ALTERING THE EFFECTIVENESS OF REMINDERS:
POSITIVE AND NEGATIVE MNEMONIC EFFECTS OF
ASSOCIATIVE RETRIEVAL PROCESSES IN HUMAN EPISODIC
MEMORY

PhD thesis
Attila Keresztes

Supervisor:
Dr. Mihály Racsmány

Budapest, 2013
# Table of Contents

Table of Contents........................................................................................................................................ 3
Acknowledgments........................................................................................................................................ 5
Abbreviations............................................................................................................................................... 6
Abstract..................................................................................................................................................... 7
Kivonat......................................................................................................................................................... 8

**Introduction**............................................................................................................................................... 9
Reminders in everyday life ............................................................................................................................. 9
Associations in memory, and the concept of cues ......................................................................................... 10
The cue driven nature of retrieval processes in human episodic memory .............................................. 12
The central theoretical debate: Interference accounts versus inhibitory accounts of forgetting and remembering cue-target associations ................................................................................ 13
Outline of the dissertation ............................................................................................................................ 20

**Memory inhibition in cued recall** .......................................................................................................... 23
Inhibition theories inspired by the works of Freud .................................................................................... 23
Inhibition in cognition ................................................................................................................................. 24
Memory inhibition ...................................................................................................................................... 25

**The think/no-think paradigm** ................................................................................................................ 26
Logic of research using the TNT paradigm ............................................................................................... 27
Major behavioral findings ............................................................................................................................ 28
Major neuroimaging and electrophysiological findings ............................................................................ 29
Mainstream interpretation of findings ........................................................................................................ 32
Problems with the mainstream interpretation ............................................................................................ 33
  *The issue of cue-independent memory inhibition* ................................................................................. 33
  *Getting the TNT effect with independent probes without suppression* .............................................. 35
  *The issue of effect size* .......................................................................................................................... 36
  *The issue of the missing baseline in neuroimaging studies* ................................................................. 36
  *Explaining the DLPFC-hippocampus inhibitory orchestra with default network processing* ........ 36
A cautious interpretation ............................................................................................................................... 37
A new perspective on the TNT effect ........................................................................................................ 38

**The retrieval practice paradigm** .......................................................................................................... 40
Retrieval-induced forgetting (RIF) ............................................................................................................. 40
The retrieval practice paradigm .................................................................................................................. 40
Theories explaining RIF ............................................................................................................................. 41
Testable hypotheses of the inhibitory control theory .............................................................................. 42
Logic of research using the retrieval practice paradigm ............................................................................ 43
Behavioral findings ...................................................................................................................................... 44
Neuroimaging and electrophysiological findings ..................................................................................... 48
Evidence for and against the inhibitory control hypothesis .................................................................... 49
  *Retrieval specificity, strength independence and interference dependence* .................................... 49
  *Cue independence* ............................................................................................................................... 51
A coherent picture involving inhibition and the challenge of interference accounts ............................ 53
A potential cautious interpretation ............................................................................................................. 54
The positive effect of retrieval practice - The testing effect.................................................. 58
The testing effect.................................................................................................................. 59
Behavioral findings.............................................................................................................. 60
Moderating factors .............................................................................................................. 62
Theoretical explanations - yet without a coherent theory.................................................. 62
Neuroimaging results contributing to a future theory of the testing effect.......................... 66

Main objectives and thesis points....................................................................................... 68
Main objectives ................................................................................................................... 68
Thesis points ...................................................................................................................... 70

Thesis I. Suppressing target memory retrieval leads to changes in cue-target associations, but not in target memory representations. (Study 1).............................................................................. 70

Thesis II. Retrieval-induced forgetting is a non-monotonic function of competitor interference during retrieval. (Study 2)...................................................................................... 71

Thesis III. Retrieval-induced forgetting depends on competitive retrieval, but manipulations of learning strategy cannot provide evidence for strength independence of RIF. (Study 2 and 3).... 72

Thesis IV. Repeated retrieval enhances later retrieval by stabilizing activity patterns in a network of brain areas related to cognitive control. (Study 4)................................................................. 73

Studies................................................................................................................................ 75
Discussion............................................................................................................................. 125
References............................................................................................................................ 133
**Acknowledgments**

First, I thank my family, Anya, Apa, and Péter, for your understanding and supporting my commitment to psychology without which all these years could not have ended up in a dissertation (and maybe more).

