Trends in Neuroscience and Education 5 (2016) 52–66

Contents lists available at ScienceDirect

Trends in Neuroscience and Education

journal homepage: www.elsevier.com/locate/tine

Review article

Neurocognitive mechanisms of the "testing effect": A review

Gesa van den Broek [a,*,1], Atsuko Takashima [a,1], Carola Wiklund-Hörnqvist [b,c], Linnea Karlsson Wirebring [b,c,d], Eliane Segers [a], Ludo Verhoeven [a], Lars Nyberg [b,d,e]

[a] Radboud University, Behavioural Science Institute, 6500 HC Nijmegen, The Netherlands
[b] Umeå University, Umeå Center for Functional Brain Imaging (UFBI), 901 87 Umeå, Sweden
[c] Umeå University, Department of Psychology, 901 87 Umeå, Sweden
[d] Umeå University, Department of Integrative Medical Biology, 901 87 Umeå, Sweden
[e] Umeå University, Department of Radiation Sciences, 901 87 Umeå, Sweden

ARTICLE INFO

Article history:
Received 16 October 2015
Received in revised form
27 May 2016
Accepted 30 May 2016
Available online 2 June 2016

Keywords:
Testing effect
Retrieval
Test-potentiated learning
fMRI
Semantic elaboration
Search set restriction

ABSTRACT

Memory retrieval is an active process that can alter the content and accessibility of stored memories. Of potential relevance for educational practice are findings that memory retrieval fosters better retention than mere studying. This so-called *testing effect* has been demonstrated for different materials and populations, but there is limited consensus on the neurocognitive mechanisms involved. In this review, we relate cognitive accounts of the testing effect to findings from recent brain-imaging studies to identify neurocognitive factors that could explain the testing effect. Results indicate that testing facilitates later performance through several processes, including effects on semantic memory representations, the selective strengthening of relevant associations and inhibition of irrelevant associations, as well as potentiation of subsequent learning.

© 2016 Elsevier GmbH. All rights reserved.

Contents

1. Memory retrieval as an active process: the testing effect

Memory is typically viewed as a three-step process that begins with the encoding of information, followed by storage and later retrieval of fixed, stable memories. However, this view is incomplete. Retrieval is not a simple read-out process but an *active* process that can change the content and accessibility of memories [1,2]. Of particular interest for educational practice is that prompting retrieval with practice-tests enhances the retention of to-be-learned information over time, as shown in studies on the so-called *testing effect*:"taking a test enhances later performance on the material relative to rereading it or to having no re-exposure at all" [3, p20]. Surprisingly, given the plethora of empirical studies demonstrating the testing effect (see Box 1), there is still limited knowledge of the specific neurocognitive mechanisms involved. In this review, we relate existing cognitive accounts of the testing

* Corresponding author.
E-mail address: g.vandenbroek@pwo.ru.nl (G. van den Broek).
[1] Equal contributions.

http://dx.doi.org/10.1016/j.tine.2016.05.001
2211-9493/© 2016 Elsevier GmbH. All rights reserved.

Box 1.

Benefits of memory retrieval: a robust phenomenon

The testing effect is a well-investigated phenomenon in cognitive psychology. For a comprehensive review of behavioral studies, readers are referred to literature overviews in [3,4,49]. Here, we provide a brief introduction to the effect to show that its robustness across different populations, study designs and materials makes it relevant for educational practice.

A typical behavioral testing effect study includes a baseline exposure, followed by either a practice-test or further restudying of the materials, and later a final test to measure learning outcomes (see Fig. 1A). For example, in a study by Roediger and Karpicke [5], students read two prose passages which covered scientific topics, and then restudied one passage and took a practice-test of the other. Learning was assessed five minutes, two days or one week later. Restudying led to better immediate results but practice-testing led to better results on the delayed final tests. This is a common finding in testing effect studies, which often show that the benefits of practice-tests are stronger when the final test is given after a delay rather than immediately after practice (for further information see [43,50–53]).

The testing effect has been replicated across different laboratories and also been documented to reliably improve learning outside the laboratory.

(1) ***The testing effect holds in authentic educational settings using course materials***
Studies have demonstrated the testing effect with course materials [54–60] and real university exams [60], using on-line testing [54], in-class testing [55,56], and classroom response systems ('clickers') [57].

(2) ***The testing effect holds when compared to other pedagogical methods and for different materials***
Testing is more beneficial than pedagogical methods such as mind mapping [12] and group discussions [61], and a better tool for self-study than techniques like reading and highlighting text [62]. The effect was documented with different materials, including materials about geography [63], statistics [60], and medical education [64].

(3) ***The testing effect generates transfer of learning***
Testing enhances the transfer of learning from the specific questions from practice to new problems [63,65–67], enhances re-learning of information [68], and results in higher exam scores [57].

(4) ***The testing effect is beneficial for different populations***
The testing effect has been demonstrated in different age groups, ranging from children [57,58,63] to older adults [69]. Recently, an equally sized testing effect has been demonstrated for individuals suffering from severe traumatic brain injury as compared with healthy individuals [70].

effect to findings from recent brain-imaging studies in order to gain a better understanding of the beneficial effects of memory retrieval on the long-term retention of information. In addition to studies on the testing-effect, available neuroimaging data for the closely related phenomenon of test-potentiated encoding will also be discussed.

2. Cognitive processes underlying the testing effect

Different ideas have been put forward regarding the cognitive processes underlying the testing effect [3–5]. Many of these explanations focus on the way in which testing affects memory representations of the to-be-learned materials. Because most studies on testing effects use verbal materials (e.g., vocabulary or word-pairs), these memory representations are typically conceptualized as (parts of) semantic networks, in which activation spreads among related pieces of information [cf. 6]. Testing is thought to enhance the accessibility of target information by changing the connections within semantic networks, for example, between the representations of two words that are encoded as a word pair [7–10].

Broadly speaking, two different theories exist about the nature of changes in semantic networks. On the one hand, elaboration accounts suggest that semantic networks become richer through testing because additional associations and alternative retrieval routes are formed [7,8]. On the other hand, search-set restriction accounts hold that testing reduces the number of associations that are activated in response to retrieval cues because cue-target associations are selectively strengthened and irrelevant representations are suppressed [9,10].

Carpenter et al. introduced the elaboration account of testing based on the assumption that mental elaboration during the search for the correct answer to a test question extends the semantic network of the tested information by creating or strengthening connections with related concepts [7,8]. These changes in semantic associations are thought to facilitate later recall by providing additional retrieval routes. Support for such accounts comes from studies showing that practice-tests[2] enhance not only memory for presented information, but also for related semantic information that learners generate to associate cue and target information. For example, participants who studied word-pairs like *Mother:Child*, showed better target recall ("*Child*") in response to related semantic mediators like "*Father*" after practice-testing (*Mother:* _____) than after restudying (*Mother:Child*) [11]. In short, representations are thought to get increasingly elaborate with practice-testing so that target information can later be activated through different alternative retrieval routes.

The search-set restriction accounts focus more on the selective nature of retrieval processes during testing, in particular, on the way in which the activation and selection of target information among competing (incorrect) responses influence future retrieval. One theory is that cue-target associations become selectively strengthened such that the memory search hones in on target information while competing associations are suppressed over the course of repeated testing [9,10]. In other words, testing is thought to refine memory representations to selectively strengthen the target response [9,10,12,13]. These ideas have also been linked to the literature on retrieval-induced forgetting. For example, repeatedly retrieving "pineapple" to the cue "fruit- p…..? " facilitates the response "pineapple" but inhibits the alternative response "pear" [14,15]. Selective retrieval, thus, seems to strengthen target responses while inhibiting related but undesired responses.

Recently, Karpicke et al. [16] presented a possible mechanism that could underlie the selection processes during repeated testing. According to their "episodic context account", items become associated with the episodic context in which they are studied. During retrieval, the context from earlier presentations is re-activated and becomes integrated with current contextual

[2] In this article, we use the term "practice-testing" when we describe experimental paradigms, to distinguish testing in the practice phase in which learners engage in retrieval, from the final (performance) test used to measure the outcomes of practice. See also Fig. 1A.

