org/10.1177/0305735608100377>
- Chan, J. C. K., Meissner, C. A., & Davis, S. D. (2018). Retrieval potentiates new learning: A theoretical and meta-analytic review. *Psychological Bulletin*, *144*, 1111–1146. DOI: <https://doi.org/10.1037/bul0000166>
- Curran, T., & Keele, S. W. (1993). Attentional and nonattentional forms of sequence learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*, 189–202. <https://doi.org/10.1037/0278-7393.19.1.189>
- Destrebecqz, A., & Cleeremans, A. (2001). Can sequence learning be implicit? New evidence with the process dissociation procedure. *Psychonomic Bulletin & Review*, *8*, 343–350. DOI: <https://doi.org/10.3758/BF03196171>
- 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. DOI: <https://doi.org/10.1162/08992998563130>
- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). GPower 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*, 175–191. DOI: <https://doi.org/10.3758/BF03193146>
- Feltz, D. L., & Landers, D. M. (1983). The effects of mental practice on motor skill learning and performance: A meta-analysis. *Journal of Sport Psychology*, *5*, 25–57.
- Hanslmayr, S., Volberg, G., Wimber, M., Oehler, N., Staudigl, T., Hartmann, T., Raabe, M., Greenlee, M. W., & Bäumel, K.-H. T. (2012). Prefrontally driven down-regulation of neural synchrony mediates goal-directed forgetting. *Journal of Neuroscience*, *32*, 14742–14751. DOI: <https://doi.org/10.1523/JNEUROSCI.1777-12.2012>
- Hoffmann, J., & Koch, I. (1997). Stimulus–response compatibility and sequential learning in the serial reaction time task. *Psychological Research*, *60*, 87–97. doi:<https://doi.org/10.1007/BF00419682>
- Irtel, H. (2007). PXLab: The Psychological Experiments Laboratory [Version 2.1.11]. Mannheim: University of Mannheim. Retrieved from <http://www.pxlab.de>
- Jiménez, L., Vaquero, J.M.M., & Lupiáñez, J. (2006). Qualitative differences between implicit and explicit sequence learning. *Journal of Experimental Psychology: Learning, Memory & Cognition*, *32*, 475–490. DOI: <https://doi.org/10.1037/0278-7393.32.3.475>
- Karpicke, J. D., & Blunt, J. R. (2011). Retrieval practice produces more learning than elaborative studying with concept mapping. *Science*, *331*, 772–775. DOI: <https://doi.org/10.1126/science.1199327>

- Landin, D., & Herbert, E. P. (1997). A comparison of three practice schedules along the contextual interference continuum. *Research Quarterly for Exercise and Sport*, 68, 357–361. DOI: <https://doi.org/10.1080/02701367.1997.10608017>
- Monsell, S. (2003). Task switching. *Trends in Cognitive Sciences*, 7, 134–140. DOI: [https://doi.org/10.1016/S1364-6613\(03\)00028-7](https://doi.org/10.1016/S1364-6613(03)00028-7)
- Nissen, M.J., & Bullemer, P. (1987). Attentional requirements of learning: Evidence from performance measures. *Cognitive Psychology*, 19, 1–32. DOI: [https://doi.org/10.1016/0010-0285\(87\)90002-8](https://doi.org/10.1016/0010-0285(87)90002-8)
- Pastötter, B., & Bäuml, K.-H. T. (2010). Amount of postcue encoding predicts amount of directed forgetting. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 36, 54–65. DOI: <https://doi.org/10.1037/a0017406>
- Pastötter, B., & Bäuml, K.-H. T. (2014). Retrieval practice enhances new learning: the forward effect of testing. *Frontiers in Psychology*, 5, 286. <https://doi.org/10.3389/fpsyg.2014.00286>
- Pastötter, B., Schicker, S., Niedernhuber, J., & Bäuml, K.-H. T. (2011). Retrieval during learning facilitates subsequent memory encoding. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37, 287–297. DOI: <https://doi.org/10.1037/a0021801>
- Pastötter, B., Tempel, T., & Bäuml, K.-H. (2017). Long-term memory updating: the reset-of-encoding hypothesis in list-method directed forgetting. *Frontiers in Psychology*, 8, 2067. doi: <https://doi.org/10.3389/fpsyg.2017.02076>
- Pastötter, B., Weber, J., & Bäuml, K.-H. T. (2013). Using testing to improve learning after severe traumatic brain injury. *Neuropsychology*, 27, 280–285. DOI: <https://doi.org/10.1037/a0031797>
- Roediger, H. L. III, & Gallo, D. A. (2002). Levels of processing: Some unanswered questions. In M. Naveh-Benjamin, M. Moscovitch, & H. L. Roediger III (Eds.), *Perspectives on human memory and cognitive aging: Essays in honour of Fergus Craik* (pp. 28–47). New York: Psychology Press.
