Memory skill: The proceduralization of declarative memory through retrieval practice

Mihály Racsmány & Ágnes Szőllősi

Learning and Memory Disorders Research Group; Institute of Cognitive Neuroscience and Psychology, Research Centre for Natural Sciences & Department of Cognitive Science, Budapest University of Technology and Economics

Two outstanding achievements of experimental memory research over the past fifty years have been the description of multiple memory systems and the demonstration that recall is the most effective form of long-term learning. The most important contribution of the former was to show that so-called declarative forms of memory can be described with different psychological characteristics and neurological background mechanisms than procedural memory, which plays a fundamental role in skill learning. Perhaps the most successful research trend in recent years has pointed out that memory recall, a declarative test, fundamentally changes memory representation and its long-term accessibility. In this chapter, we try to combine these two research fields and we aim to present the results of behavioral and brain research that support the assumption that declarative recall causes changes in retrieved memory content that are best understood within the mechanisms of the procedural memory system.

Introduction

In the next chapter, we summarize the pedagogical implications of the experimental psychological phenomenon called the testing effect, and show that the test has proven to be an effective long-term learning method in both laboratory experiments and field studies. We show that it is effective in relation to many curricula (e.g. foreign language learning, acquisition of scientific concepts, scientific knowledge, etc.) and in groups with varied abilities and educational background. The test not only helps to acquire the given curriculum, but also provides a transfer for the acquisition of a similar level of knowledge with a similar structure. We show that the test is not only a more effective learning method as opposed to simple re-learning, but also compared to popular elaborative learning methods such as concept mapping widely used in open educational systems. We then present our experiments that demonstrate that the test proceduralizes learning and memory in a manner similar to that of skill learning widely used in efficient training programs of closed systems. That is, test-based learning methods must take into account the aspects and boundary conditions explored in skill learning trainings in order to be effective enough.

Translational educational science

Perhaps, one of the greatest interdisciplinary challenges for education is how to translate the results of cognitive psychological research in the topic of learning and memory into the practice of everyday education. Roediger (2013) articulated the importance of this issue and stressed the necessity to develop a so-called "translational educational science" that would be capable of translating basic research findings related to memory processes into educational practices and techniques.

Strongly related to this topic, Csaba Pléh and colleagues examined the classical and new relationships between pedagogy and psychology in two comprehensive studies (Pléh, 2013, Pléh & Faragó, 2016). Following Rothkopf (2008), they drew attention to the fact that conflicts over the perception and operation of pedagogical institutions – called provocatively *open systems* (such as the common schools, universities, and many commercial vocational schools), *arcade products* (such as books, computer programs, and videos) and *closed systems* (such as corporate or military training) – can only be achieved through applying psychological scientific models of memory and learning. According to Rothkopf (2008), only closed systems are strongly dominated by concern about efficiencies. Open systems aim to be consistent with cultural expectations, whereas arcade products include attractiveness and the avoidance of egregious errors. As Pléh and Faragó (2016) highlighted, the great pedagogical-psychological dilemma of today is that many forces try to make the open system closed or to be arcade products in the name of efficiency. Moments of arcade products appear in the open system (let the university be full of joy), however, due to the lack of criteria for arcade products, it is difficult to lead anywhere. A clear criterion analysis of closed systems is difficult to implement in the open system as a whole. Pléh (2013) concluded that from a pedagogical point of view, one of the most significant conceptual innovations in cognitive research has been the emergence of increasingly sophisticated models of multiple memory systems. As Pléh (2017) pointed out, this could have been accomplished through the synthesis of old philosophical and pedagogical principles.

Following the thoughts of Aristotle and Plato, Ryle (1949) proposed a distinction between declarative ("knowing that") and procedural ("knowing how") systems. Perhaps the best-known modern version of this theory was developed by Squire and colleagues (see e.g., Squire 1987; Squire, Knowlton, & Musen, 1993) who made a distinction between procedural (skills and abilities) and declarative (episodic and semantic) forms of learning and memory. The key factor here, according to Pléh (2017), would be to incorporate procedural learning forms and trainings used in closed systems and arcade products into educational programs and to build complex semantic knowledge in open systems.

The core properties of automaticity and skill learning

Skill acquisition is suggested to be based on the process of automatization that refers to the progress when one forms a direct association between a stimulus (cue) and a response (for overviews, see Anderson, 1992; Logan, 1985, 1988a, 1988b, 2005; Schneider & Shiffrin, 1977; Shiffrin & Schneider, 1977). Automatization can be described along a set of properties that makes it distinguishable from other types of learning. First of all, automatic behaviours do not require (extra) attentional resources (Hasher & Zacks, 1979; Schneider, Dumais, & Shiffrin, 1984). Since automatic behaviour is not limited by attentional capacities, these processes are fast (see Logan, 1988a; Newell & Rosenbloom, 1981). In fact, there is a consensus among theories of skill learning that response time at learning (or training) follows a nonlinear function (e.g., Logan, 1988a; Josephs, Silvera, & Giesler, 1996; Logan, 1988a; Newell & Rosenbloom, 1981; Palmeri, 1997; Thorndike, 1913). Among others, the "power law of practice" represents the quantitative properties of automatization that became a benchmark test for studies investigating automatization in terms of response speed (Logan, 1988a; Newell & Rosenbloom, 1981). Because there is no need for attentional capacities during automatic behaviour, they require no effort (Hasher & Zacks, 1979; Schneider & Shiffrin, 1977). In other words, one core feature of automatization is the reduction of processing load during learning. Due to the fact that automatic behaviour is (unconsciously) triggered whenever the stimulus is given, performance in tasks assessing automaticity is not reduced in dual task situations (see Logan, 1992, for an overview). In contrast, controlled processes are capacity limited, effortful, and slow (Schneider & Shiffrin, 1977; Shiffrin & Schneider, 1977).

In fact, several authors propose that automatization is a memory phenomenon meaning that automatic behaviour is based on the direct retrieval of a response following the presentation of a cue (stimulus) (see Logan, 1988, Newell & Rosenbloom, 1981; Schneider, 1985). Importantly, this kind of retrieval can occur as the consequence of the continuous strengthening of stimulus-response connections during training (LeBerge & Samuels, 1974). These latter findings underline the essential role of practice in the process of automatization.