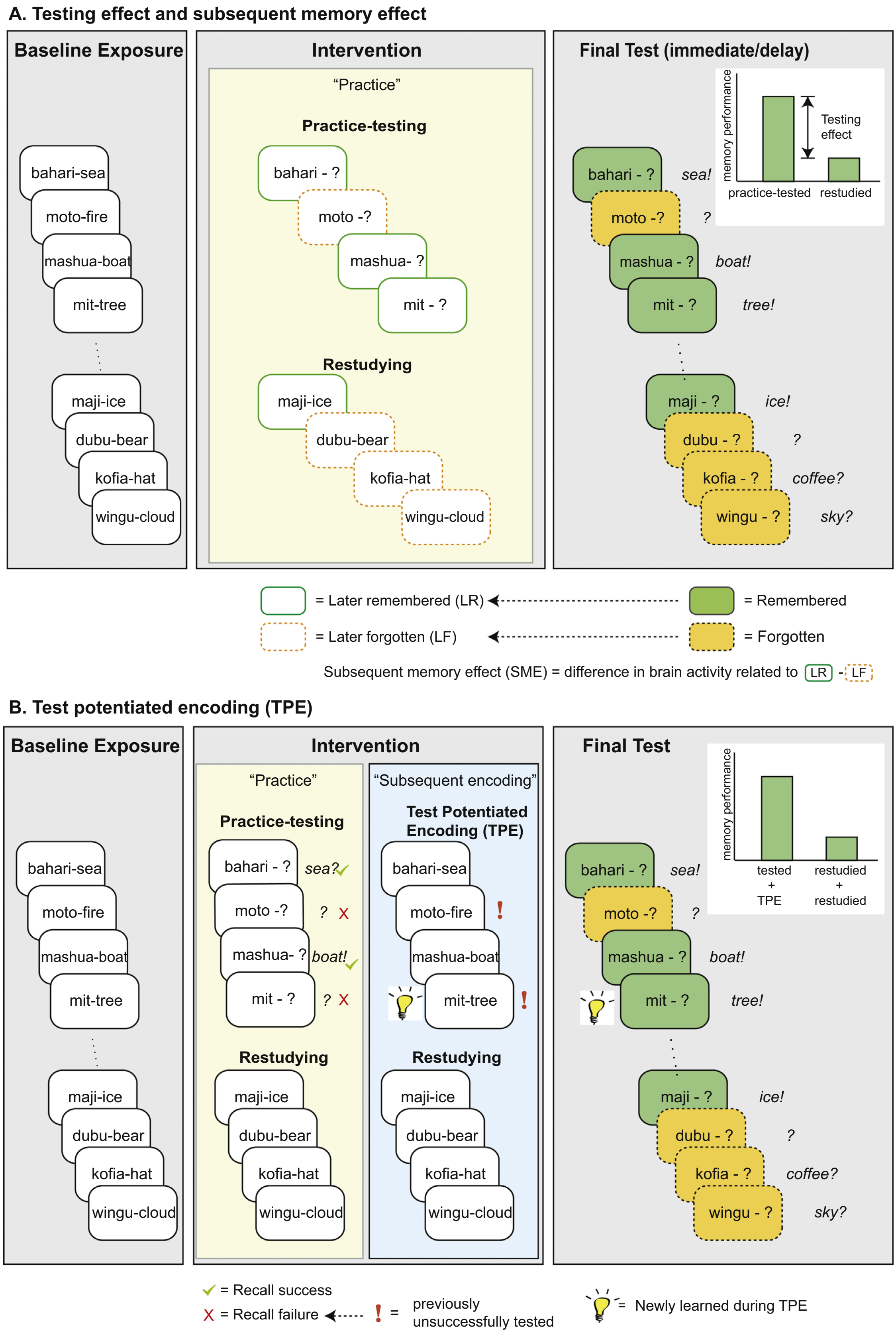

Fig. 1. A. Schematic view of a typical testing effect paradigm, with subsequent memory contrast. The typical set-up of a testing effect experiment is that participants undergo (1) a *baseline exposure* before (2) the critical intervention period in which they *practice* the items through *restudying* (encoding of the complete information) or *practice-testing* (retrieval of part of the information from memory). There can be several rounds of practice with repeated restudying or practice-testing. (3) To measure the testing effect on memory performance, a *final test* is administered either immediately after the intervention or after a delay. fMRI analyses: Testing effect fMRI studies have measured changes in brain activation both during the intervention and the final test period. The most common contrasts are between restudied and practice-tested items, and between *later remembered* and *later forgotten* items using the so-called *subsequent memory effect (SME).* B.Test-potentiated encoding (TPE) paradigm. In TPE studies, the intervention period includes practice (practice-testing and restudying) followed by "*test-potentiated encoding*" to observe the effect of prior practice-testing on *subsequent encoding*. Brain activations during test-potentiated encoding are often contrasted between items that were previously successfully and unsuccessfully tested, and can also be related to performance on the final test using SMEs as explained in Fig. 1A.

information. This refines the context representation associated with an item because those contextual features that serve as effective retrieval cues are strengthened the most. As a result, the search set of candidates that are activated increasingly zooms in on the target response, while competing responses are suppressed.

The elaboration and search-set restriction accounts are both compatible with another popular account in the cognitive literature, namely that practice-testing is a more effortful process than restudying [18]. The amount of effort during practice-testing has been related to the size of testing effects, with more difficult tests producing better outcomes than easier tests [17,18]. Mental effort is thus likely to be important for testing, but the definition and interpretation of the term "effort" is complex. Roediger and Butler called it "an index of the amount of reprocessing of the memory trace that occurs during retrieval" [3, p5]. Hence, the term is only vaguely defined and more effort could reflect both more elaboration and an increasing number of available retrieval routes (following the elaboration account), or higher selection demands and more or stronger suppression of competing incorrect responses (following the search-set restriction account). Whichever process underlies the effort during testing, both accounts predict that practice-testing changes cue-target associations in such a way that future memory retrieval is facilitated. With repetition, demands on effortful retrieval processes are thought to decrease because the target information becomes increasingly accessible through changes in the available retrieval routes. This facilitation of memory retrieval is then thought to lead to better long-term performance because the final performance test involves similar retrieval processes as those engaged during the practice-testing, and because a greater overlap of the cognitive processes involved in a final performance test with the cognitive processes engaged during practice is thought to enhance performance (an idea known as *transfer-appropriate processing* [19]).

In addition to direct benefits of testing, indirect effects of testing on subsequent learning are also highly relevant for educational practice. Testing enhances the efficiency of subsequent encoding in comparison to pure restudy conditions [20–22], which is known as *test-potentiated encoding* (TPE). Explanations of TPE broadly fall into two categories. First, testing could enable the learner to better discriminate between information that is successfully recalled on the test (and thus likely already well learned) and information that is not recalled (and thus needs further studying) [e.g., 4, 5]. In this way, testing enables a more efficient focus of attention during subsequent encoding, and the refinement of learning strategies. For example, mnemonic mediators that link cue and target information could become refined over the course of repeated testing [23]. A second explanation is that the testing context is reactivated during TPE when learners are reminded of prior testing during the encoding opportunity. This could enrich the memory representation and create additional retrieval cues when the different testing contexts become integrated [24,25].

In summary, cognitive explanations of the testing effect resolve around three lines: 1) memory representations change due to elaboration of relevant and/or suppression of irrelevant associations, 2) testing is an effortful process that becomes facilitated by repetition, and 3) testing leads to more efficient subsequent learning (TPE). In the following, we will relate these ideas to available neuroimaging studies.

3. Neural correlates of the testing effect

Although the accounts outlined above cannot be directly translated into predictions about brain activity (or vice versa), results from functional magnetic resonance imaging (fMRI) studies

Box 2.

Candidate brain regions of potential importance for the testing effect

Cognitive accounts of the testing effect suggest that practice-testing compared to restudying leads to a) changed *memory representations*, and b) facilitated *selective retrieval*. Here we discuss which brain regions are likely to support these cognitive processes. Note: All anatomical areas that are mentioned can be found in Fig. 2.

Semantic memory representations

It is well established that patterns of brain activity during memory retrieval partly overlap with those seen during initial encoding [71–73]. Thus, visually encoded material will evoke visual processing areas, and auditory memory will evoke auditory processing areas at retrieval, even when memory is probed with a different sensory input. For encoding and retrieval of semantic memory, as used in all fMRI studies on testing effects, it is likely that brain areas related to semantic processing will be involved. Semantic representations contain multi-sensory information widely distributed across cortical and subcortical regions of the brain [74–76]. Candidate brain regions for semantic representations include posterior brain areas such as the lateral temporal cortex, and the inferior parietal lobe (IPL) including the supramarginal (SMG) and the angular (AG) gyri (see Fig. 2). These are involved in the storage and conceptual integration of semantic representations [74,77]. The inferior (ITG) and the middle (MTG) temporal gyri, as well as the anterior temporal lobe, are thought to integrate different sensory inputs into multi-modal representations and are therefore expected to be active during concept retrieval [75,78,79]. The AG in the IPL is a higher-order associative area thought to integrate different components of semantic concepts into a coherent meaning [74,75,78–81].

Selective retrieval of semantic memory

In order to retrieve target representations from memory, processes related to control and monitoring are necessary. The ventral lateral prefrontal cortex (VLPFC) has been discussed as one key brain region underlying cognitive control during memory retrieval [82,83], and is thought to direct attention to goal-relevant information and inhibit irrelevant information during selective retrieval [14,84]. The repeated selection of target information and inhibition of competing information during testing that are thought to facilitate later retrieval, are likely mediated by this area. This is supported by the involvement of the VLPFC during selective retrieval that weakens competing memories (cf. retrieval-induced forgetting, [85]): VLPFC activity increases during retrieval when there is competition between several responses and decreases when a specific memory trace is selectively strengthened.

The anterior cingulate cortex (ACC) and the dorsal lateral prefrontal cortex (DLPFC) are other brain areas that are activated when demands on cognitive control are high and are suggested to be a part of the attention-control network. The ACC is thought to detect conflicts in information processing and plays a role in outcome evaluation and decision making [86]. The DLPFC is thought to play a role in directing attention to task-relevant representations [42], selectively mediate the resolution of response conflicts [87], and to become deactivated with practice as a function of decreased demands on selection [42].

can be related to the explanations (see Box 2). To date, ten fMRI studies have investigated beneficial effects of practice-testing for learning (see Table 1). Four of these studies measured classic testing effects by comparing performance after restudying and practice-testing[3] and related brain activation during these different practice conditions to later performance [26–29]. Three other studies investigated the consequences of practice-testing by measuring changes in brain activation after practice, or related brain activation during practice-testing to results on a subsequent final test [30–32]. Two studies focused at TPE [24,25]. One study [33] analyzed brain activations during restudying directly following practice-testing, and measured brain activation patterns during interleaved restudying and practice-testing. In the following, we review these neuroimaging studies in light of the cognitive accounts outlined above to summarize how reports of brain activity related to practice-testing can tentatively inform ideas from the cognitive literature on testing effects. In the course of reviewing, some other literature related to, but not directly focusing on testing effects, will also be discussed. A schematic overview of the experimental paradigms employed in the reviewed studies and the terms that are used in this article to describe them are presented in Fig. 1. More detailed descriptions of the paradigms of the studies can be found in Table 1.

3.1.1. *How does testing affect memory representations?*

The cognitive account that memory traces are semantically elaborated through testing predicts greater activation of semantic representations during practice-testing than restudying. If this is the case, we would expect greater involvement of brain areas related to semantic memory retrieval, such as the left temporo-parietal areas (see Box 2 and Fig. 2 for candidate brain structures) during practice-testing than restudying. Furthermore, the activation in these areas would be expected to increase with repetition. If, on the other hand, search-set restriction underlies testing effects, then one would expect that retrieval leads to less activation with each repetition.

The four studies that directly compared brain activations during practice-testing and restudying (see Table 1) reported *lower* activation in putative semantic memory storage areas during practice-testing relative to restudying. Wing et al. [29] had participants practice English word-pairs in an MRI scanner. After baseline exposure, half of the pairs were practiced through testing (cued recall of the second word) and half through restudying (re-exposure to the complete pair). Practice-testing led to significantly better memory performance than restudying one day after practice. However, brain activity was lower during practice-testing than restudying in the left superior temporal gyrus (STG) and in the bilateral middle temporal gyrus (MTG). Similarly, van den Broek et al. [28] observed less activation in the right MTG during practice-testing than restudying. Moreover, brain activity was lower during practice-testing than restudying in the inferior parietal lobe (IPL), including the supramarginal gyrus (SMG), and the angular gyrus (AG) in three studies [27–29], although in one study [26], in which participants retrieved semantically associated words from prior knowledge, activation in bilateral SMG/AG was *higher* during practice-testing than during restudying.

In order to understand which processes contribute to the testing effect, the relation between brain activity during practice and later performance is critical. In brain imaging studies, differences in brain activation during studying of items that are later remembered in contrast to later forgotten items are often taken as evidence for successful encoding and termed *subsequent memory effect* (SME; [34]; see also Fig. 1A). Within the putative semantic memory representation areas, increased activation predictive of later successful memory performance was found in the bilateral AG and SMG [28], the right [28] or the left [29] inferior temporal gyrus (ITG), the left MTG [28,33], the left STG [26,33] and the right STG [28,29,33]. However, these effects differed between practice-testing and restudying. Whereas activity was consistently *higher* during practice-testing of items that were later remembered than for items that were later forgotten [26,28,29,33] (cf. also the SME in the right superior parietal lobe in [31]), no difference [26,28] or even a reversed effect [29] was found during restudying. Taken together, these results suggest that the engagement of semantic memory storage areas during practice-testing is different from that during restudying. Tentatively, activity in these areas appears to be higher during restudying than practice-testing, but only predictive of later performance when measured during practice-testing.