I am especially grateful to Mihály Racsmány, my supervisor, who has given me more inspiration than he probably knows. Besides his vast factual knowledge and creative insights, the many projects he involved me in granted me an invaluable breadth of theoretical views, and methodological skills. Without your help, I could not have come this far.

Working together with Gyula Demeter, his friendship, his empathy, and our conversations, gave the PhD years a touch of nice collegiality that was further strengthened by Péter Pajkossy, and Ágnes Szöllősi. I am deeply thankful to all of you, it has been great fun to work together, and I hope this will continue.

I am also grateful to Gyula Kovács who introduced me to functional magnetic resonance imaging, and the idea of a predictive brain which now keeps me fascinated, and who also made my stays in Regensburg possible, fun, and successful.

Of course, it was Csaba Pléh who made all this possible by pioneering cognitive science in Hungary, and establishing the graduate school I have been a student of. His continued mentorship, with his striking insights, and extremely specific yet thoughtful advices on what to read, who to listen to, played a great role in finding a field that suits my interests most, channeling it from psychoanalytic theory to the cognitive neuroscience of memory that is evident also in this dissertation. I am deeply grateful to you for all this.

I am also thankful to all the colleagues and students I have worked together with during these years, and especially thankful to all of you who have given me feedback on what to do better.

Thanks to Balázs Aczél for shaping my interests and keeping renewing my motivation as an implicit mentor.

And many thanks to Lilla for your invaluable, full and loving support throughout the last years of my PhD.
### Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>BOLD</td>
<td>Blood Oxygenation Level Dependent</td>
</tr>
<tr>
<td>DLPFC</td>
<td>Dorsolateral prefrontal cortex</td>
</tr>
<tr>
<td>DM</td>
<td>Difference in memory</td>
</tr>
<tr>
<td>DRM</td>
<td>Deese-Roediger-McDermott</td>
</tr>
<tr>
<td>ERP</td>
<td>Event related potentials</td>
</tr>
<tr>
<td>fMRI</td>
<td>Functional magnetic resonance imaging</td>
</tr>
<tr>
<td>IAPS</td>
<td>International affective picture system</td>
</tr>
<tr>
<td>IFG</td>
<td>Inferior frontal gyrus</td>
</tr>
<tr>
<td>IPS</td>
<td>Intraparietal sulcus</td>
</tr>
<tr>
<td>LTS</td>
<td>Long-term store</td>
</tr>
<tr>
<td>MTL</td>
<td>Mediotemporal lobe</td>
</tr>
<tr>
<td>OCD</td>
<td>Obsessive-compulsive disorder</td>
</tr>
<tr>
<td>OSPAN</td>
<td>Operation span</td>
</tr>
<tr>
<td>PFC</td>
<td>Prefrontal cortex</td>
</tr>
<tr>
<td>PMDr</td>
<td>Dorsal premotor cortex</td>
</tr>
<tr>
<td>PTSD</td>
<td>Posttraumatic stress disorder</td>
</tr>
<tr>
<td>pre-SMA</td>
<td>Pre supplementary motor area</td>
</tr>
<tr>
<td>REM</td>
<td>Retrieving effectively from memory</td>
</tr>
<tr>
<td>RIF</td>
<td>Retrieval-induced forgetting</td>
</tr>
<tr>
<td>RT</td>
<td>Reaction time</td>
</tr>
<tr>
<td>SAM</td>
<td>Search of associative memory</td>
</tr>
<tr>
<td>STS</td>
<td>Short-term store</td>
</tr>
<tr>
<td>TNT</td>
<td>Think/no think</td>
</tr>
<tr>
<td>VLPFC</td>
<td>Ventrolateral prefrontal cortex</td>
</tr>
<tr>
<td>WM</td>
<td>Working memory</td>
</tr>
</tbody>
</table>
Abstract