The testing effect

Perhaps the most attention-grabbing experimental results in the field of cognitive psychology research and in the topic of memory and learning in recent decades have been those that have shown that testing or retrieval leads to more robust long-term knowledge than any other form of declarative learning, a phenomena called the testing-effect (see Roediger, Putnam, & Smith, 2011). In a typical experimental procedure of the testing effect phenomenon, following an initial learning phase, participants took part either in a retrieval practice or in a repeated study task with the initially studied items (see e.g., Karpicke, Lehman, & Aue, 2014). The final retrieval of all items could be either a few minutes or days after the practice phase. The most robust finding is that items practiced through retrieval show decreased forgetting rate and that retrieval practice has a long-term benefit relative to study practice (Roediger & Butler, 2011; Roediger & Karpicke, 2006). Specifically, most studies found an interaction between the lengths of delay before final recall (minutes or days) and the form of practice (repeated retrieval or study) and showed a short-term advantage of repeated study and a long-term retrieval practice benefit (Thompson, Wenger, & Bartling, 1978; Wheeler, Ewers, & Buonanno, 2003; but see Karpicke et al., 2014).

Overall, an extensive amount of research has shown that taking a memory test on some learning material can improve long-term retention relative to repeatedly studying the material, a phenomenon known as the testing effect (e.g., Carrier & Pashler, 1992; Roediger & Butler, 2011; Roediger & Karpicke, 2006; Wheeler & Roediger, 1992). Moreover, knowledge acquired by retrieval practice is more resistant to interference effects (Kliegl & Bauml, 2016; Racsmány & Keresztes, 2015; Szpunar, McDermott, & Roediger, 2008; but see Siler and Benjamin [in press] for evidence that under certain conditions, testing does not appear to reduce forgetting, but it is a potent means of enhancing inference). Additionally, retrieval practice leads to the better organization of the acquired knowledge, enhances transfer of information to new contexts, and produces faster access to learned information (Jacoby, Walheim, & Coane, 2010; Racsmány, Szőllősi, & Bencze, 2018; Zaromb & Roediger, 2010). Altogether these characteristics of retrieval-based learning make test a potential powerful tool for improving learning in everyday educational practice (Dunlosky, Rawson, Marsh, Nathan, & Willingham, 2013; McDermott, Agarwal, D’Antonio, Roediger, & McDaniel, 2014; Roediger et al., 2011).

Furthermore, retrieval practice is more effective in terms of long-term memory retention than restudy even when repeated study involves the elaborative encoding of the material (Karpicke & Blunt, 2011; Karpicke & Smith, 2012). Naturally, many learning strategies have been developed in the past decades other than restudy (or re-learning) and test practice. Based on the empirical findings, however, it seems that retrieval practice is more beneficial than various types of learning/practice strategies that have spread in educational practice. One of the most popular learning techniques nowadays is the so-called concept mapping strategy. In this method, individuals construct nodes and links that represent concepts and relations among these concepts, respectively. Based on recent empirical findings, retrieval practice is even more effective than concept mapping (Karpicke & Blunt, 2011). In a review article, Dunlosky and colleagues (2013) compared the efficiency of different learning techniques including various types of retrieval practice (e.g., distributed and interleaved retrieval) and other frequently used learning strategies. The authors suggest that some techniques have very low success rate, such as when students are instructed to mentally imagine the contents of texts (see Leutner, Leopold, & Sumfleth, 2009) or to highlight/underline different parts of verbal materials. Other techniques seem to be more beneficial, such as when students are asked to generate explanations for facts (see e.g., Pressley, McDaniel, Turnure, Wood, & Ahmad, 1987) or for some aspects of processing/problem solving during learning (see e.g., Schworm & Renkl, 2006). The authors finally concluded that using memory tests is the most effective way, especially when the tests are separated by delays (Cepeda, Vul, Rohrer, Wixted, & Pashler, 2008; Rawson & Dunlosky, 2012).

Prominent and partly conflicting accounts of the testing effect proposed that high strengths of memory traces are due to the effort required to retrieve a specific item, or assumed that every act of retrieval adds new semantically appropriate cues or new temporal/contextual features to retrieved memories (Carpenter, 2009; Karpicke et al., 2014; Kornell, Bjork, & Garcia, 2011). However, the long-term benefit of retrieval-based learning is still without unanimous explanation.

Similarities between skill learning and retrieval-enhanced learning

It leads to an interesting observation when one compares the experimental literatures of skill-learning and the testing effect. Specifically, information acquired through repeated retrieval is characterized with different attributes in comparison with repeatedly studied information. For instance, information learned by repeated retrieval is (1) more resistant to interference effects (Racsmány & Keresztes, 2015; Szpunar et al., 2008), (2) shows a lower forgetting rate following weeks or months (Roediger & Butler, 2011; Roediger & Karpicke, 2006), and (3) remains accessible in multitasking situations where attentional processes are heavily loaded (Mulligan & Picklesimer, 2016). Compellingly, the above properties of retrieval-based learning are also the characteristics of skill learning (Kuhl, Dudukovic, Kahn, & Wagner, 2007; Newell & Rosenbloom, 1981; Schneider & Chein, 2003; Squire & Zola, 1996).

Moreover, it is also known that spaced initial retrieval practice produces greater memory benefits than massed initial retrieval (Jacoby, 1978; Whitten & Bjork, 1977). Similarly, distributed practice in simple and complex skill learning is also superior to massed learning (Lee & Genovese, 1988). Another finding is that retrieval practice with relatively infrequent and weak retrieval cues produces better memory than the same practice with strong and frequent cues (Carpenter, 2009). Similarly, it was found that reduced frequency of knowledge of results enhances motor skill learning (Winstein & Schmidt, 1990). The number of practice trials is important in both retrieval-based learning and skill learning, as it was found that one or two practice trials are less beneficial than a higher amount of repeated practice (Hanawalt, 1937; Logan, 1988a).

Along with these similarities, two recent functional magnetic resonance imaging (fMRI) studies drew attention to another dependent variable that should be relevant to all accounts of human memory (Ratcliff & McKoon, 2000), namely, the reaction time of retrieval (Keresztes, Kaiser, Kovács, & Racsmány, 2014; van den Broek, Takashima, Segers, Fernández, & Verhoeven, 2013). Both of these studies found that retested items were retrieved faster than restudied ones both 20 minutes (Keresztes et al., 2014) and 7 days (Keresztes et al., 2014; van den Broek et al., 2013) following practice. These results were in line with previous studies that found decreased retrieval latencies during selected retrieval practice (Keresztes & Racsmány, 2013; Román, Soriano, Gómez-Ariza, & Bajo, 2009), retrieval of semantic facts (Pirolli & Anderson, 1985) and list recall (Lehman, Smith, & Karpicke, 2014).