In a different type of study relevant to the question how practice-testing affects memory representations, Keresztes et al. [32] reported brain activity on an immediate and a delayed final test after restudying or practice-testing. They found that activity in the IPL decreased from the immediate to the delayed test for restudied items but not for items that were tested during practice. Since one hypothesis concerning the level of activation in the IPL is that it might reflect the strength or richness of retrieved representations [35,36], the comparably stable activation level over time for tested items might indicate that practice-testing made memory traces more resistant to decay than restudying.

More direct investigations of changes in memory representations over time have recently become available from two publications that apply multivariate analysis techniques to the study of repeated retrieval [31,37]. Karlsson Wirebring et al. [31] used *representational similarity analysis* [RSA; 38] to correlate patterns of brain activation between multiple trials to determine under which conditions brain activations are more similar or dissimilar from each other. The study focused on successful repeated retrieval across three consecutive practice-tests, and related patterns of brain activations during practice-testing to performance at a final test one week later (see Table 1). Remembered items elicited higher activity in different areas of the brain, including the left lateral temporal cortex and bilateral posterior parietal cortices one week after practice. Notably, the same right parietal region identified as important for retrieval success one week after practice showed higher BOLD activity already at the day of practice for items that were later successfully remembered. An RSA in this region revealed that activation patterns were less correlated over the three consecutive tests for items later remembered compared to those forgotten. This can be interpreted as a sign of more altered or elaborated semantic representations during repeated practice-testing for items later remembered, which tentatively supports the idea that semantic elaboration is one key mechanism fostering long-term retention after repeated practice-testing.

So far, there has not been a study that compared changes in brain activations during repeated testing to changes during repeated studying. However, Xue et al. [39] investigated brain activations during repeated studying and it is interesting to note that

[3] These include two studies [26,27] in which participants either read a pair of semantically related words or generated a word when presented with a cue word and the initial letter of a semantically associated word (e.g., salt - p*****). We do not distinguish between such retrieval of pre-existing semantic associations and the retrieval of recently learned associations as employed in the other testing effect studies in this review, because both involve the activation of target information in response to a cue, focus processing on cue-target relations, and can be influenced in a similar way by certain experimental manipulations (cf. [47]). Therefore, we refer to the “read” and “generate” conditions as “restudying” and “practice-testing”, respectively in the main text. However, because some authors argue that the generation of semantic associations is qualitatively different from the retrieval of recently learned episodic associations (e.g., [9]), the type of retrieval is mentioned in the text if results differ between the two types of studies.

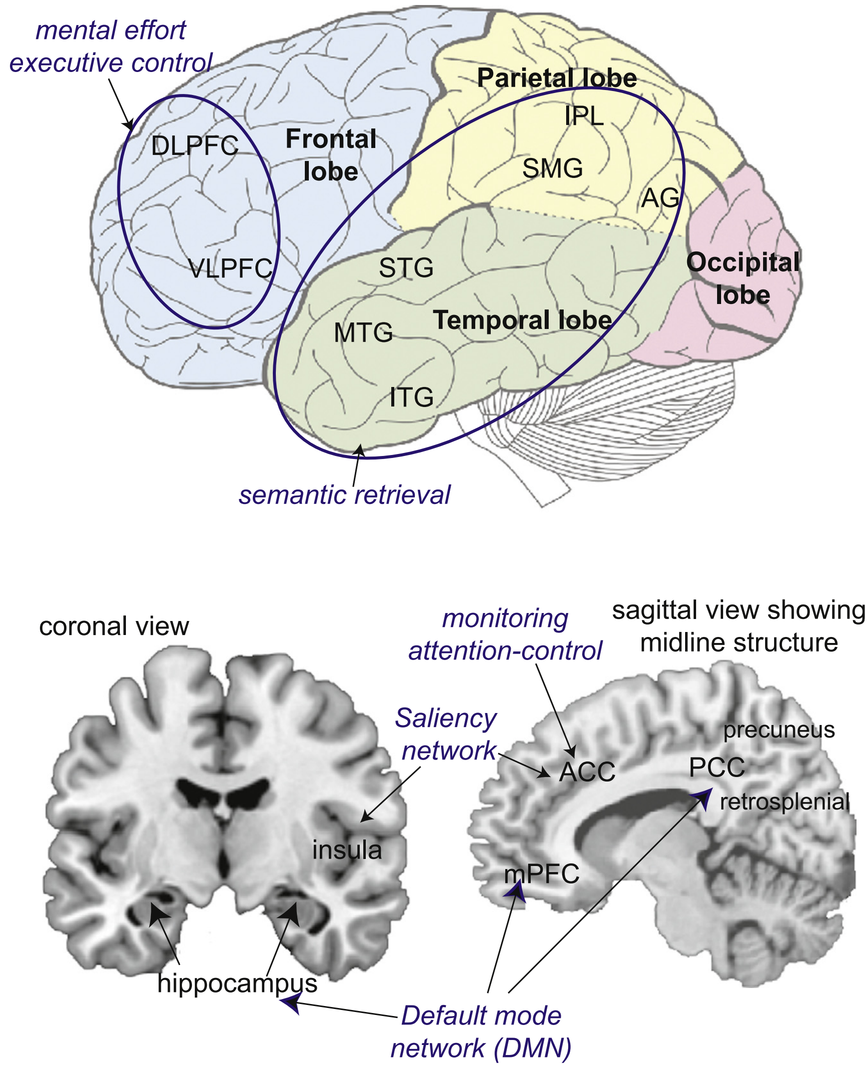

Fig. 2. Key anatomical areas of the brain for the testing effect. The upper panel shows a lateral view of the brain. The temporal lobe (superior temporal gyrus STG; middle temporal gyrus MTG; inferior temporal gyrus ITG), and the inferior parietal lobe (IPL) including the supramarginal (SMG) and the angular (AG) gyri, have often been found to increase in activation during semantic tasks (see Box 2). The dorsal lateral prefrontal cortex (DLPFC) and the ventral lateral prefrontal cortex (VLPFC) have been found to increase in activation when tasks require mental effort and executive control, such as selective memory retrieval (see Box 2). The lower panel shows coronal (left) and sagittal midline (right) sections of the brain. The anterior cingulate cortex (ACC) has been related to attention control and conflict monitoring. Furthermore, this area together with the anterior insula is a part of the *Saliency network* (see Box 3). The medial prefrontal cortex (mPFC), the posterior cingulate cortex (PCC), the retrosplenial, the precuneus and the hippocampus along with the inferior parietal lobe (IPL) and lateral temporal cortices have been associated with the *default mode network* (DMN; see Box 3). The Figure in the upper panel was adapted from https://commons.wikimedia.org/wiki/File:Lobes_of_the_brain_NL.svg.

unlike Karlsson Wirebring et al. [31], they reported that a *greater* similarity in patterns of neural activation across study trials predicted better memory performance. Xue et al. concluded that items for which repeated studying leads to a consistent neural representation are better remembered. The different conclusions from these two studies suggest that the effect of variations of the amount of semantic activation across repetitions might be different during testing [31] and studying [39]. The role of semantic elaboration in explaining the testing effect thus needs further exploration.

The second report of fMRI measurements during repeated retrieval is available from a study by Wimber et al. that focused on retrieval-induced forgetting [37]. The authors employed what they call a *canonical template tracking method* to determine the activation state of target memories and competing information during testing. Participants initially learned to associate one cue word (e.g., *sand*) with two different pictures (e.g., *sand – Marilyn Monroe, sand – hat*) and were then repeatedly instructed to retrieve only one of the pictures (e.g., *Marilyn Monroe*). The authors investigated the overlap between brain activations during this selective retrieval and template brain responses to the target (*Marilyn Monroe*) and competitor (*hat*) information in order to infer to what extent the target and competitor information were activated when cued with the word (*sand*). Results suggest that while participants initially tended to activate both target and competitor, the competitor was progressively suppressed over the course of repeated testing. Moreover, target information was reinstated increasingly over repetitions. Although this study does not directly address testing effects, it is informative for the evaluation of search-set restriction accounts of the testing effect because it shows that selective retrieval can lead to the suppression of competing information while strengthening target information. However, it is

Table 1
Paradigms, behavioral testing effect and key brain activations related to beneficial effects of testing in the 10 reviewed studies.