The act of recall is an act of construction (Schacter, Norman, & Koutstaal, 1998), and remembering can even modify memories (Roediger, 2000). During associative retrieval, i.e. recalling a target memory to a cue present in the environment, cue-target associations are changed to meet current goals of remembering (Bjork & Bjork, 1992). In four studies, this dissertation explored how associative memories can be changed by competitive and non-competitive retrieval processes as well as by intentionally stopping retrieval. In Study 1, stopping retrieval of cue-target associations led to forgetting of targets by weakening unidirectional associations between cues and targets. Study 2 and 3 showed that competitive associative retrieval led to forgetting of non-retrieved target memories that were related to the retrieved ones. In addition, Study 2 showed that there is a non-monotonic relationship between the amount of competition during retrieval and later forgetting of competing targets. Study 3 demonstrated that this type of competitive retrieval, as compared to restudying associations, attenuates forgetting of retrieved target memories for a period as long as five days. In Study 4, using an event-related fMRI experiment, we found that such benefits of effortful associative retrieval are correlated with neural activity in brain areas related to cognitive control. We suggest that with changes that incorporate the potential role of cognitive control processes, two models, the SAM-REM theory, and the episodic inhibition theory might accommodate all modifications of episodic cue-target associations reported in this dissertation.
Kivonat

Az előhívás aktusa konstrukció (Schacter, Norman, & Koutstaal, 1998), sőt néha az emlékezés módosíthatja emlékeinket (Roediger, 2000). Asszociatív előhívás során – amikor egy, a környezetben jelen lévő hívóingerre adott reakcióként egy célemléket hívunk elő – a hívóinger-célemlék asszociációk megváltoznak az emlékezés aktuális céljainak megfelelően. A disszertáció négy tanulmányban azt vizsgálta, hogyan módosíthatóak asszociatív emlékek kompetitív és non-kompetitív előhívás során, illetve az előhívási folyamat szándékos leállítása révén. Az első tanulmányban a hívóinger-célemlék asszociációk előhívásának leállítása, a hívóingerek és a célemlékek közötti egyirányú kapcsolatok gyengítése révén, a célemlékek felejtéséhez vezetett. A második és harmadik tanulmány kimutatta, hogy a kompetitív asszociatív előhívás a nem előhívott, de a célemlékekhez kapcsolódó emlékek felejtéséhez vezet. A második tanulmány ezen kívül megmutatta, hogy az előhívás alatti versengés mértéke és a kapcsolódó, versengő emlékek felejtése között nonmonotonikus kapcsolat van. A harmadik tanulmány meghatározza, hogy az ilyen típusú versengő előhívás, szemben az asszociációk újratanulásával, lecsökkenti az előhívott célemlékek felejtését, akár egy ötnapos késleltetés mellett is. A negyedik tanulmányban, egy esemény alapú fMRI vizsgálatban kimutattuk, hogy az erőfeszítést igénylő asszociatív előhívásnak ez az előnye kognitív kontrollhoz társított agyterületek aktivitásával korrelál. Feltételezéseink szerint az epizodikus hívóinger-célemlék asszociációk e disszertációban bemutatott módosulásait - a háttérben lehetségesen meghúzódó kognitív kontroll folyamatok hatásainak beépítésével - két elmélet, a SAM-REM, illetve az epizodikus gátlás elmélete is megmagyarázhatja.
Introduction

Reminders in everyday life

We constantly stumble on reminders. Faces of long forgotten relatives in a photo album, objects popping up while you move from one flat to another trigger memories you have not thought about for a long time. At an exam, the questions on your blank sheet trigger memories of painful hours of bending over your books to learn for the exams. The view of an ATM, a bank card, an entry phone on a door trigger memories of numbers, this time well-learnt, almost automatic knowledge that you might find impossible to remember without the reminders.

Sometimes you wish the reminders not trigger anything: You go to the bar to order a drink, and a red label on a bottle on the shelf behind the bartender reminds you of a party you hardly want to think about, so you slightly shift the line of thoughts to something more pleasant. If you ever had an accident you might want to avoid reminders of that too, and if you still happen to face such reminders, you might try not to think about the event itself, and stop the retrieval process triggered by the reminder.

A reminder is associated to several memories, and these memories compete for retrieval every time you see the reminder. Whenever you switch your phone, address, etc. to a new one, the reminders 'dial', 'type' will trigger memories of your old phone number, address, etc, and you will need a few occasions of reminding until the reminder successfully leads you to your new number, location.