Moreover, van den Broek and colleagues (2013) scanned participants during retrieval/restudy practice and found increased activity in the striatal cortex, in the thalamus, and in the associative cortex during retrieval practice, a finding resembling patterns typically observed in skill learning studies (Raichle, Fiez, Videen, MacLeod, Pardo, Fox, & Petersen, 1994). In another study, Keresztes and colleagues (2014) scanned participants during a cued-recall task either twenty minutes or 1 week following retest/restudy practice, and found decreased control network activity for retested items in comparison with restudied items, with no change in activation level after a 7-day delay. These results suggest that retrieval practice produces faster retrieval at the final test, and that this reduction in response speed is associated with increased basal ganglia and decreased control network activities, again, a typical finding in skill learning literature (Hikosaka, Nakamura, Sakai, & Nakahara, 2002; Kuhl et al., 2007; Newell & Rosenbloom, 1981; Schneider & Chein, 2003). Altogether, these attributes of test-enhanced learning could point to the hypothesis that while retrieval practice is a declarative learning strategy, it shows a similar automatization pattern for item retrieval that is usually observed in skill learning (Squire & Zola, 1996).

The automatization of retrieval during practice: Evidence from response latency, psychophysiology, and brain stimulation

Learning involves processes when one forms associations between pieces of information (e.g., associating a name to a face) among a variety of other processes. Therefore, laboratory studies prefer to use paired associates as stimuli in learning paradigms (see Calkins, 1984). In a set of three experiments we also used paired associates (word pairs) as study materials to investigate response latency during retrieval practice and its relationship with long-term memory retention (Racsmány et al., 2018). Following the tradition of previous studies (see e.g., Karpicke & Roediger, 2008; Wheeler et al., 2003), after initial (intentional) learning, participants practiced the word pairs either by repeated study or by retrieval practice in six subsequent blocks. Former refers to the repeated representations of the study material (restudy condition), whereas latter refers to a cued recall task when subjects were instructed to recall the target words in response to the cue words (retest condition). Following either a short or a longer time delay (few minutes and one week, respectively), memory for all word pairs was tested. The task is illustrated in Figure 1a.

In line with the findings of several former studies (for an overview, see Roediger & Karpicke, 2006), results showed better long-term memory for the tested word pairs relative to the recall of the restudied items – indicating a significant testing effect. Beside this replication, results revealed that response latency during retrieval practice followed a power function. Specifically, reduction in response speed showed a substantial gain in the early phase of practice (see Figure 1b). Most importantly, the measure of goodness of fit to this power function was associated with short- and long-term final recall success suggesting a positive relationship between the level of automatization and memory retention. Furthermore, at final recall, response latency was longer in the restudy than it was in the retest condition (this finding was later replicated by others, e.g., Kubik, Jonsson, Knopf, & Mack, 2018; Marián, Szőllősi, & Racsmány, 2018), indicating that the accessibility of tested elements was easier and probably automatic. Altogether these results point to the conclusion that the automatization of retrieval is a core component of the testing effect phenomenon and lay the basis of a new explanatory framework to better understand the nature of retrieval-based learning and its long-term efficiency.

Neil Mulligan and Daniel Peterson also pointed out that retrieval practice strengthens the associations between targets and cues (Mulligan & Peterson, 2013). This finding is consistent with our results showing that test practice decreases response latency at the retrieval of cue-target associations (Racsmány et al., 2018). However, the authors highlight that while retrieval strengthens cue-target relations, it disrupts the processing of inter-item relational information (see also Mulligan & Peterson, 2015a, 2015b). Consequently, if final recall needs to be based on inter-item relationships, retrieval practice leads to lower memory performance as compared to restudy practice (a phenomenon termed the negative testing effect). Therefore, it should be mentioned that under special circumstances repeated study is a more effective learning strategy as compared to test practice.

Take into account that the process of automatization can be described along a set of core properties (other than with the reduction of response speed), the question arises whether there are further similarities between skill acquisition and retrieval-based learning. One crucial property of automatization is the reduction of processing load during practice (for an overview, see Logan, 1988a). Accordingly, Neil Mulligan and Milton Picklesimer (2016) reported that recall performance was reduced by divided attention after restudy practice but not after retest practice indicating that less attention during a memory test does not reduce the benefit of retrieval on subsequent recall (for similar findings, see Gaspelin, Ruthruff, & Pashler, 2013). This study, however, did not directly measure the processing load of recall during practice. While dual task experimental designs (such as in the study of Mulligan & Picklesimer, 2016) give an opportunity to indirectly estimate processing load, the detection of changes in pupil diameter (pupillometry) is an online measure of processing requirements. Specifically, an increase in task-evoked pupillary response (TEPR) is thought to be associated with higher processing load (Hess & Polt, 1964; Kahneman, 1973; Unsworth & Robison, 2015). Therefore, in two experiments, we used pupillometry to examine processing load during retrieval practice and its relationship with subsequent recall performance (Pajkossy, Szőllősi, & Racsmány, 2019). As expected, TEPR decreased as a function of repeated retrieval cycles during practice (see Figure 1c). Additionally, TEPR was smaller at final recall in the retest condition than it was in the restudy condition after both a brief delay of minutes and a relatively long delay of one week. In sum, retrieval reduced processing load and the involvement of the attentional control network at practice and at later recall.

Strongly related to processing load, one important neuroanatomical structure that plays a key role in effortful attentional control processes is the dorsolateral prefrontal cortex (DLPF)

(Barbey, Koenigs, & Grafman, 2013; Curtis & D'Esposito, 2003). Recent research has shown that the initial retrieval of memories was associated with an increase in the activity of the DLPFC indicating that initial memory retrieval requires effortful attentional control processes. Furthermore, following initial recall, subsequent repeated retrieval was associated with a gradual decrease in DLPFC activity and this decrease was related to better subsequent memory performance (Karlsson Wirebring, Wiklund-Hörnqvist, Eriksson, Andersson, Jonsson, & Nyberg, 2015). Consequently, it seems plausible that the maintenance of DLPFC activity during repeated retrieval should have a negative impact on memory retention. To test this assumption, we used a brain stimulation technique (transcranial Direct Current Stimulation, tDCS) (Marián et al., 2018). At first, participants studied a list of word pairs, such as in our previous experiments listed above. As tDCS is suggested to exert its effect not only at stimulation but also after that (Nitsche & Paulus, 2003), the anodal (excitatory) stimulation of the DLPF occurred before a practice phase while subjects practiced the study material by either repeated study or cued recall. Our findings revealed that the anodal stimulation of the DLPFC disrupted the long-term retention of memories one week after the practice phase (as compared to a control condition where no stimulation occurred).