Study	Overview of experiment		Behavioral testing effect	Key brain regions involved
Studies focusing on brain activity during the intervention				
Karlsson Wirebring et al. 2015 [31]	Baseline exposure	Intentional study of 60 Swahili-Swedish word-pairs ten consecutive times	n/a (no comparison with a study condition)	During the intervention:SME during repeated correct retrieval: ∘ LR > LF: right SPL – RSA analysis in the right SPL during repeated correct retrieval: ∘ LR > LF: Lower pattern similarity for LR items compared to LF items ∘ Repetition x subsequent memory interaction: a monotonic decrease in the left DLPFC over the course of repeated testing for LR items, but not for LF items During the final test: – Retrieval success effect (remembered > forgotten items) in bil. PPC (right SPL/AG), bil. ITG, bil. IFG, right HC and PHG, left putamen (no differences for the reversed contrast)
	Intervention*	Each word-pair was repeatedly tested three times. The Swahili word was used as a probe, and participants indicated whether they knew the Swedish word, believed they knew or did not know. Immediately after, they chose among four alternatives the second letter of the word.		
	Final test*	After 7 days, cued recall of the translation of the Swahili words as during the intervention		
Rosner et al. 2014 [27]	Baseline exposure	none (participants had prior knowledge of the 100 pairs of semantically related words that were practiced)	Yes, at immediate test higher correct recognition rate for tested (87%) than for restudied words (65%)	– Activation difference ∘ T > RS for LR condition: VLPFC, bil VLPFC, LOC, ITG, IPS, PrC and ACC. ∘ RS > T: none – Correlation between behavioural and neural testing effect: Participants who showed a stronger behavioral testing effect tended to show a larger difference in neural activity during testing and restudying of LR words in paracingulate, frontal pole, left ACC, and right SFG – Brain activity related to successful retrieval during the final test did not differ between previously T and RS
	Intervention*	Half of the pairs were generated (= tested) once ("garbage = w_st_"), and half were read (= estudied) once ("garbage = waste"); i.e. the task was to either retrieve a target word that was semantically related to the cue from memory or to read the complete word pair.		
	Final test*	Immediately after the intervention, a recognition test with confidence rating included all target words and 100 lures.		
van den Broek et al. 2013 [28]	Baseline exposure	Intentional study of 100 Swahili words with translations through writing task and repeated exposure with judgments of learning	Yes, at delayed test: Higher cued recall and shorter response times for tested (58%) than for restudied words (49%)	– Activation difference ∘ T > RS: bil VLPFC/insula, bil striatum ∘ RS > T : bil IPL, right MTG ∘ LR > LF for T but not RS: left MTG, left IPL(AG/SMG) ∘ LF > LR for T but not RS: left calcarine, left SMA
	Intervention*	Half of the pairs were tested three times ("mit –translate!"), and half were restudied ("wingu - cloud") three times;		
	Final test	After 7 days, cued recall of the translation of the Swahili words		
Vannest et al. 2012 [26]	Baseline exposure	none (participants had prior knowledge of the 60 pairs of semantically or phonologically related words that were practiced)	Yes, at immediate test better multiple choice performance for tested (80%) than for restudied words (72%)	– Activation difference ∘ T > RS: bil VLPFC/MFG/insula, ACC/MedFG, bil caudate, bil AG/SMG/IOG/MOG/SOG, ∘ RS > T: Left M/SFG, right MFG, bil insula, bil IPL, right PrC, right lingual, left cuneus ∘ LR > LF for T but not RS: Left STG/SMG, left insula, right MedFG ∘ LF > LR for T but not RS: Right insula/VLPFC – Correlations between neural testing effect and recognition performance at the final test across participants: ∘ Degree of T > RS in Left M/STG predicted better recognition of tested items. ∘ Degree of T > RS in Right M/SFG, bil ACC, left PCC,right insula,right paracentral lobule predicted *worse* recognition of tested items. ∘ Degree of RS > T in bil cuneus. Right PrC predicted better recognition of restudied items. ∘ Degree of RS > T in Right M/MedFG, left postcentral, right cingulate, left MTG predicted *worse* recognition of restudied items.
	Intervention*	Half of the pairs were generated (=tested; "salt - p*****") once and half were read (=restudied; "salt – pepper") once. Similar to Rosner et al. study, participants either retrieved associated word from memory or read the word pair that was presented on the screen. Participants knew that a memory test would follow.		
	Final test	Immediately after the intervention; forced-choice task in which one word was presented and the associated word was selected among two lures		

Table 1 (*continued*)

Study	Overview of experiment		Behavioral testing effect	Key brain regions involved
Wing et al. 2013 [29]	**Baseline exposure*** **Intervention***	Rating semantic relatedness of 192 weakly related noun pairs Half of the pairs were tested ("TUSK -?") once and half were restudied ("TUSK – HORN") once. ° Note that this study alternated between baseline exposure and intervention. It consisted of 12 blocks in which 16 words were first presented for baseline exposure and then for one testing or restudy trial.	Yes, at delayed test higher cued recall for tested (63%) than for restudied words (51%)	– Activation difference ○ T > RS: bil VLPFC/insula, ACC, left ITG, left PrC, left PHG, left MOG ○ RS > T: bil MFG, bil MTG, bil IPL, right PrC ○ LR > LF: SMedFG – Activity interaction ○ LR > LF for T and reverse for RS: Left I/MTG, bil hippocampus ○ Interaction effect (no direction info): ACC, right STG, left insula/claustrum – Connectivity with hippocampus interaction ○ LR > LF for T only: PCC, vmPFC, left VLPFC, ○ LR > LF vs T > RS (no direction reported): Right MFG, right MedFG, bil STG, right insula/STG, right MTG, left ITG, left PHG, right SMG, bil postcentral, right PrC
	Final test	After 1 day, surprise cued recall test by presenting the first word as cue for recall of the second word		
Studies focusing on brain activity during test-potentiated encoding				
Nelson et al. 2013 [24]	**Baseline exposure*** **Intervention***	Intentional study of 126 weakly associated word pairs Interim testing ("crater - ?") or restudying ("crater – lake") followed by subsequent (test-potentiated) encoding (crater – lake).	No At delayed test, no difference between tested and restudied only word pairs	– Activation difference ○ baseline exposure > test-potentiated encoding: Left VLPFC, ○ baseline exposure < test-potentiated encoding: Left IPL/AG, PrC, MCC ○ Previously T > RS at test-potentiated encoding: Left IPL/AG, PrC, MCC – Activity level at test-potentiated encoding for items not recalled previously in left IPL correlated with the amount of new learning: left IPL/AG – Overlap of time course of brain activation in left IPL/AG during test-potentiated encoding with time course of activation during successful recognition found in previous meta-analysis
	Final test	After 1 day, cued-recall combined with detection of new words, by presenting the first words intermixed with 42 new words to prompt the cued recall of the second word (or recognition as new)		
Vestergren et al. 2014 [25]	Baseline exposure	24 h before the intervention, intentional study of 120 Swahili words with translations during 5 presentations. Immediately preceding the intervention one more presentation of all word pairs.	No At immediate test, no difference between tested and restudied only word pairs	– Activation difference during test-potentiated encoding ○ previously T > RS: Bil VLPFC/insula, left hippocampus ○ previously RS > T: Left MCC, bil SMG, bil PrC, bil MTG ○ Previously unsuccessfully recalled T: recalled later > not recalled later: anterior insula – No general activation difference during test-potentiated encoding between items that were previously tested unsuccessfully and items that were previously tested successfully
	Intervention*	Testing (wingu _____) or restudying (wingu – cloud) followed by subsequent (test-potentiated) encoding (wingu – cloud).		
	Final test	Immediately after intervention, cued recall of the translation of the Swahili words, followed by a multiple choice task to select the translation among 3 lures		
Liu et al., 2014 [33]	**Baseline exposure*** **Intervention***	Intentional study of 45 high-frequency unrelated Chinese-Chinese word pairs One testing trial (T1) per word pair immediately followed by restudy. Cue words were shown with a prompt to recall the second word before selecting it among all possible targets. After each test trial, the complete pair was shown for 3 seconds for restudying.	N/A All trials were both tested and subsequently restudied There was an increase in recall accuracy from T1 to T2	– LR > LF for successful test trials at T1 (brain activity during T1 dependent on performance at T2) ○ ROI analysis: left PFC, right PFC, right PPC, and left hippocampus. Marginally significant left PPC and right hippocampus ○ whole brain analyses:left SFG, MFG, IFG, right IFG; left IPL, SMG, MTG; left STG; right STG, MTG – Significant correlation between activation for LR items during T1 and performance at T2: ○ ROI analysis: in right PFC (r = .64, p = .022) and right PPC (r = .57,p = .012) – LR > LF for *unsuccessful* test trials at T1 (brain activity during T1 dependent on performance at T2) ○ MFG, Precuneus, Cingulate gyrus
	Final test	Immediately after the intervention, cued recall of the target words (T2). The delay between the intervention and the final test of single items was approximately 20 minutes.		

Table 1 (*continued*)

Study	Overview of experiment		Behavioral testing effect	Key brain regions involved
				– SME for baseline exposure ○ PFC, left PPC, and bilateral hippocampus, left – Test-potentiated learning during restudy after a failed test: trials that became correct on T2 > trials that were again incorrect on T2 ○ ROI: Marginally significant differences in left hippocampus and left PFC ○ Whole brain analysis: Putamen/caudate
Studies focusing on brain activity after the intervention				
Eriksson et al. 2011 [30]	Baseline exposure	Intentional study of 40 Swahili words with translations (This was combined with intervention)	Yes **Delayed test:** Positive correlation between number of successful practice tests and memory performance 1 week later	– Activation difference during final test for items previously ○ more retrieved < less retrieved items : Right SPL, right VLPFC ○ more retrieved > less retrieved items: ACC ○ more repetition of successful retrieval led to higher ACC activation – Participants who benefited from testing during intervention (more retrieval of items at intervention led to heightened ACC activation during final test) were the ones who had better memory performance 5 months later
	Intervention	Alternating restudying and testing until a word was tested successfully, then only testing continued and restudying stopped for that word for at least 4 cycles		
	Final test*	After 1 day, cued recall of the translation of the Swahili words (fMRI). The same test was given again after 7 days and 5 months (no fMRI)		
Keresztes et al. 2013 [32]	Baseline exposure	Intentional study of 60 Swahili words with translations	Yes, at delayed test higher cued recall for tested (50%) than for restudied words (39%) At immediate test, no difference between tested and restudied words.	– Analyses were focused at a set of brain areas (regions of interest) activated by a working memory task. – 2 way interaction Condition x Testing Moment: ○ Activation difference RS > T after 20min, T > RS after 1 week: bil DLPFC, bil insula, bil IPL, right Mid orb FG, right SPL, left fusiform, right thalamus ▪ This pattern of activations resembled the pattern found for behavioural effects (RS= T after 20min, T > RS after 1 week) ▪ Interaction appeared driven by different changes over time : all areas decreased in activation over time after RS, no areas decreased over time after T
	Intervention	6 rounds which each consisted of 1 block of 30 words being tested and one block of 30 words being restudied, and one (subsequent) encoding block of all words		
	Final tests*	Immediately (half of the participants) and 7 days (other half of the participants) after the intervention. Cued recall of the translation of the Swahili words.		

Note. The overview of the experiment refers to the same baseline, intervention, and final test phases as depicted in Figure 1. Phases printed in bold and marked with an asterisk were conducted in the MR scanner. Abbreviations: T = testing; RS = restudying; T > RS = contrast between testing trials and restudy trials; LR = later remembered (see Figure 1); LF = later forgotten. Abbreviations of anatomical regions in alphabetical order: ACC = anterior cingulate cortex; AG = angular gyrus; IFG = inferior frontal gyrus; IPL = inferior parietal lobe; IPS = inferior parietal sulcus; ITG = inferior temporal gyrus; LOC = lateral occipital cortex; MCC = middle cingulate cortex; MedFG = medial frontal gyrus; MFG = middle frontal gyrus; Mid orb FG = middle orbital frontal gyrus; MTG = middle temporal gyrus; PHG = parahippocampal gyrus; PPC= posterior parietal cortices; PrC = precuneus; SFG = superior frontal gyrus; SMedFG = superior medial frontal; SMG = supramarginal gyrus; SPL = superior parietal lobe; STG = superior temporal gyrus; VLPFC = ventrolateral prefrontal cortex

unclear how such competitor suppression is related to the retention of target information because the study focused largely on competitor suppression rather than target enhancement.