Learning in education is also overwhelmingly about reminders: trying to retrieve memories in the face of reminders, names of species, countries, titles of works, names of authors, words to translate, concepts, etc. are all reminders you associated some knowledge to, and you will
always be tested using reminders. During learning, you constantly and repeatedly retrieve memories to be learned.

Reminders guide most of our memories, and we wish they were always there for us at the right moment and disappear when unwanted. Indeed, reminders and their effectiveness to find memories change dynamically. In this dissertation, I will present studies on how these changes occur due to three processes highlighted in the above examples: stopping retrieval, competitive retrieval, and repeated retrieval. Before focusing on the central question of the dissertation, an introduction to the study of the retrieval process is inevitable. In the next sections, I will introduce the concept of associations between reminders and memories, and the concept of cues in episodic memory research. Then, I will give an overview of the whole retrieval process in human memory, emphasizing the critical role of cues in it.

**Associations in memory, and the concept of cues**

The study of reminders is closely related to the study of associations. Since the classical works of Pavlov (1927), associative learning refers to processes linking the mental representation of at least two physical stimuli (Mitchell, De Houwer, & Lovibond, 2009), e.g. the sound of a bell, and the smell of a ham. In the state-of the art dichotomy of memory, that distinguishes unconscious and conscious memory (Schacter, 1987; Squire, 1992), most of the unconscious learning phenomena is referred to as associative learning (Shrager & Squire, 2009). The only exception is habituation and sensitization that are termed together non-associative learning forms. This distinction is clearly due to the distinction between associative and non-associative nature of the underpinning neural mechanisms. In the realm of conscious memory, the term association, or associative learning refers loosely to the phenomenon of one knowing that two events occurred in association to each other (Mitchell et al., 2009).
Interestingly, the associative nature of retrieval in conscious memory was largely neglected, until the nineteen sixties, when Endel Tulving introduced the concept of retrieval cue to the experimental psychology of memory (Baddeley, 1999). Tulving and colleagues ran several experiments that showed that retrieval cues can largely aid retrieval. For instance, in one experiment, Tulving and Osler (1968) presented to be remembered words, such as 'short' or 'stomach', together with cues that were associateable to these words, such as 'girl' or 'hurt', respectively. Subsequently, they asked their participants to recall the to be remembered words, and either provided them with the cues (aided condition) or not (unaided condition). Participants performed significantly better in the aided condition. In another experiment (Tulving & Pearlstone, 1966) participants learnt lists of words belonging to categories, and their attention was explicitly drawn to categories. Again, giving the cues as reminders at a final test increased recall performance drastically. From today's understanding of memory processes these results might seem trivial, however the role of retrieval cues in successful or unsuccessful retrieval processes have only been brought to the spotlight of experimental psychology after these groundbreaking experiments by Tulving and colleagues.

Baddeley (1999, pp. 174.) defines a retrieval cue as a fragment 'of the original learning that can be used to evoke the learning experience'. In this sense, retrieval cues aid all conscious remembering, and successful retrieval depends on the available cues (Dudai, 2002). In this dissertation I will focus more specifically on the phenomenon of episodic cued recall. Here, a to be remembered memory (referred to as target memory) is paired with a cue and later on the cue and/or the cue and fragments of the target memory are given to aid conscious recall of the target memory. As described in the next section, episodic memory retrieval critically depends on the effectiveness of cues in the environment (Tulving & Osler, 1968; Tulving & Pearlstone, 1966).
The cue driven nature of retrieval processes in human episodic memory

Since the earliest treaties of memory, the process of retrieving consciously accessible memories has been described as a dynamic interaction of cues and target memories. There are some components that are inherent in most theories of retrieval: A search process, a cue-target match, and an evaluation process. For instance, Norman and Bobrow (1979) observed retrieval processes in natural, ecologically valid situation. They collected a large amount of records on how their participants were trying to retrieve memories learnt a few weeks earlier. They noted that participants were 'struggling' to complete a partial specification of the target memories (which they termed description) by continuously monitoring the environment for novel and more useful cues. They identified three different processes in this process: first, individuals identify or create a retrieval specification with a partial description of the target memory and verification criteria. Second, some process must search for candidate target memories, and an other process should select those that might be good matches for the target. Third, an evaluation process should determine whether a given target under inspection is an appropriate candidate by the verification criteria.