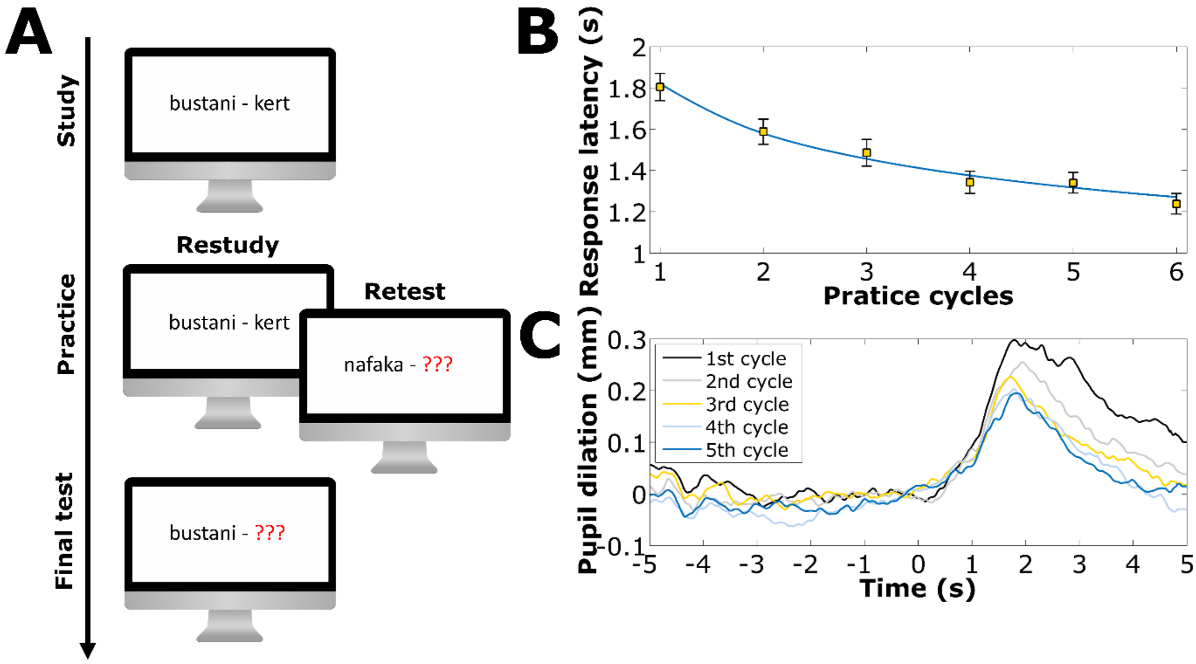

Figure 1. (A) The procedure of the memory task developed to investigate the testing effect phenomenon. Following the initial study of paired associates (Swahili-Hungarian word pairs) participants practice the material with either repeated study (restudy) or cued recall (retest) in a number of subsequent cycles. Finally, memory for all word pairs is tested where participants are required to recall the target items (Hungarian words) in response to the cues (Swahili words). (B) Response latency during retrieval (retest) practice (Racsmány et al., 2018). The points represent the observed data (means with error bars indicating the standard error of the mean); the line represents the model predicted data ($RL = a + bN$^-c, where RL: response latency, a: minimum time required to retrieve the target, b: the amount to be learned, N: number of practice cycles, c: learning rate). (C) Task-evoked pupillary response during five retrieval

(retest) practice cycles (Pajkossy et al., 2019). The *x*-axis gives the time relative to the “onset” of retrieval (0 sec). The *y*-axis gives baseline corrected values (means in mm).

Highlights and suggestions for educational practice: The role of feedback and the effect of acute stress

Reviewing the most important results of our experiments listed above, it has been demonstrated that response latency (Racsmány et al., 2018) and pupil diameter (Pajkossy et al., 2019) showed a gradual decrease during retrieval practice and that the maintenance of attentional control activity prevented memories from automatization and led to lower long-term recall performance (Marián et al., 2018). Based on these observations, we propose that individuals benefit from those learning strategies that allow the automatization of cue-target associations and retrieval. Importantly, there are additional useful suggestions for educational practice on how to further improve students’ learning efficiency and study skills. For example, a relatively short delay between initial learning and the first retrieval attempt promotes memory retention (Jacoby, 1978). Moreover, wide spacing between retrieval attempts (i.e., practice cycles) further improves memory performance (Karpicke & Bauernschmidt, 2011; Landauer & Bjork, 1978; Pyc & Rawson, 2009). Other authors emphasize the critical role of (elaborative) feedback during practice that enhances the benefit of retrieval on long-term retention. Specifically, it seems to be crucial not only to inform students whether the response is correct or not, but to provide them the correct answers (e.g., Agarwal, Karpicke, Kang, Roediger, & McDermott, 2008).

However, a recent study pointed out that feedback could have a negative impact on the testing effect (Storm, Friedman, Murayama, & Bjork, 2014). In this study of Benjamin Storm and colleagues, subjects completed not only one but six subsequent final tests one week after restudy/retest practice. Crucially, at final recall, participants received feedback after each and every retrieval attempt by getting the right answers. Interestingly, feedback at long-term final recall reversed the testing effect (i.e., subjects showed better memory after restudy practice relative to retest practice) after one single final test cycle (see also Pastötter & Bäuml, 2016). However, if one has a closer look on these findings, memory performance was extremely low on the first final test in this study (18% and 25% in the restudy and test conditions, respectively). Therefore, we aimed to improve subjects’ retrieval success at the beginning of practice and to assess the impact of this experimental manipulation on later recall. For this reason, we applied multiple pre-practice learning trials (Racsmány, Szőllősi, & Marián, 2020). As expected, when participants saw the study material three/six times in the initial study phase (and not only once as in the study of Storm and colleagues), no reversed testing effect was found in the final test trials. Instead, the tested items were better remembered. Thus, we suggest what other authors also stressed (for an overview, see e.g., Karpicke et al., 2014) that students should reach a criterion level at initial learning in order to be successful at retrieval practice.

The New Theory of Disuse is one of the most influential theories in the scientific literature of memory research (Bjork & Bjork, 1992; see also Bjork & Bjork, 2013). This theory makes a distinction between two dimensions (or features) of memories. While storage strength is a measure of learning, retrieval strength is an index of the accessibility of memory representations. This theory predicts an increase in storage strength following successful retrieval attempts and states that without continuous access to learnt information forgetting occurs. Therefore, we can assume that when participants’ recall success improved during practice as a consequence of multiple pre-practice learning trials in our study (Racsmány et al., 2020), it led to an increase in storage strength and this increase in storage strength resulted in better long-term memory retention.