In summary, neural responses in brain areas related to semantic memory retrieval lead to mixed conclusions about the competing cognitive models of testing effects. Overall, the available studies indicate a different role for temporo-parietal regions during practice-testing and restudying. Tentatively, activity in these areas appears to be higher during restudying than practice-testing, but only predictive of later performance when measured during practice-testing. Practice-testing, if successful, seems to strengthen the neural representations of target information in temporo-parietal regions. Restudying on the other hand seems to evoke semantic information that is less relevant for learning, possibly related to mind-wandering (cf. Box 3). The question of whether semantic networks become elaborated or restricted over the course of repeated practice-testing remains open. The results from Karlsson Wirebring et al. [31] suggest that successful repeated retrieval that fosters long term retention might be characterized by semantic elaboration, as indicated by reduced pattern similarity within the parietal lobe for items subsequently remembered. This supports the idea that semantic elaboration might underlie the benefits of testing [7,8]. However, in the study by Xue et al. [39], reduced pattern similarity during repeated studying predicted worse outcomes. Variations in semantic activation during repeated practice may thus not always be beneficial, and may have different effects during practice-testing and restudying. Regarding search-set restriction accounts, one study on retrieval-induced forgetting [37] showed that information that directly competes with relevant cue-target associations becomes increasingly suppressed over the course of repeated testing.

3.1.2. The role of mental effort during retrieval

The cognitive accounts that testing is an effortful process predict an involvement of brain regions that support controlled, selective memory retrieval. Executive control over effortful retrieval is often associated with the involvement of the prefrontal cortex, especially the ventral and dorsal lateral prefrontal cortex (VLPFC, DLPFC). Furthermore, the DLPFC and the anterior cingulate cortex (ACC) are often related to attention, control, and conflict monitoring processes (see Box 2 and Fig. 2). Indeed, the four studies that analyzed brain activity during practice-testing and restudying [26–29] consistently reported higher activation in the VLPFC and the ACC during practice-testing compared to restudying.

Concerning the association between brain activation during practice and later performance at final test, different patterns of results were reported for the VLPFC and the ACC. Enhanced hippocampal functional connectivity with the VLPFC during practice-testing but not restudying predicted later performance [29], which could reflect an interaction of executive control processes through the VLPFC with core memory processing areas in the hippocampus [40,41] during practice-testing. There were no reports that higher VLPFC activation itself predicted better performance, and in a generation effect study by Vannest et al., VLPFC activation during practice-testing even predicted later forgetting of words [26]. Further in line with this, Karlsson Wirebring et al. [31] reported a decrease in activity of the left DLPFC over the course of repeated successful practice-testing which was predictive of better later memory performance. Similarly, in the previously mentioned retrieval-induced forgetting study by Wimber et al. [37], activity in the left and right VLPFC decreased over the course of repeated selective retrieval. Moreover, activations predicted how much competing information was suppressed but not how much target information was enhanced: more activation in the left and right VLPFC during the retrieval of specific memories predicted stronger suppression of that memory's competitors. At the same time, activations decreased over the course of repeated selective retrieval, and the stronger the decrease was, the more competing information was suppressed [37]. Overall, the available results are in line with the idea that activations in the VLPFC reflect the need to engage control mechanisms to select target information among competing memories [42], and that this process occurs during practice-testing more than restudying. Selection processes seem to become facilitated after repeated practice-testing, resulting in reduced recruitment of effortful control processes.

Cognitive accounts that testing facilitates later retrieval predict that brain activity during a later memory test (final test) should change as a function of prior testing. Indeed, VLPFC activity on a final test one day after combined restudying/practice-testing was lower the more often items had been successfully tested during prior practice [30], possibly reflecting that prior testing made the retrieval on the final test less demanding and reduced the need for competitor suppression. Behavioral reports of faster reaction times at the final test for previously tested compared to previously restudied items further support this interpretation [28,32,43].

The interpretation of ACC involvement in testing effects is not straightforward. Higher ACC activity during practice-testing than restudying predicted a larger behavioral testing effect in a study by Rosner et al. [27], as participants who benefited more from practice-testing compared to restudying showed larger increases in ACC activation during practice-testing than restudying. Higher ACC activation measured during practice-testing was also more predictive of (better) performance on the final test than ACC activation measured during restudying [29]. Although these results point to a positive effect of ACC involvement during practice-testing, one study in which highly associated words were recalled (generation) reported that participants who showed higher ACC activity during practice-testing than restudying tended to perform *worse* on the final test [26]. A possible explanation for these contradictory outcomes is that ACC activation during practice could reflect the detection of competition between target responses and related competitors. Such conflict could, on the one hand, correlate with effortful, beneficial retrieval processes during practice-testing and thereby predict better performance at final test. On the other hand, high levels of conflict could also be a consequence of higher item difficulty, and therefore be related to worse performance.

Unlike activity in the VLPFC, ACC activity did not decrease but *increased* as a function of prior testing. In one study, the strength of this effect predicted performance five months later [30]: Participants who showed a larger ACC response to previously tested items thus performed better later on. In a different study in which participants were scanned immediately and a week after testing/restudy practice, practice-testing led to better performance than restudying on the delayed final test but not on the immediate final test [32] (a typical result in testing effect studies, see Box 1). Activation in several brain areas, including the ACC, showed a similar interaction between time and condition as they were more active during the delayed final test of previously practice-tested than restudied materials, but not during the immediate final test.

In summary, activations in the VLPFC that are likely to reflect cognitive control processes required for the selection of target information among competing memories [37,42,44,45] were enhanced during practice-testing compared to restudying. Higher demands on this putative executive-control system did not, however, predict better learning. Instead, evidence suggests that a reduction of activation may correlate with the facilitation of retrieval processes, as competing information becomes less activated and demands on inhibition processes are reduced over the course of repeated testing. Results with respect to ACC involvement are mixed and cannot fully be explained by the idea that ACC activation is involved in conflict detection. This area has been related to long-term benefits of testing in two studies, and might play a role in consolidation processes that enhance retention [30,32].

3.2. Neural correlates of test-potentiated encoding (TPE)

Three fMRI experiments [24,25,33] addressed TPE. In two experiments [24,25], participants encoded word-pairs and practiced the word-pairs between encoding episodes by practice-testing or restudying (see Fig. 1). In the third study, all word-pairs were tested and then immediately restudied [33]. For all TPE experiments, the neural response of interest is that during subsequent encoding *after* testing.

Some TPE accounts predict that previously unsuccessfully tested items receive extra attention during subsequent encoding [17]. Vestergren and Nyberg [25] specifically addressed this idea by comparing brain activity during the encoding of previously unsuccessfully tested items to successfully tested and restudied items. In their study, no differences in brain activity were found that were specific to re-encoding of *unsuccessfully* tested items, although several areas were more activated overall after practice-testing than restudying. This finding suggests that TPE might affect all previously tested items regardless of testing success.

Both Liu et al. [33] and Vestergren and Nyberg [25] investigated which brain activations correlated with encoding success during TPE of previously unsuccessfully tested items. Liu et al. found that the left putamen and the caudate, the left hippocampus and the left PFC were more active during restudying of previously unsuccessfully tested items when these items were subsequently remembered than when they were not remembered. The authors suggested that the striatal (caudate, putamen) activity reflects changes in memory representations in response to negative feedback from the failed testing trial. Regarding the striatal involvement, van den Broek et al. [28] also reported activity increases during successful practice-testing compared to restudying. They attributed this effect to highlighting of motivationally significant information through the dopaminergic system. Vestergren and Nyberg [25] found that activity in the anterior insula reflected successful encoding and related this to the possible involvement of a saliency detecting network (see Box 3 and Fig. 2).

A consistent result across the two experiments that included comparisons of activations during encoding of previously tested and restudied items was higher activation in the VLPFC for previously tested items than un-tested items [24,25]. Vestergren and Nyberg speculated that this increased VLPFC activity along with activity in the anterior insula and the ACC could reflect deep processing due to the involvement of a saliency network in the brain that detects tested items as relevant and reduces distraction and mind-wandering during practice (see Box 3). Moreover, these authors found higher hippocampal activity for previously practice-tested than for restudied items, which suggests that practice-tested items underwent further re-encoding. Although they did not compare practice-tested and restudied items, Liu et al. [33] reported (marginally) higher activation in the left hippocampus and the left PFC for word pairs that benefited from re-encoding after unsuccessful prior testing. Activity in these areas was higher for successful re-encoding relative to unsuccessful re-encoding after a failed test, but not after a successful test. These results are in line with cognitive accounts of TPE that suggest that items receive extra attention after retrieval failure [17] (but see [25]).

Among areas found to be more active during encoding of previously tested than during encoding of previously restudied words, a study by Nelson et al. [24] focused analyses on the left IPL/AG, an area that had shown retrieval-related activity in an earlier meta-analysis [46]. They found that activity in the IPL/AG correlated with the amount of new learning during TPE. Moreover, the authors plotted the average time course of brain activity for TPE trials from the IPL/AG showing signal fluctuation of a typical TPE trial. Based on a comparison of this plot to the results of a previous meta-analysis on memory recognition, the authors argued that the time course of brain activity during the re-encoding of previously tested words (TPE) resembled that of successful recognition. The activity of re-encoding of previously restudied pairs, in contrast, resembled that of seeing unknown items. From this finding, the authors concluded that there is increased retrieval-like activation in the IPL/AG during TPE. This could be due to the reinstatement of prior testing experiences, in line with the idea that prior testing becomes incorporated into the memory representation. This interpretation is somewhat in line with the "episodic context account" proposed by Karpicke et al.[16].

In sum, the three studies that investigated TPE revealed that after unsuccessful testing, activity in the insula [25] and the PFC, as well as the parietal cortex and the hippocampus [33] correlated with successful TPE. Only one study by Vestergren and Nyberg directly compared previously successfully and unsuccessfully tested items during subsequent encoding [25] and concluded that TPE might affect all previously tested items regardless of prior testing success.

Box 3.

Reverse inference: The role of Default-Mode and Saliency networks in testing practice

Imaging studies have revealed patterns of activation that may stimulate ways to think about the testing effect that go beyond existing cognitive accounts. The role of default mode and saliency networks is an example of such results (see Fig. 2 for candidate brain regions).

The default mode network/resting state network

Imaging studies have identified the *Default Mode Network* (DMN), a set of brain regions that are coactive during rest but become relatively deactivated during demanding tasks [88,89]. Brain regions commonly implicated in the DMN include the ventro- and dorso-medial prefrontal cortex, the hippocampus, the posterior midline structures (posterior cingulate; retrosplenial, precuneus), as well as the inferior parietal lobe (IPL) and lateral temporal cortices. The DMN was initially identified during rest, and it has been interpreted as reflecting internally driven thoughts that are, at least partly, task-unrelated. Indeed, unsuccessful memory encoding is accompanied by increased DMN activation [90], alluding to the idea that increased DMN activity reflects lapses in the focus of attention towards the encoding task. However, DMN activity has also been found when tasks involve self-referential thought such as (autobiographical) episodic memory retrieval [91,92].