The description of Norman and Bobrow is very similar to later cognitive and computational models of the retrieval process. In all these models the role of search process is to find the best match between the cue(s) and the target(s). The search itself might be an active search process (Norman & Bobrow, 1979; Raaijmakers & Shiffrin, 1983) or a process of spreading activation (J. R. Anderson, 1983) in a network of representations. Recall occurs when the target memory is matched (Norman & Bobrow, 1979), recovered (Raaijmakers & Shiffrin, 1981) or leads to ephory (Semon as cited by Schacter, Eich, & Tulving, 1978). This is paralleled with monitoring processes that evaluate the appropriateness of the target memory.
In addition, Tulving (1983) proposed that for any stimulus to be treated as a cue, the individual must be in a specific cognitive state: retrieval mode. This would imply that the three processes (search, match, evaluation) of retrieval are all supported by retrieval mode. Interestingly, all these processes have been supported by neuroimaging and electrophysiological studies. Correlates of matching or recovery have been termed in the neuroimaging literature as the retrieval success effect, whereas search and monitoring processes have been associated to neural processes related to retrieval effort (for summaries see Purves et al. 2008 and Rugg & Wilding, 2000). Importantly, in addition to identifying brain correlates of retrieval mode in the right frontopolar cortex (Düzel et al., 1999), the above mentioned component processes of retrieval have also been associated to distinct brain networks.

Having showcased the importance of cues in the retrieval process in episodic memory, let's turn now to how different theories conceptualize the way this retrieval process can be biased by altering the effectiveness of cues.

**The central theoretical debate: Interference accounts versus inhibitory accounts of forgetting and remembering cue-target associations**

How do changes in the effectiveness of reminders come by? During the twentieth century, several theories have been laid down to explain how cue-target associations might be changed. These theories were mainly concerned about the less trivial part, forgetting, but also postulated hypotheses about positive changes in cue-target associations. In the following paragraphs, I will briefly introduce the most influential views, based on two excellent and exhaustive reviews of these theories (Anderson and Bjork, 1994; Racsmány, 2003). This section will highlight some key points in the debate between interference and inhibitory accounts. These generally refer to a range of theories and frameworks that differ in one major point: whether they involve any cognitive inhibition or not. Here, for the sake of simplicity, I
will use the term 'inhibitory' for theories that involve an active inhibitory process. This is important to note, because such a use will exclude some theories that have the word 'inhibition' in their name, from the group of inhibitory theories. A schematic, and simplified illustration of the two theoretical lines can be seen in Figure 1.

**Figure 1.** Simplified dynamics of processes leading to changes in cue-target associations according to two groups of models: interference accounts and inhibitory accounts.