Further corroborating the universality of the testing effect phenomenon, the beneficial influence of retrieval practice on long-term memory retention has been demonstrated outside the laboratory as well, such as in the classic study of Herbert F. Spitzer from the early 20th century who tested more than 3500 students from almost 100 elementary schools (Spitzer, 1939; later replicated on a smaller sample by Sones & Stroud in 1940). Recent studies further strengthened the ecological validity of the testing effect by showing superior memory after retest practice using a variety of test formats (including multiple-choice, short-answer, group, and individual quizzes) (e.g., Agarwal et al., 2008; Cranney, Ahn, McKinnon, Morris, & Watts, 2007; McDermott et al., 2014; for additional examples, see e.g., the meta-analytic review paper of Bangert-Drowns, Kulik, & Kulik, 1991; for an overview, see e.g., Butler, 2018). Also, the positive impact of testing has been shown outside the laboratory as well in college/university courses (Cranney et al., 2009; McDaniel, Anderson, Derbish, & Morrisette, 2007; Vojdanoska, Cranney, & Newell, 2010), high school courses (Nungester & Duchastel, 1982), and elementary school courses (Lipowski, Pyc, Dunlosky, & Rawson, 2014). Relatedly, memory tests seem to be advantageous in terms of memory retention in various age groups including pre-schoolers (Fritz, Morris, Nolan, & Singleton, 2007), elementary school children (Lipowski et al., 2014), high school students (Nungester & Duchastel, 1982), as well as in healthy (Meyer & Logan, 2013; Tse, Balota, & Roediger, 2010) and pathological ageing (Camp, Foss, O'Hanlon, & Stevens, 1996). Furthermore, although a long line of research investigated the testing effect phenomenon in laboratory settings using associative learning paradigms, the advantageous effect of testing is present for a wide range of stimulus types, such as for word lists (Wheeler et al., 2003), short passages from textbooks (Rawson & Dunlosky, 2011), and articles (Spitzer, 1939).

One further crucial aspect of testing is that it enhances the transfer of learning (for overviews, see Carpenter, 2012; Karpicke, 2017), for example the transfer across different test formats. In other words, test enhances memory performance even when the types of retrieval (e.g., free recall, cued recall, recognition) differ between practice and the final test (e.g., Glover, 1989; Kang, McDermott, & Roediger, 2007; Karpicke & Blunt, 2011). However, it seems that recall is a more beneficial practice strategy as compared to various types of recognition memory tasks (Kang et al., 2007). Based on these findings, memory researchers suggest to use open-ended questions in educational practice instead of multiple-choice questions also because students are encountered with incorrect responses as well when they complete a recognition memory test. Additionally, retrieval practice enhances transfer across different domains of knowledge. That is, when the target information is tested in the form of retrieval practice, it leads to better memory for related materials that were never tested (Chan, McDermott, & Roediger, 2006). This facilitative effect is not present following restudy practice with no opportunity of testing. Relatedly, test has a beneficial influence on the retrieval of factual knowledge for the tested material and also when individuals respond to conceptual questions that needs the abstraction of concepts from different parts of a text (Butler, 2010). Several furthers studies have shown that the initial retrieval of a study material helps to respond problem-solving and inference questions at final tests (see Blunt & Karpicke, 2014; Karpicke & Blunt, 2011). Finally, testing promotes the acquisition and application of rules for novel materials never tested before (Kang, McDaniel, & Pashler, 2011).

One substantial difference between laboratory studies and educational practice, however, that typical testing situations in everyday life (e.g., exams, job interviews, competitions) are stressful experiences. In these situations, individuals tend to be motivated to perform well despite the considerable stress they experience. For that reason alone, it seems to be a thought provoking finding that the testing effect can be eliminated when practice occurs under high pressure (Hinze & Rapp, 2014). Thus, it seems to be also essential to find out whether retrieval-based learning is an effective learning strategy if final recall occurs under stress (as in an exam, by way of only one example). To answer this question, we conducted an experiment where participants were exposed to a psychosocially stressful situation immediately before the final recall of paired associates. Our results showed better memory at final recall after retest practice (when compared to restudy practice) despite subjects' enhanced acute stress levels, as measured by salivary cortisol levels, alpha-amylase activity as well as subjective ratings of negative affective states (Szőllősi, Keresztes, Novák, Szászi, Kéri, & Racsmány, 2017; but see Smith, Floerke, & Thomas, 2016). Based on these findings we believe that retrieval-based learning is an especially efficient learning strategy even if individuals have to/want to perform well in stressful situations (see also Pastötter, von Dawans, Domes, & Frings, 2020).

Conclusions

In light of all these results, it seems that retrieval-enhanced learning can be an educational tool in the hands of open systems that can achieve similar efficiencies in building complex declarative knowledge as that achieved by closed systems and arcade products. Educational programs built on the principle of retrieval-enhanced learning can result in long-term knowledge that is resistant to stress and interference with more and more learning, while facilitating the transfer of new knowledge.

Acknowledgment

This work was supported by the 2017-1.2.1-NKP2017-00002 Research Grant (Hungarian Brain Research Program) and by the NKFI (National Research, Development and Innovation Office, Hungary) K124098 Research Grant.

References

Agarwal, P. K., Karpicke, J. D., Kang, S. H., Roediger, H. L., & McDermott, K. B. (2008). Examining the testing effect with open-and closed-book tests. *Applied Cognitive Psychology, 22*, 861-876.

Anderson, J. R. (1992). Automaticity and the ACT theory. *The American Journal of Psychology, 105*, 165-180.

Bangert-Drowns, R. L., Kulik, J. A., & Kulik, C. L. C. (1991). Effects of frequent classroom testing. *The Journal of Educational Research, 85*, 89-99.

Barbey, A. K., Koenigs, M., & Grafman, J. (2013). Dorsolateral prefrontal contributions to human working memory. *Cortex, 49*, 1195-1205.

Bjork, R. A., & Bjork, E. L. (1992). A new theory of disuse and an old theory of stimulus fluctuation. In A. F. Healy, S. M. Kosslyn, & R. M. Shiffrin (Eds.), *From learning processes to cognitive processes: Essays in honor of William K. Estes* (Vol. 2, pp. 35-67). Hillsdale, NJ, Erlbaum.