Several testing effect fMRI studies reported higher activation in the DMN during restudying compared to practice-testing [26,28,29]. This may indicate that testing reduces mind-wandering and distraction in comparison to restudying and, as a consequence, enhances attention to the task, which is in accordance with behavioral studies [93,94]. At the same time, a somewhat puzzling result is that increased activity in lateral temporal and parietal areas of DMN during practice-testing (but not during restudying) predicted successful retention [26,28,29]. A possible reason for this is that testing requires retrieval, which involves task-related self-referential thoughts that engage parts of the DMN.

Saliency network

Activity in the DMN is negatively correlated with activity in a network that becomes active when attention must be directed to external stimuli [95] instead of internal thoughts [96,97]. This *saliency network* (including as important nodes the anterior insula and the anterior cingulate cortex; ACC) might play a crucial role in responding to salient cues in the environment that require attention and reducing internally directed thought [98,99]. During practice-testing more than restudying, and also as a function of the amount of prior testing, the saliency network is reported to become activated [25,26,28–30]. Thus, testing possibly increases the learners' attention to the materials through the saliency network. This could be an additional explanation of the testing effect revealed by neuroimaging studies, thus adding to the accounts suggested in the cognitive literature.

4. Towards a neurocognitive account of the testing effect

In the behavioral literature, the testing effect has been explained in terms of changes in semantic representations through elaboration or the restriction of the search-set to relevant associations, effortful retrieval processes that become easier with repetition, and the potentiation of subsequent encoding. Can these cognitive accounts be informed by the neuroimaging studies published to date?

A relatively consistent finding across the testing effect fMRI studies is that even though the engagement of temporo-parietal semantic memory storage areas, such as the IPL and the MTG, was greater during restudying compared to practice-testing [26–29], activation increases only predicted later performance when measured during practice-testing and not when measured during restudying [26,28,29]. Engagement of these areas could reflect how testing alters memory representations such that they become more resistant to decay. A simple interpretation of the observation that temporo-parietal storage areas were less active during practice-testing than during restudying would be that testing does not generally enhance elaboration in comparison to restudying. However, as van den Broek et al. [28] argued, activations in these areas could reflect both relevant and irrelevant semantic information processing, and reduced activations during practice-testing could also reflect a beneficial focus of attention on relevant information. Testing might direct elaboration to relevant associations that improve retention more than unfocused elaboration that takes place during restudy. Support for the notion that testing elaborates memory representations comes from the finding that successful repeated retrieval that fostered long-term retention was characterized by higher BOLD signal but lower pattern similarity in the parietal cortex [31]. However, there is also support for an involvement of selection processes during testing. During repeated selective retrieval, target information seems to become increasingly activated, whereas competing information becomes increasingly suppressed [37]. These results raise the question of whether an alternative cognitive model is needed that can accommodate both elaboration and selection processes. For example, elaboration during testing could be selective and focus on associations that strengthen the cue-target link, whereas associations that compete with the target response are suppressed.

The fMRI findings related to selective retrieval and inhibition processes are mixed. There is evidence that the VLPFC and the ACC are more involved during practice-testing than restudying [26–29], and several studies demonstrated a link between activity in the VLPFC and the ACC and later memory performance at final test, but the direction of this link differed among studies. Moreover, different effects of prior testing on VLPFC and ACC activations during the final test suggest a differential role of these two areas during testing. With regard to the VLPFC, one explanation could be that the VLPFC does not directly influence cue-target associations, but only through interactions with the core memory system in the hippocampus [29]. The VLPFC may act during the first instances of testing when selection demands during retrieval are high but with repeated successful retrieval, the VLPFC engagement might decrease. This claim is supported by the relation between changes in brain activity over the course of repeated retrieval and later memory performance [31,37]. With regard to the ACC, activation increases on the final test after prior practice-testing are difficult to interpret under the assumption that the ACC detects conflicting response options during retrieval. An alternative interpretation of the role of the ACC is that it reflects the amount of attention evoked by the tested materials in a similar way as during TPE (see Box 3 on the ACC as a part of the saliency network). By this view, attention is higher during practice-testing than restudying [29], more attention is paid to previously tested items than previously restudied items [25], and through this heightened attention there is a higher chance for these items to become well-consolidated for better retention [30].

Imaging studies on the testing effect have also revealed patterns of activation that cannot easily be linked to existing cognitive accounts. Certain structures of the brain were not covered in this review, such as the posterior cingulate cortex (PCC) and the thalamus. At least some of these areas are likely to also play a functional role in testing, and future findings could reveal additional explanations of the testing effect. As a case in point, we outline in Box 3 how the testing effect could be explained in terms of reduced mind-wandering or saliency detection.

Results from TPE studies show how testing influences subsequent learning. Depending on the timing of the re-encoding occasion, encoding success was predicted by the involvement of different areas. Immediate feedback after a failed practice-test evoked extra brain activity in the left PFC when items were later remembered (compared to forgotten) on the final performance test [33]. When re-encoding occurred after a whole round of practice-testing, activation in the anterior insula correlated with successful re-encoding [25]. Analyses of patterns of brain activation suggest that practice-testing could lead to the engagement of retrieval processes in the IPL/AG regions during subsequent encoding when subjects are reminded of prior testing, and thereby form an alternative learning context that enriches memory representations [24]. Alternatively, or in addition, testing could increase attention via the engagement of the saliency network through the VLPFC/anterior insula and the ACC activity (see Box 3).

5. Future perspectives and conclusion

Several cognitive theories exist that make predictions about the nature of the testing effect and its constituent processes. However, there is no consensus yet about how to best explain the effect. This review adds a neurocognitive perspective to the literature by summarizing the evidence available from fMRI studies on the testing effect.

It is not trivial to link neuroimaging results to the behavioral literature and a number of limitations need to be taken into account. Cognitive theories are usually not constrained by the way the human brain works, and can have a rather abstract level of description. This makes one-to-one mapping between cognitive accounts and neural responses challenging. In addition, typical paradigms in neuroimaging and in behavioral research differ, which can make it more difficult to compare results. On the one hand, these differences in paradigms are due to methodological constraints of imaging studies, such as the required high numbers of comparable trials and the types of responses that participants can make in the scanner. On the other hand, imaging studies have different paradigms to allow analyses that are not possible with behavioral data, such as tracking changes in patterns of brain activity over the course of repetitions or test-potentiated encoding. In fMRI studies, it is also possible to categorize trials retrospectively, such as into later-remembered and later-forgotten trials (i.e., SME) and observe brain responses that are predictive of memory outcomes at a later test. As in other areas of educational neuroscience, continued exploration and testing of the translation between and incorporation of theories from cognitive and neuroscientific fields are called for.

The number of imaging studies of beneficial effects of practice-testing on learning outcomes is still comparably small. For this review, we identified ten fMRI studies that related practice-testing to later performance or explicitly focused at the testing effect or TPE. These studies employed a variety of approaches, measuring brain activations during practice-testing and restudying, over the course of repeated practice-testing, during subsequent encoding after practice-tests, and during final performance tests. Each of these approaches allows the evaluation of different predictions from the cognitive literature, but the number of studies of each approach is still small. More research is therefore needed before more definite conclusions can be drawn. In addition, a number of limitations of the literature base should be noted. First, two of the reviewed TPE-studies did not report a behavioral effect [24,25]. This did not influence the analysis of the imaging data because brain activations were related to prior retrieval success and later performance on a by-item basis, but a behavioral effect would confirm that the chosen paradigms caused TPE. Second, two studies [31,33] did not include a restudy comparison condition, so it is unclear to what extent the reported neural activations during practice-testing also occur during other forms of practice such as restudying. This however, is not necessarily problematic for cognitive theories that predict that quantitatively *more* elaboration or mental effort is involved during practice-testing compared to restudying, rather than qualitatively different processes. Third, we reviewed two studies that implemented retrieval of pre-existing semantic associations rather than recently learned associations [26,27]. Although similarities exist between these two forms of retrieval (cf. [47]), it would be informative to take into account the nature of the associations that are activated during practice-tests in future studies. The testing effect has been found with many different materials (Box 1) but all testing-effect fMRI studies so far used visual word-pairs. It is an open question if similar effects are obtained when participants study different materials (e.g., auditory, non-verbal).

Regarding the fMRI methods employed, the focus of the available imaging studies on testing effects has been largely on activity changes in specific (more or less isolated) nodes of the brain. We suggest that in future studies, the neural mechanisms underlying the testing effect may be better described by examining patterns of interactions across several brain areas. Connectivity analyses of the interactions between brain regions may help develop a network perspective on the mechanisms of testing, as illustrated by the results of hippocampal connectivity by Wing et al. [29]. Extending such analyses to other brain regions may reveal interactions in a broader network of areas underlying the testing effect. Furthermore, changes in memory representations as a consequence of practice-testing may be better documented by changes in *patterns* of neural activation over the course of and after practice rather than net activation differences during practice-testing and restudying. Using techniques like multivariate pattern analysis/RSA [38,48], researchers have only recently begun to investigate testing effects in this way [31]. More research along these lines might further improve insight into the mechanisms of testing effects.

Imaging studies offer a unique way to test predictions from cognitive theories and may suggest new ways to think about well-known behavioral phenomena like the testing effect. This review of recent fMRI studies on the testing effect informs the literature about its neurocognitive substrates by highlighting several different processes that might be important for testing effects. First, the available data support the idea that practice-testing indeed engages the memory representational areas in the posterior cortices in a different way than restudying does, possibly in a more focused way that stabilizes the relevant memory trace. Examining variation in patterns of brain activity during successful repeated retrieval in relation to subsequent memory, produced support for the semantic elaboration view. However, more studies are needed to establish how these results fit with studies that show the suppression of competing information during selective retrieval. An alternative cognitive model might be needed that can accommodate both elaboration and selection processes, such as selective elaboration. Second, there is evidence for effortful, controlled retrieval processes reflected in the engagement of the prefrontal cortex during practice-testing, which reduces over the course of repeated practice-testing, although the link of this reduction with later performance at final test needs further investigation. Third, TPE is not restricted to materials that were previously not recalled and could involve the reactivation of testing experiences and extra attention to previously tested information. Finally, neuroimaging studies point at other possible mechanisms that have not been covered by cognitive accounts yet, such as enhanced attention through the engagement of motivation or saliency network or less mind-wandering (see Box 3).