**Interference theories**

The dynamics extractable from all interference based theories can be modeled in a linear chain of events. This chain would involve a first encounter with different cue-target associations, i.e., a study event. Then, repeated encounters with any of these cue-target associations will induce a redistribution of one or some of the following attributes: absolute associative strengths, associative weights, memory strengths, cue-context associations, target-context associations. The type of redistribution depends on the given model's assumption. Some of these model assumption are depicted in Figure 2. In a third moment in this chain of events, at a later time point, memory performance will be determined by the redistributed pattern.
Figure 2. Assumptions of five interference based models, and one inhibitory model (bottom row, right) on how cue-target and cue-competitor associations change due to repeated encounters with cue-target associations. Note that in the case of the inhibitory model, only encounters involving competitive retrieval of targets is assumed to lead to the depicted changes. The upper left state illustrates the initial state of associative memories after a first encounter with a cue-target and a cue-competitor association. The six panels below indicate the assumption of specific models on how repeated encounter changes the pattern of associative memories (based on Anderson and Bjork, 1994). These changes will then determine later recall probabilities. Occlusion: reencountering a target in the presence of a cue will strengthen the association between these two, facilitating access to target memory at the cost of a decreased accessibility of competitors. Resource diffusion: These models assume that there is a limited amount of activation that can spread from one memory item to another, and reencountering cue-target associations will allocate more activation to these associations with less activation remaining for cue-competitor associations. Associative decrement: These models assume that association strength can both increase and decrease just like synaptic associations can be potentiated and depressed. In this framework, strengthening cue-target associations while not strengthening cue-competitor associations leads to an increase in cue-target associative strength, and a decrease in cue-competitor associative strength. Cue bias: These models assume that reencounters with a cue-target association will strengthen attributes of the cue that are related to the target. For instance seeing the 'fruit-banana' association will bias the meaning of the cue fruit to tropical fruit. This cue bias can lead to better memory for the target and worse memory for competitors that are not sharing the same attributes of the cue (e.g., apple). Context bias: These models incorporate the fact that each reencounter is set in a different context than the original study event. As such, reencounters with a cue-target association will create an context which biases later memory in favor of the cue-target association rather than the cue-competitor association. Item-level inhibition: Contrary to the above interference-based models, this one assumes that
competitors are actively inhibited when they interfere with recall of the target in the presence of the cue. This active inhibition leads to a decreased activation of the competitor at the item level. Not all possible models are illustrated. For instance, the episodic inhibition model could have been illustrated as a combination of the item-level inhibition and the context bias models. This model assumes that repeated retrievals of cue-target associations change the episodic pattern of representations of the study event. Memory items get inhibited and facilitated but this pattern of inhibition/facilitation is bound to specific episodic contexts. Retrieval will be biased in favor of a given episodic pattern, anytime the given episode is accessed. Note that only the item-level inhibition model assumes that changes occur at the item level independently of cues and/or contexts, and not only the level of associative links.

A now prototypical interference model is the Search of Associative Memory (SAM) model of Raaijmakers and Shiffrin (1981), further expanded by Mensink and Raaijmakers (1988) and then Shiffrin and Steyvers (1997), and Malmberg and Shiffrin (2005). This model assumes an active search process that is largely determined by association strengths between memory representations. According to the SAM theory, the retrieval process is fully cue-driven: there is no retrieval independent of cues. A background assumption, based on Atkinson and Shiffrin (1968), of the SAM model is that separate short and long-term stores (STS and LTS respectively) exist. For a target memory to be retrieved from LTS, a search plan is constructed, then cues are assembled in the STS. Next, the search is first restricted to an LTS subregion in the association networks. Here, potential targets are sampled based on a ratio-rule, i.e. based on the strength of their association to the cue (or assembly of cues) relative to the strength of association of other targets to the same cue in the same set. If sampling is complete, recovery (i.e. recall) of the target memory occurs based on the absolute strength of the cue-target association. Finally, recovered memories are evaluated, and this evaluation gives input to the decision about reiterating the search process or to terminate it (together with either providing an output or not).

Contrary to the model of Norman and Bobrow (1979), the search processes in SAM operate in serial steps in cycles until a decision to stop the search process is made. It is assumed that sampling and recovery are automatic processes, whereas the construction of the search plan, the evaluation of the recovered target memory, as well as decision-making about when to stop, and what to output, is aided by control processes.
To explain diverse memory phenomena, such as interference related forgetting, the SAM model was further specified by Mensink and Raaijmakers (1988). This model incorporates changes in context due to passing of time and other factors, and hence is referred to as the contextual fluctuation model. The important change in this model is that it assumes that the overlap between the context of encoding and the context of retrieval affects the associative strength between the given the cue and the target memory. Learning occurs through strengthening of associations, with each recovery leading to incremental learning, i.e. strengthening the association.

Interestingly, this version of the SAM model, explains memory phenomena without reference to any suppression-like processes. Earlier theories, such as that of Melton and Irwin (1940) included the concept of unlearning. This process resembled and was partly based on the concept of extinction from research on conditioning: unlearning was thought to occur when a memory associated to the cue was not retrieved because it was not a target of the search process.

A further development in the SAM theory occurred late in the nineties, by the development of the Retrieving effectively from memory (REM) model (Shiffrin & Steyvers, 1997) and then its further specifications, named the SAM-REM model, to include several novel model assumption by Malmberg and Shiffrin (2005). The most important one in these later modifications for the current discussion is that these models assumed that different types of reencounters with cue-target material lead to differential effects on sampling and recovery. Importantly, it is assumed that context-cue-target associations are strengthened more by retrieval than by repeated exposure to the same cue-target associations.