Bjork, R. A., & Bjork, E. L. (2013). Optimizing treatment and instruction: Implications of a new theory of disuse. In L. G. Nilsson, & N. Ohta (Eds.), *Memory and society: Psychological perspectives* (pp. 119-143). New York, NY, Psychology Press.

Blunt, J. R., & Karpicke, J. D. (2014). Learning with retrieval-based concept mapping. *Journal of Educational Psychology, 106*, 849-858.

Butler, A. C. (2010). Repeated testing produces superior transfer of learning relative to repeated studyings*Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*, 1118-1133.

Butler, A. C. (2018). Multiple-choice testing in education: Are the best practices for assessment also good for learning? *Journal of Applied Research in Memory and Cognition, 7*, 323-331.

Calkins, M. W. (1894). Association: I. *Psychological Review, 1*, 476-483.

Camp, C. J., Foss, J. W., O'Hanlon, A. M., & Stevens, A. B. (1996). Memory interventions for persons with dementia. *Applied Cognitive Psychology, 10*, 193-210.

Carpenter, S. K. (2009). Cue strength as a moderator of the testing effect: The benefits of elaborative retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*, 1563-1569.

Carpenter, S. K. (2012). Testing enhances the transfer of learning. Current Directions in Psychological Science, 21(5), 279-283.

Carrier, M., & Pashler, H. (1992). The influence of retrieval on retention. *Memory and Cognition, 20*, 633-642.

Cepeda, N. J., Vul, E., Rohrer, D., Wixted, J. T., & Pashler, H. (2008). Spacing effects in learning: A temporal ridgeline of optimal retention. *Psychological Science, 19*, 1095-1102.

Chan, J. C., McDermott, K. B., & Roediger, H. L. (2006). Retrieval-induced facilitation: Initially nontested material can benefit from prior testing of related material. *Journal of Experimental Psychology: General, 135*, 553-571.

Cranney, J., Ahn, M., McKinnon, R., Morris, S., & Watts, K. (2009). The testing effect, collaborative learning, and retrieval-induced facilitation in a classroom setting. *European Journal of Cognitive Psychology, 21*, 919-940.

Curtis, C. E., & D'Esposito, M. (2003). Persistent activity in the prefrontal cortex during working memory. *Trends in Cognitive Sciences*, *7,* 415-423.

Duchastel, P. C., & Nungester, R. J. (1982). Testing effects measured with alternate test forms. *The Journal of Educational Research, 75,* 309-313.

Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). Improving students' learning with effective learning techniques: Promising directions from cognitive and educational psychology. *Psychological Science in the Public Interest, 14*, 4-58.

Gaspelin, N., Ruthruff, E., & Pashler, H. (2013). Divided attention: An undesirable difficulty in memory retention. *Memory and Cognition, 41*, 978-988.

Glover, J. A. (1989). The" testing" phenomenon: Not gone but nearly forgotten. *Journal of Educational Psychology, 81,* 392-399.

Hasher, L., & Zacks, R. T. (1979). Automatic and effortful processes in memory. *Journal of Experimental Psychology: General, 108,* 356-388.

Hanawalt, N. G. (1937). Memory trace for figures in recall and recognition. *Archives de Psychologie, 31,* 216-223.

Hess, E. H., & Polt, J. M. (1964). Pupil size in relation to mental activity during simple problem solving. *Science, 143,* 1190-1192.

Hinze, S. R., & Rapp, D. N. (2014). Retrieval (sometimes) enhances learning: Performance pressure reduces the benefits of retrieval practice. *Applied Cognitive Psychology, 28,* 597-606.

Hikosaka, O., Nakamura, K., Sakai, K., & Nakahara, H. (2002). Central mechanisms of motor skill learning. *Current Opinion in Neurobiology*, *12*, 217-222.

Jacoby, L. L. (1978). On interpreting the effects of repetition: Solving a problem versus remembering a solution. *Journal of Verbal Learning and Verbal Behavior, 17,* 649-667.

Jacoby, L. L., Wahlheim, C. N., & Coane, J. H. (2010). Test-enhanced learning of natural concepts: Effects on recognition memory, classification, and metacognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*, 1441-1451.

Josephs, R. A., Silvera, D. H., & Giesler, R. B. (1996). The learning curve as a metacognitive tool. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*, 510-524.

Kahneman, D. (1973). *Attention and effort*. Prentice Hall, Engelwood Cliffs.

Kang, S. H., McDaniel, M. A., & Pashler, H. (2011). Effects of testing on learning of functions. *Psychonomic Bulletin and Review, 18,* 998-1005.

Kang, S. H., McDermott, K. B., & Roediger, H. L. (2007). Test format and corrective feedback modify the effect of testing on long-term retention. *European Journal of Cognitive Psychology, 19*, 528-558.

Karlsson Wirebring, L., Wiklund-Hörnqvist, C., Eriksson, J., Andersson, M., Jonsson, B., & Nyberg, L. (2015). Lesser neural pattern similarity across repeated tests is associated with better long-term memory retention. *Journal of Neuroscience, 35,* 9595-9602.

Karpicke, J. D. (2017). Retrieval-based learning: A decade of progress. J. H. Byrne (Ed.), Learning and memory: A comprehensive reference, 2nd edition (pp. 487-514). Cambridge, MA: Elsevier.

Karpicke, J. D., & Bauernschmidt, A. (2011). Spaced retrieval: absolute spacing enhances learning regardless of relative spacing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37,* 1250-1257.

Karpicke, J. D., & Blunt, J. R. (2011). Retrieval practice produces more learning than elaborative studying with concept mapping. *Science, 331,* 772-775.

Karpicke, J. D., Lehman, M., & Aue, W. R. (2014). Retrieval-based learning: An episodic context account. *Psychology of Learning and Motivation, 61*, 237-284.

Karpicke, J. D., & Roediger, H. L. (2008). The critical importance of retrieval for learning. *Science, 319,* 966-968.

Karpicke, J. D., & Smith, M. A. (2012). Separate mnemonic effects of retrieval practice and elaborative encoding. *Journal of Memory and Language, 67*, 17-29.

Keresztes, A., Kaiser, D., Kovács, G., & Racsmány, M. (2014). Testing promotes long-term learning via stabilizing activation patterns in a large network of brain areas. *Cerebral Cortex, 24*, 3025-3035.

Keresztes, A., & Racsmány, M. (2013). Interference resolution in retrievalinduced forgetting: Behavioral evidence for a nonmonotonic relationship between interference and forgetting. *Memory and Cognition, 41,* 511–518.