Glossary

- **Default mode network (DMN):** anatomically defined interconnected brain system that is activated when individuals are engaged in self-referential thoughts such as daydreaming, envisioning the future, retrieving memories, and gauging others' perspectives [91]. It is negatively correlated with brain systems that focus on external signals.
- **Functional connectivity:** a method to define multiple regions that are co-active during a certain brain processing state. It is often measured as the statistical association or dependency among time-series derived from two or more anatomically distinct regions of the brain [100].
- **Functional magnetic resonance imaging (fMRI):** indirect measure of neuronal activity assessed over time. The most common type of fMRI measurement is sensitive to the amount of deoxygenated hemoglobin in the blood, which is known to peak around 4–8 s following the onset of neuronal activity; this effect is known as the blood oxygenation level-dependent (BOLD) contrast.
- **Hippocampus:** a region of the brain that is essential for learning and memory formation [40].
- **(mnemonic) Mediator:** information that links cue and target information, such as a key word that is recallable when prompted with a cue and elicits target information; usually studied in the context of association learning.
- **Representational Similarity Analysis (RSA):** a multivariate pattern analysis technique that makes it possible to investigate whether patterns of activity within sets of voxels differ between experimental conditions or item categories [38,48].
- **Retrieval-induced forgetting:** refers to the phenomenon that if one studies A–B and A–C associations, and selectively retrieves A–B associations through testing, the memory of A–C associations is weakened [85].
- **Saliency network:** a network of brain areas thought to detect the most relevant information among internal and external stimuli in order to focus attention onto relevant information [99].
- **Testing effect:** the phenomenon that *testing*, or *retrieval practice*, enhances later memory performance of a material, typically relative to *restudying* or no re-exposure.
  Restudying: a practice condition in which complete materials are presented for study; often used as a comparison condition in testing effect studies because it requires no memory retrieval
  Testing: a practice condition in which the learner is asked to retrieve information from memory.
- **Test-potentiated encoding (TPE):** the phenomenon that testing prior to a later encoding session increases the effectiveness of encoding
- **Transfer appropriate processing:** a notion that the overlap between the way in which information is initially encoded and the way in which it is later retrieved influences memory performance [19]. Usually, more overlap induces better performance.

Acknowledgments

This research was supported by the National Initiative Brain & Cognition, Netherlands Organization for Scientific Research (NWO Grant number 056-33-014), the Swedish Research Council (VR

Grant number 721-2014-2099) and the Torsten and Ragnar Söderberg's Foundation (to LN) (Grant number KVA/2011/88/65).

References

[1] Y. Dudai, The neurobiology of consolidations, or, how stable is the engram? Annu. Rev. Psychol. 55 (2004) 51–86.
[2] G. Winocur, M. Moscovitch, Memory transformation and systems consolidation, J. Int. Neuropsychol. Soc. 17 (2011) 766–780.
[3] H.L. Roediger, A.C. Butler, The critical role of retrieval practice in long-term retention, Trends Cognit. Sci. 15 (2011) 20–27.
[4] H.L. Roediger, J.D. Karpicke, The power of testing memory: Basic research and implications for educational practice, Perspect. Psychol. Sci. 1 (2006) 181–210.
[5] H.L. Roediger, J.D. Karpicke, Test-enhanced learning: taking memory tests improves long-term retention, Psychol. Sci. 17 (2006) 249–255.
[6] D.E. Rumelhart, J.L. McClelland, PDP Research Group, Parallel Distributed Processing, 1, The MIT Press, Cambridge, 1986.
[7] S.K. Carpenter, E.L. Delosh, Impoverished cue support enhances subsequent retention: support for the elaborative retrieval explanation of the testing effect, Mem. Cogn. 34 (2006) 268–276.
[8] S.K. Carpenter, Cue strength as a moderator of the testing effect: the benefits of elaborative retrieval, J. Exp. Psychol.: Learn. Mem. Cogn. 35 (2009) 1563–1569.
[9] J.D. Karpicke, F.M. Zaromb, Retrieval mode distinguishes the testing effect from the generation effect, J. Mem. Lang. 62 (2010) 227–239.
[10] R.C. Thomas, M.A. McDaniel, Testing and feedback effects on front-end control over later retrieval, J. Exp. Psychol.: Learn. Mem. Cogn. 39 (2013) 437–450.
[11] S.K. Carpenter, Semantic information activated during retrieval contributes to later retention: Support for the mediator effectiveness hypothesis of the testing effect, J. Exp. Psychol.: Learn. Mem. Cogn. 37 (2011) 1547–1552.
[12] J.D. Karpicke, J.R. Blunt, Retrieval practice produces more learning than elaborative studying with concept mapping, Science 331 (2011) 772–775.
[13] J.D. Karpicke, M.A. Smith, Separate mnemonic effects of retrieval practice and elaborative encoding, J. Mem. Lang. 67 (2012) 17–29.
[14] B. Storm, B. Levy, A progress report on the inhibitory account of retrieval-induced forgetting, Mem. Cogn. 40 (2012) 827–843.
[15] K. Murayama, T. Miyatsu, D. Buchli, B.C. Storm, Forgetting as a consequence of retrieval: a meta-analytic review of retrieval-induced forgetting, Psychol. Bull. 140 (2014) 1383–1409.
[16] J.D. Karpicke, M. Lehman, W.R. Aue, Retrieval-based learning: an episodic context account, in: B.H. Ross (Ed.), Psychology of learning and motivation, Elsevier Academic Press, San Diego, CA, 2014, pp. 237–284.
[17] M.A. Pyc, K.A. Rawson, Why testing improves memory: mediator effectiveness hypothesis, Science 330 (2010) 335.
[18] M.A. Pyc, K.A. Rawson, Testing the retrieval effort hypothesis: does greater difficulty correctly recalling information lead to higher levels of memory? J. Mem. Lang. 60 (2009) 437–447.
[19] R.S. Lockhart, Levels of processing, transfer-appropriate processing, and the concept of robust encoding, Memory 10 (2002) 397–403.
[20] P. Grimaldi, J. Karpicke, When and why do retrieval attempts enhance subsequent encoding? Mem. Cogn. 40 (2012) 505–513.
[21] L.E. Richland, N. Kornell, L.S. Kao, The pretesting effect: Do unsuccessful retrieval attempts enhance learning? J. Exp. Psychol.: Appl. 15 (2009) 243–257.
[22] C. Izawa, The test trial potentiating model, J. Math. Psychol. 8 (1971) 200–224.
[23] M.A. Pyc, K.A. Rawson, Why is test–restudy practice beneficial for memory? An evaluation of the mediator shift hypothesis? J. Exp. Psychol.: Learn. Mem. Cogn. 38 (2012) 737.
[24] S.M. Nelson, K.M. Arnold, A.W. Gilmore, K.B. McDermott, Neural signatures of test-potentiated learning in parietal cortex, J. Neurosci. 33 (2013) 11754–11762.
[25] P. Vestergren, L. Nyberg, Testing alters brain activity during subsequent restudy: Evidence for test-potentiated encoding, Trends Neurosci. Educ. 3 (2014) 69–80.
[26] J. Vannest, K.P. Eaton, D. Henkel, M. Siegel, R.K. Tsevat, J.B. Allendorfer, B. K. Schefft, C. Banks, J.P. Szaflarski, Cortical correlates of self-generation in verbal paired associate learning, Brain Res. 1437 (2012) 104–114.
[27] Z.A. Rosner, J.A. Elman, A.P. Shimamura, The generation effect: activating broad neural circuits during memory encoding, Cortex 49 (2013) 1901–1909.
[28] G.S.E. van den Broek, A. Takashima, E. Segers, G. Fernández, L. Verhoeven, Neural correlates of testing effects in vocabulary learning, Neuroimage 78 (2013) 94–102.
[29] E.A. Wing, E.J. Marsh, R. Cabeza, Neural correlates of retrieval-based memory enhancement: an fmri study of the testing effect, Neuropsychologia 51 (2013) 2360–2370.
[30] J. Eriksson, G. Kalpouzos, L. Nyberg, Rewiring the brain with repeated retrieval: a parametric fmri study of the testing effect, Neurosci. Lett. 505 (2011) 36–40.
[31] L. Karlsson Wirebring, C. Wiklund-Hörnqvist, J. Eriksson, M. Andersson, B. Jonsson, L. Nyberg, Lesser neural pattern similarity across repeated tests is associated with better long-term memory retention, J. Neurosci. 35 (2015) 9595–9602.
[32] A. Keresztes, D. Kaiser, G. Kovács, M. Racsmány, Testing promotes long-term learning via stabilizing activation patterns in a large network of brain areas, Cereb. Cortex (2013).
[33] X.L. Liu, P. Liang, K. Li, L.M. Reder, Uncovering the neural mechanisms underlying learning from tests, PLoS One 9 (2014) e92025.
[34] T.F. Sanquist, J.W. Rohrbaugh, K. Syndulko, D.B. Lindsley, Electrocortical signs of levels of processing: Perceptual analysis and recognition memory, Psychophysiology 17 (1980) 568–576.
[35] E.T. Reas, J.B. Brewer, Retrieval search and strength evoke dissociable brain activity during episodic memory recall, J. Cognit. Neurosci. 25 (2012) 219–233.
[36] M.D. Rugg, K.L. Vilberg, Brain networks underlying episodic memory retrieval, Curr. Opin. Neurobiol. 23 (2013) 255–260.
[37] M. Wimber, A. Alink, I. Charest, N. Kriegeskorte, M.C. Anderson, Retrieval induces adaptive forgetting of competing memories via cortical pattern suppression, Nat. Neurosci. 18 (2015) 582–589.
[38] N. Kriegeskorte, M. Mur, P.A. Bandettini, Representational similarity analysis - connecting the branches of systems neuroscience, Front. Syst. Neuroscience 2 (2008).
[39] G. Xue, Q. Dong, C. Chen, Z. Lu, J.A. Mumford, R.A. Poldrack, Greater neural pattern similarity across repetitions is associated with better memory, Science 330 (2010) 97–101.
[40] W.B. Scoville, B. Milner, Loss of recent memory after bilateral hippocampal lesions, J. Neurol. Neurosurg. Psychiatry 20 (1957) 11–21.
[41] L.R. Squire, S. Zola-Morgan, The medial temporal lobe memory system, Science 253 (1991) 1380–1386.
[42] B.A. Kuhl, N.M. Dudukovic, I. Kahn, A.D. Wagner, Decreased demands on cognitive control reveal the neural processing benefits of forgetting, Nat. Neurosci. 10 (2007) 908–914.
[43] G.S.E. van den Broek, E. Segers, A. Takashima, L. Verhoeven, Do testing effects change over time? Insights from immediate and delayed retrieval speed, Memory 22 (2014) 803–812.
[44] B.A. Kuhl, I. Kahn, N.M. Dudukovic, A.D. Wagner, Overcoming suppression in order to remember: Contributions from anterior cingulate and ventrolateral prefrontal cortex, Cogn. Affect. Behav. Neurosci. 8 (2008) 211–221.
[45] M. Wimber, K.-H. Bäuml, Z. Bergström, G. Markopoulos, H.-J. Heinze, A. Richardson-Klavehn, Neural markers of inhibition in human memory retrieval, J. Neurosci. 28 (2008) 13419–13427.
[46] S.M. Nelson, A.L. Cohen, J.D. Power, G.S. Wig, F.M. Miezin, M.E. Wheeler, K. Velanova, D.I. Donaldson, J.S. Phillips, B.L. Schlaggar, S.E. Petersen, A parcellation scheme for human left lateral parietal cortex, Neuron 67 (2010) 156–170.
[47] D.J. Peterson, N.W. Mulligan, The negative testing effect and multifactor account, J. Exp. Psychol.: Learn. Mem. Cogn. 39 (2013) 1287–1293.
[48] N. Kriegeskorte, R. Goebel, P. Bandettini, Information-based functional brain mapping, Proc. Natl. Acad. Sci. USA 103 (2006) 3863–3868.
[49] C.A. Rowland, The effect of testing versus restudy on retention: a meta-analytic review of the testing effect, Psychol. Bull. 140 (2014) 1432–1463.
[50] L.C. Coppens, Verkoeijen PPJL, Rikers RMJP, Learning adinkra symbols: the effect of testing, J. Cognit. Psychol. 23 (2011) 351–357.
[51] N. Kornell, R.A. Bjork, M.A. Garcia, Why tests appear to prevent forgetting: a distribution-based bifurcation model, J. Mem. Lang. 65 (2011) 85–97.
[52] T.C. Toppino, M.S. Cohen, The testing effect and the retention interval, Exp. Psychol. 56 (2009) 252–257.
[53] V. Halamish, R.A. Bjork, When does testing enhance retention? A distribution-based interpretation of retrieval as a memory modifier, J. Exp. Psychol.: Learn. Mem. Cogn. 37 (2011) 801–812.
[54] M.A. McDaniel, J.L. Anderson, M.H. Derbish, N. Morrisette, Testing the testing effect in the classroom, Eur. J. Cognit. Psychol. 19 (2007) 494–513.
[55] C. Wiklund-Hörnqvist, B. Jonsson, L. Nyberg, Strengthening concept learning by repeated testing, Scand. J. Psychol. 55 (2014) 10–16.
[56] F.C. Leeming, The exam-a-day procedure improves performance in psychology classes, Teach. Psychol. 29 (2002) 210–212.
[57] M.A. McDaniel, R.C. Thomas, P.K. Agarwal, K.B. McDermott, H.L. Roediger, Quizzing in middle-school science: Successful transfer performance on classroom exams, Appl. Cognit. Psychol. 27 (2013) 360–372.
[58] S.K. Carpenter, H. Pashler, N.J. Cepeda, Using tests to enhance 8th grade students' retention of U.S. history facts, Appl. Cognit. Psychol. 23 (2009) 760–771.
[59] S.K. Carpenter, Testing enhances the transfer of learning, Curr. Dir. Psychol. Sci. 21 (2012) 279–283.
[60] K.B. Lyle, N.A. Crawford, Retrieving essential material at the end of lectures improves performance on statistics exams, Teach. Psychol. 38 (2011) 94–97.
[61] T. Stenlund, F.U. Jönsson, B. Jonsson, Group discussions and test-enhanced learning: Individual learning outcomes and personality characteristics, Educ. Psychol. (2016) 1–15.
[62] J. Dunlosky, K.A. Rawson, E.J. Marsh, M.J. Nathan, D.T. Willingham, Improving students' learning with effective learning techniques: Promising directions from cognitive and educational psychology, Psychol. Sci. Public Interest 14 (2013) 4–58.
[63] D. Rohrer, K. Taylor, B. Sholar, Tests enhance the transfer of learning, J. Exp. Psychol.: Learn. Mem. Cogn. 36 (2010) 233–239.
[64] C.B. Kromann, M.L. Jensen, C. Ringsted, The effect of testing on skills learning, Med. Educ. 43 (2009) 21–27.
[65] A.C. Butler, Repeated testing produces superior transfer of learning relative