- Roediger, H. L. III, Agarwal, P. K., McDaniel, M. A., & McDermott, K. B. (2011). Test-enhanced learning in the classroom: Long-term improvements from quizzing. *Journal of Experimental Psychology: Applied*, 17, 382–395. <https://doi.org/10.1037/a0026252>
- Sahakyan, L., & Delaney, P. F. (2003). Can encoding differences explain the benefits of directed forgetting in the list-method paradigm? *Journal of Memory and Language*, 48, 195–201. DOI: [https://doi.org/10.1016/S0749-596X\(02\)00524-7](https://doi.org/10.1016/S0749-596X(02)00524-7)
- Shanks, D.R., & Perruchet, P. (2002). Dissociation between priming and recognition in the expression of sequential knowledge. *Psychonomic Bulletin & Review*, 9, 362–367. DOI: <https://doi.org/10.3758/BF03196294>
- Symons, C. S., & Johnson, B. T. (1997). The self-reference effect in memory: A meta-analysis. *Psychological Bulletin*, 121, 371–394. DOI: <https://doi.org/10.1037/0033-2909.121.3.371>
- Szpunar, K. K., Khan, N. Y., & Schacter, D. L. (2013). Interpolated memory tests reduce mind wandering and improve learning of on-line lectures. *Proceedings of the National Academy of Sciences*, 110, 6313–6317. DOI: <https://doi.org/10.1073/pnas.1221764110>
- Szpunar, K. K., McDermott, K. B., & Roediger, H. L., III (2008). Testing during study insulates against the buildup of proactive interference. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34, 1392–1399. DOI: <https://doi.org/10.1037/a0013082>
- Tempel, T., Aslan, A., & Frings, C. (2016). Competition dependence of retrieval-induced forgetting in motor memory. *Memory & Cognition*, 44 (4), 671–680. DOI: <https://doi.org/10.3758/s13421-015-0578-3>
- Tempel, T., & Frings, C. (2013). Resolving interference between body movements: Retrieval-induced forgetting of motor sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 39, 1152–1161. DOI: <https://doi.org/10.1037/a0030336>
- Tempel, T., & Frings, C. (2014a). Interference within hands: Retrieval-induced forgetting of left and right hand movements. *Acta Psychologica*, 148, 1–5. DOI: <https://doi.org/10.1016/j.actpsy.2014.01.003>
- Tempel, T., & Frings, C. (2014b). Forgetting motor programmes: Retrieval dynamics in procedural memory. *Memory*, 22, 1116–1125. DOI: <https://doi.org/10.1080/09658211.2013.871293>
- Tempel, T., & Frings, C. (2015). Categorization by movement direction: Retrieval-induced forgetting of motor sequences grouped by motion features. *Quarterly Journal of Experimental Psychology*, 68, 473–486. DOI: <https://doi.org/10.1080/17470218.2014.945098>
- Tempel, T., & Frings, C. (2016a). How motor practice shapes memory: Retrieval but not extra study can cause forgetting. *Memory*, 24(7), 903–915. DOI: <https://doi.org/10.1080/09658211.2015.1059858>
- Tempel, T., & Frings, C. (2016b). Directed forgetting benefits motor sequence encoding. *Memory & Cognition*, 44 (3), 413–419. DOI: <https://doi.org/10.3758/s13421-015-0565-8>
- Tempel, T., & Frings, C. (2017). Retrieval-induced forgetting is retrieval-modality specific: Evidence from motor memory. *Cognition*, 162, 143–152. DOI: <https://doi.org/10.1016/j.cognition.2017.02.005>
- Tempel, T., & Kubik, V. (2017). Test-potentiated learning of motor sequences. *Memory*, 25, 326–334. DOI: <https://doi.org/10.1080/09658211.2016.1171880>
- Tempel, T., Loran, I., & Frings, C. (2015). Dancing your moves away: How memory retrieval shapes complex motor action. *Journal of Experimental Psychology: Applied*, 21, 300–312. DOI: <https://doi.org/10.1037/xap0000052>
- Unsworth, N., & Engle, R., W. (2005). Individual differences in working memory capacity and learning: Evidence from the serial reaction time task. *Memory & Cognition*, 2005, 33, 213–220. DOI: <https://doi.org/10.3758/BF03195310>
- Weinstein, Y., Gilmore, A. W., Szpunar, K. K., & McDermott, K. B. (2014). The role of test expectancy in the build-up of proactive interference in long-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 40, 1039–1048. <https://doi.org/10.1037/a0036164>
- Weinstein, Y., McDermott, K. B., & Szpunar, K. K. (2011). Testing protects against proactive interference in face-name learning. *Psychonomic Bulletin & Review*, 18, 518–523. DOI: <https://doi.org/10.3758/s13423-011-0085-x>
- Wissman, K. T., Rawson, K. A., & Pyc, M. A. (2011). The interim test effect: Testing prior material can facilitate the learning of new material. *Psychonomic Bulletin & Review*, 18, 1140–1147. DOI: <https://doi.org/10.3758/s13423-011-0140-7>
- Yang, C., Potts, R., & Shanks, D. (2017). The Forward Testing Effect on Self-Regulated Study Time Allocation and Metamemory Monitoring. *Journal of Experimental Psychology: Applied*. DOI: <https://doi.org/10.1037/xap0000122>
- Yang, C., Potts, R., and Shanks, D. R. (2018). Enhancing learning and retrieval of new information: a review of the forward testing effect. *Science of Learning*, 3, 8. DOI: <https://doi.org/10.1038/s41539-018-0024-y>

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