Kliegl, O., & Bäuml, K. H. T. (2016). Retrieval practice can insulate items against intralist interference: Evidence from the list-length effect, output interference, and retrieval-induced forgetting. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 42,* 202-214.

Kornell, N., Bjork, R. A., & Garcia, M. A. (2011). Why tests appear to prevent forgetting: A distribution-based bifurcation model. *Journal of Memory and Language, 65,* 85-97.

Kubik, V., Jönsson, F. U., Knopf, M., & Mack, W. (2018). The direct testing effect is pervasive in action memory: Analyses of recall accuracy and recall speed. *Frontiers in Psychology, 9*: 1632.

Kuhl, B. A., Dudukovic, N. M., Kahn, I., & Wagner, A. D. (2007). Decreased demands on cognitive control reveal the neural processing benefits of forgetting. *Nature Neuroscience, 10,* 908-914.

LaBerge, D., & Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology, 6*, 293-323.

Landauer, T. K., & Bjork, R. A. (1978). Optimum rehearsal patterns and name learning. In M. M. Gruneberg, P. E. Morris, & R. N. Skyes (Eds.), *Practical aspects of memory* (pp. 625-632). London, Academic Press.

Lee, T. D., & Genovese, E. D. (1988). Distribution of practice in motor skill acquisition: Learning and performance effects reconsidered. *Research Quarterly for Exercise and Sport, 59,* 277-287.

Lehman, M., Smith, M. A., & Karpicke, J. D. (2014). Toward an episodic context account of retrieval-based learning: Dissociating retrieval practice and elaboration. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40*, 1787-1794.

Leutner, D., Leopold, C., & Sumfleth, E. (2009). Cognitive load and science text comprehension: Effects of drawing and mentally imagining text content. *Computers in Human Behavior, 25*, 284-289.

Lipowski, S. L., Pyc, M. A., Dunlosky, J., & Rawson, K. A. (2014). Establishing and explaining the testing effect in free recall for young children. *Developmental Psychology, 50*, 994-1000.

Logan, G. D. (1985). Skill and automaticity: Relations, implications, and future directions. *Canadian Journal of Psychology, 39,* 367-386.

Logan, G. D. (1988a). Toward an instance theory of automatization. *Psychological Review, 95*, 492-527.

Logan, G. D. (1988b). Automaticity, resources, and memory: Theoretical controversies and practical implications. *Human Factors, 30,* 583-598.

Logan, G. D. (2005). Attention, automaticity, and executive control. In A. F. Healy (Ed.), *Decade of behavior*. *Experimental cognitive psychology and its applications* (pp. 129-139). Washington, DC, American Psychological Association Press.

Marián, M., Szőllősi, Á., & Racsmány, M. (2018). Anodal transcranial direct current stimulation of the right dorsolateral prefrontal cortex impairs long-term retention of reencountered memories. *Cortex, 108*, 80-91.

McDaniel, M. A., Anderson, J. L., Derbish, M. H., & Morrisette, N. (2007). Testing the testing effect in the classroom. *European Journal of Cognitive Psychology, 19*, 494-513.

McDermott, K. B., Agarwal, P. K., D'Antonio, L., Roediger, H. L., & McDaniel, M. A. (2014). Both multiple-choice and short-answer quizzes enhance later exam performance in middle and high school classes. *Journal of Experimental Psychology: Applied, 20*, 3-21.

Meyer, A. N., & Logan, J. M. (2013). Taking the testing effect beyond the college freshman: Benefits for lifelong learning. *Psychology and Aging, 28*, 142-147.

Mulligan, N. W., & Peterson, D. J. (2013). The negative repetition effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39*, 1403-1416.

Mulligan, N. W., & Peterson, D. J. (2015a). Negative and positive testing effects in terms of itemspecific and relational information. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 41*, 859-871.

Mulligan, N. W., & Peterson, D. J. (2015b). The negative testing and negative generation effects are eliminated by delay. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 41*, 1014-1025.

Mulligan, N. W., & Picklesimer, M. (2016). Attention and the testing effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 42*, 938-950.

Newell, A., & Rosenbloom, P. S. (1981). Mechanisms of skill acquisition and the law of practice. J. R. Anderson (Ed.), *Cognitive skills and their acquisition* (pp. 1-55). Hillsdale, NJ, Erlbaum.

Nitsche, M. A., & Paulus, W. (2001). Sustained excitability elevations induced by transcranial DC motor cortex stimulation in humans. *Neurology, 57*, 1899-1901.

Pajkossy, P., Szőllősi, Á., & Racsmány, M. (2019). Retrieval practice decreases processing load of recall: Evidence revealed by pupillometry. *International Journal of Psychophysiology, 143*, 88-95.

Palmeri, T. J. (1997). Exemplar similarity and the development of automaticity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*, 324-354.

Pastötter, B., & Bäuml, K. H. T. (2016). Reversing the testing effect by feedback: Behavioral and electrophysiological evidence. *Cognitive, Affective, and Behavioral Neuroscience, 16*, 473-488.

Pastötter, B., von Dawans, B., Domes, G., & Frings, C. (2020). The forward testing effect is immune to acute psychosocial encoding/retrieval stress. *Experimental Psychology* (in press).

Pirolli, P. L., & Anderson, J. R. (1985). The role of practice in fact retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 11*, 136-153.

Pléh, C. (2013). A pszichológia és a pedagógia klasszikus és új kapcsolatai. *Iskolakultúra, 13*, 76-81. (Title in English: “The classical and new relationships of psychology and pedagogy”)

Pléh, C. (2017). Tudásfajták a kognitív kutatásban, az iskolában és a köznapokban. *Magyar Tudomány, 178*, 1387-1398. (Title in English: “Types of knowledge in cognitive research, education, and everyday life”)

Pléh, C., & Faragó, B. (2016). Az új technológia és a tanulás világa a pszichológus szemével. *Magyar Pszichológiai Szemle, 71*, 3-23. (Title in English: “New technology and the world of learning from the psychologist’s perspective”)

Pressley, M., McDaniel, M. A., Turnure, J. E., Wood, E., & Ahmad, M. (1987). Generation and precision of elaboration: Effects on intentional and incidental learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13*, 291.

Pyc, M. A., & Rawson, K. A. (2009). Testing the retrieval effort hypothesis: Does greater difficulty correctly recalling information lead to higher levels of memory? *Journal of Memory and Language, 60*, 437-447.

Racsmány, M., & Keresztes, A. (2015). Initial retrieval shields against retrieval-induced forgetting. *Frontiers in Psychology, 6*: 657.