to repeated studying, J. Exp. Psychol.: Learn. Mem. Cogn. 36 (2010) 1118–1133.
[66] S.K. Carpenter, J.W. Kelly, Tests enhance retention and transfer of spatial learning, Psychon. Bull. Rev. 19 (2012) 443–448.
[67] S.K. Kang, M. McDaniel, H. Pashler, Effects of testing on learning of functions, Psychon. Bull. Rev. 18 (2011) 998–1005.
[68] M. de Jonge, H.K. Tabbers, R.M.J.P. Rikers, Retention beyond the threshold: Test-enhanced relearning of forgotten information, J. Cognit. Psychol. 26 (2014) 58–64.
[69] A.N.D. Meyer, J.M. Logan, Taking the testing effect beyond the college freshman: Benefits for lifelong learning, Psychol. Aging 28 (2013) 142–147.
[70] B. Pastötter, J. Weber, K.-H.T. Bäuml, Using testing to improve learning after severe traumatic brain injury, Neuropsychology 27 (2013) 280–285.
[71] L. Nyberg, R. Habib, A.R. McIntosh, E. Tulving, Reactivation of encoding-related brain activity during memory retrieval, Proc. Natl. Acad. Sci. 97 (2000) 11120–11124.
[72] L. Nyberg, K.M. Petersson, L.-G. Nilsson, J. Sandblom, C. Åberg, M. Ingvar, Reactivation of motor brain areas during explicit memory for actions, Neuroimage 14 (2001) 521–528.
[73] M.E. Wheeler, S.E. Petersen, R.L. Buckner, Memory's echo: vivid remembering reactivates sensory-specific cortex, Proc. Natl. Acad. Sci. USA 97 (2000) 11125–11129.
[74] J.R. Binder, R.H. Desai, The neurobiology of semantic memory, Trends Cogn. Sci. 15 (2011) 527–536.
[75] J.R. Binder, R.H. Desai, W.W. Graves, L.L. Conant, Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies, Cereb. Cortex 19 (2009) 2767–2796.
[76] A. Martin, The representation of object concepts in the brain, Annu. Rev. Psychol. 58 (2007) 25–45.
[77] C.J. Price, A review and synthesis of the first 20 years of pet and fmri studies of heard speech, spoken language and reading, Neuroimage 62 (2012) 816–847.
[78] K. Patterson, P.J. Nestor, T.T. Rogers, Where do you know what you know? The representation of semantic knowledge in the human brain, Nat. Rev. Neurosci. 8 (2007) 976–987.
[79] C. Whitney, E. Jefferies, T. Kircher, Heterogeneity of the left temporal lobe in semantic representation and control: Priming multiple versus single meanings of ambiguous words, Cereb. Cortex. 21 (2011) 831–844.
[80] J.R. Binder, K.A. McKiernan, M.E. Parsons, C.F. Westbury, E.T. Possing, J. N. Kaufman, L. Buchanan, Neural correlates of lexical access during visual word recognition, J. Cogn. Neurosci. 15 (2003) 372–393.
[81] W.W. Graves, R. Desai, C. Humphries, M.S. Seidenberg, J.R. Binder, Neural systems for reading aloud: a multiparametric approach, Cereb. Cortex. 20 (2010) 1799–1815.
[82] D. Badre, A.D. Wagner, Left ventrolateral prefrontal cortex and the cognitive control of memory, Neuropsychologia 45 (2007) 2883–2901.
[83] E. Race, B.A. Kuhl, D. Badre, A.D. Wagner, The dynamic interplay between cognitive control and memory, in: M.S. Gazzaniga (Ed.), He Cognitive Neurosciences, MIT Press, Cambridge, MA, 2009, pp. 705–724.
[84] R.S. Blumenfeld, C. Ranganath, Prefrontal cortex and long-term memory encoding: an integrative review of findings from neuropsychology and neuroimaging, Neuroscientist 13 (2007) 280–291.
[85] M.C. Anderson, R.A. Bjork, E.L. Bjork, Remembering can cause forgetting: retrieval dynamics in long-term memory, J. Exp. Psychol.: Learn. Mem. Cogn. 20 (1994) 1063–1087.
[86] M. Botvinick, Conflict monitoring and decision making: reconciling two perspectives on anterior cingulate function, Cognit., Affect., Behav. Neurosci. 7 (2007) 356–366.
[87] D. Badre, A.D. Wagner, Selection, integration, and conflict monitoring: assessing the nature and generality of prefrontal cognitive control mechanisms, Neuron 41 (2004) 473–487.
[88] D.A. Gusnard, M.E. Raichle, Searching for a baseline: functional imaging and the resting human brain, Nat. Rev. Neurosci. 2 (2001) 685–694.
[89] M.E. Raichle, A.Z. Snyder, A default mode of brain function: a brief history of an evolving idea, Neuroimage 37 (2007) 1083–1090.
[90] H. Kim, Neural activity that predicts subsequent memory and forgetting: a meta-analysis of 74 fmri studies, Neuroimage 54 (2011) 2446–2461.
[91] R.L. Buckner, J.R. Andrews-Hanna, D.L. Schacter, The brain's default network, Ann. N. Y. Acad. Sci. 1124 (2008) 1–38.
[92] E. Svoboda, M.C. McKinnon, B. Levine, The functional neuroanatomy of autobiographical memory: a meta-analysis, Neuropsychologia 44 (2006) 2189–2208.
[93] K.K. Szpunar, N.Y. Khan, D.L. Schacter, Interpolated memory tests reduce mind wandering and improve learning of online lectures, Proc. Natl. Acad. Sci. 110 (2013) 6313–6317.
[94] K.K. Szpunar, H.G. Jing, D.L. Schacter, Overcoming overconfidence in learning from video-recorded lectures: Implications of interpolated testing for online education, J. Appl. Res. Mem. Cogn. 3 (2014) 161–164.
[95] M. Corbetta, G.L. Shulman, Control of goal-directed and stimulus-driven attention in the brain, Nat. Rev. Neurosci. 3 (2002) 201–215.
[96] M.D. Fox, A.Z. Snyder, J.L. Vincent, M. Corbetta, D.C. Van Essen, M.E. Raichle, The human brain is intrinsically organized into dynamic, anticorrelated functional networks, Proc. Natl. Acad. Sci. USA 102 (2005) 9673–9678.
[97] M.D. Greicius, B. Krasnow, A.L. Reiss, V. Menon, Functional connectivity in the resting brain: a network analysis of the default mode hypothesis, Proc. Natl. Acad. Sci. 100 (2003) 253–258.
[98] D. Sridharan, D.J. Levitin, V. Menon, A critical role for the right fronto-insular cortex in switching between central-executive and default-mode networks, Proc. Natl. Acad. Sci. 105 (2008) 12569–12574.
[99] V. Menon, L. Uddin, Saliency, switching, attention and control: a network model of insula function, Brain Struct. Funct. 214 (2010) 655–667.
[100] K.J. Friston, C.D. Frith, P. Fletcher, P.F. Liddle, R.S.J. Frackowiak, Functional topography: multidimensional scaling and functional connectivity in the brain, Cereb. Cortex 6 (1996) 156–164.