Racsmány, M., Szőllősi, Á., & Bencze, D. (2018). Retrieval practice makes procedure from remembering: An automatization account of the testing effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 44*, 157-166.

Racsmány, M., Szőllősi, Á., & Marián, M. (2020). Reversing the testing effect by feedback is a matter of performance criterion at practice. *Memory and Cognition* (in press).

Raichle, M. E., Fiez, J. A., Videen, T. O., MacLeod, A. M. K., Pardo, J. V., Fox, P. T., & Petersen, S. E. (1994). Practice-related changes in human brain functional anatomy during nonmotor learning. *Cerebral Cortex, 4*, 8-26.

Ratcliff, R., & McKoon, G. (2000). Memory models. In E. Tulving & F. I. M. Craik (Eds.), Oxford handbook of memory (pp. 571-582). New York, NY, Oxford University Press.

Rawson, K. A., & Dunlosky, J. (2012). When is practice testing most effective for improving the durability and efficiency of student learning?. *Educational Psychology Review, 24,* 419-435.

Roediger, H. L. (2013). Applying cognitive psychology to education: Translational educational science. *Psychological Science in the Public Interest 14*, 1-3.

Roediger, H. L., & Butler, A. C. (2011). The critical role of retrieval practice in long-term retention. *Trends in Cognitive Sciences, 15*, 20-27.

Roediger, H. L., & Karpicke, J. D. (2006). The power of testing memory: Basic research and implications for educational practice. *Perspectives on Psychological Science, 1*, 181-210.

Roediger, H. L., Putnam, A. L., & Smith, M. A. (2011). Ten benefits of testing and their applications to educational practice. *Psychology of Learning and Motivation, 55*, 1-36.s

Román, P., Soriano, M. F., Gómez-Ariza, C. J., & Bajo, M. T. (2009). Retrieval-induced forgetting and executive control. *Psychological Science*, *20*, 1053-1058.

Rothkopf, E. Z. (2008). An immodest proposal: Pedagogic information supports for teachers. *Teachers College Record, 111*, 164-179.

Ryle, G. (1949). *The concept of mind.* New York, NY, Barnes and Noble.

Schneider, W. (1985). Toward a model of attention and the development of automatic processing. In M. I. Posner, & O. S. Marin (Eds.), *Attention and performance XI* (pp. 475-492). Hillsdale, NJ, Erlbaum.

Schneider, W., & Chein, J. M. (2003). Controlled & automatic processing: Behavior, theory, and biological mechanisms. *Cognitive Science, 27,* 525-559.

Schneider, W., Dumais, S. T., & Shiffrin, R. M. (1984). Automatic and control processing and attention. In R. Parasuraman, & R. Davies (Eds.), *Varieties of attention* (pp. 1-27). New York, NY, Academic Press.

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search and attention. *Psychological Review, 84,* 1-66.

Schworm, S., & Renkl, A. (2006). Computer-supported example-based learning: When instructional explanations reduce self-explanations. *Computers and Education, 46*, 426-445.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review, 84*, 127-190.

Siler, J., & Benjamin, A. S. (2019). Long-term inference and memory following retrieval practice. *Memory and Cognition* (Epub ahead of print).

Smith, A. M., Floerke, V. A., & Thomas, A. K. (2016). Retrieval practice protects memory against acute stress. *Science, 354*, 1046-1048.

Sones, A. M., & Stroud, J. B. (1940). Review, with special reference to temporal position. *Journal of Educational Psychology, 31*, 665-676.

Spitzer, H. F. (1939). Studies in retention. *Journal of Educational Psychology, 30,* 641-656.

Squire, L. R. (1987). *Memory and brain.* New York, NY, Oxford University Press.

Squire, L. R., Knowlton, B., & Musen, G. (1993). The structure and organization of memory. *Annual Review of Psychology, 44,* 453-495.

Squire, L. R., & Zola, S. M. (1996). Structure and function of declarative and nondeclarative memory systems. *Proceedings of the National Academy of Sciences of the United States of America, 93,* 13515-13522

Storm, B. C., Friedman, M. C., Murayama, K., & Bjork, R. A. (2014). On the transfer of prior tests or study events to subsequent study. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40*, 115-124.

Szőllősi, Á., Keresztes, A., Novák, B., Szászi, B., Kéri, S., & Racsmány, M. (2017). The testing effect is preserved in stressful final testing environment. *Applied Cognitive Psychology, 31*, 615-622.

Szpunar, K. K., McDermott, K. B., & Roediger, H. L. (2008). Testing during study insulates against the buildup of proactive interference. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*, 1392-1399.

Thompson, C. P., Wenger, S. K., & Bartling, C. A. (1978). How recall facilitates subsequent recall: A reappraisal. *Journal of Experimental Psychology: Human Learning and Memory, 4*, 210-221.

Thorndike, E. L. (1913). *Educational psychology: The psychology of learning* (Vol. 2). New York, NY, Teachers College Press.

Tse, C. S., Balota, D. A., & Roediger, H. L. (2010). The benefits and costs of repeated testing on the learning of face–name pairs in healthy older adults. *Psychology and Aging, 25*(4), 833-845.

Unsworth, N., & Robison, M. K. (2015). Individual differences in the allocation of attention to items in working memory: Evidence from pupillometry. *Psychonomic Bulletin and Review, 22*, 757-765.

van den Broek, G. S., Takashima, A., Segers, E., Fernández, G., & Verhoeven, L. (2013). Neural correlates of testing effects in vocabulary learning. *NeuroImage, 78*, 94-102.

Vojdanoska, M., Cranney, J., & Newell, B. R. (2010). The testing effect: The role of feedback and collaboration in a tertiary classroom setting. *Applied Cognitive Psychology, 24*, 1183-1195.

Wheeler, M., Ewers, M., & Buonanno, J. (2003). Different rates of forgetting following study versus test trials. *Memory, 11*, 571-580.

Wheeler, M. A., & Roediger, H. L. (1992). Disparate effects of repeated testing: Reconciling Ballard's (1913) and Bartlett's (1932) results. *Psychological Science, 3*, 240-246.

Whitten II, W. B., & Bjork, R. A. (1977). Learning from tests: Effects of spacing. Journal of *Verbal Learning and Verbal Behavior, 16*, 465-478.

Winstein, C. J., & Schmidt, R. A. (1990). Reduced frequency of knowledge of results enhances motor skill learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*, 677-691.

Zaromb, F. M., & Roediger, H. L. (2010). The testing effect in free recall is associated with enhanced organizational processes. *Memory and Cognition, 38*, 995-1008.