This version of the model is in line with a more theory driven framework describing changes in cue-target associations, the episodic inhibition framework (Racsmandy & Conway, 2006). This framework is explicit about the distinction between the role of semantic and episodic
memory in episodic retrieval. It assumes that activation of memories in the semantic network can vary independently of their activations in episodic memory. For instance, a memory might be more or less accessible after a given manipulation that affects episodic, context driven retrieval, together with being more activated (in both cases) in the semantic network. In brief, the framework suggests that each act of episodic retrieval encodes (and reencodes) a pattern of activation/inhibition of memories. This pattern, that is strictly bound to the currently retrieved episodic context will determine the accessibility of memories. This is one of the theories that have 'inhibition' in their name, but will not be considered as inhibitory in this dissertation. Inhibition in this framework simply means a negative change in the associative weights between episodic context-dependent cues and target memories (M. Racsmány, September 18, 2013).

Because the SAM model claims to explain all explicit memory phenomena without reference to such processes as inhibition and unlearning, it will be a good candidate to contrast inhibitory theories with.

**Inhibition theories**

The dynamics extractable from inhibitory theories involve one more step. As seen in Figure 1, these models assume that after the study event, repeated encounters with some cue-target associations might induce competition: Interference between reencountered targets and competitors (targets that have been associated to the same cue during the study event, but are not currently reencountered) has to be resolved. At this moment, cognitive control mechanisms come into play to resolve competition (interference) and inhibit the competitor's representation. The bottom right panel of Figure 2 depicts the assumption of inhibitory models. After inhibition, later memory performance is determined by the redistributed activations, with competitor representations being inhibited.
A now prototypical inhibitory theory is the inhibitory control theory proposed in its first form by Anderson and Bjork (1994), and then further specified in Anderson (2003) and, among others, Anderson (2008). This theory conceives the retrieval to be an active, effortful search process that is guided by cognitive control. It not only takes into account the strength of associations between memories, but assumes that memory items per se have memory strength. Accordingly, retrieval success is a function of cue-target associations and target memory strengths. This is important because this theory also assumes that the search process is cue driven, but also assumes that accessibility of target memories can be modified by changes that affect only target memories, and not associations. This cue-independence of target memory accessibility is a crucial point of disagreement between interference and inhibitory theories and will be discussed further in later chapters.

The inhibitory theory also assumes that selection of targets is not fully automatic and involves cognitive control. Indeed, it is an intuitive assumption that when targets, related to a cue with roughly equal association strength, have to be selected against, some kind of control process must resolve this interference. A following assumption here is that the interference resolving process is inhibition mediated by cognitive control. The theory assumes that target memories selected against (competitors) during an effortful target memory retrieval get inhibited. These memories being in a state of inhibition are then less retrievable to any cue in the future. This account is at sharp contrast to the interference account. The next section will briefly summarize the crucial differences.

**Key differences between interference and inhibitory accounts**

1) The locus of the effect: Interference accounts assume that associations change, inhibitory accounts assume that memory items also change. 2) Cue-independent changes in memory accessibility: Whereas interference accounts assume that there are no cue-independent changes in the accessibility of memories, inhibition theories suggest that activation of target
memory traces can decrease independently of cues. 3) Retrieval dynamics: Interference accounts assume that the retrieval process determines recall. In contrast, in inhibitory theories representational states are central to determining accessibility of memories. 4) The role of cognitive control: Although interference theories often assume that some form of cognitive control must play a role in the retrieval process (e.g. in the SAM model: planning the search process, selecting and changing cues, evaluating recovered information, and decide to output and/or stop), these processes are independent from matching (e.g. from sampling and recovery in the SAM model, or from the extinction-like unlearning of associations in the theory of Melton and Irwin (1940)). In contrast, in inhibitory models, the matching process itself is guided by cognitive control. 5) Formalization: the prototypical interference model of SAM is a mathematical model refined several times during its history. Inhibitory accounts are rather theoretical frameworks that have not been formalized. A computational model of changes occurring to cue-target associations and target memory representations involving inhibition has been described by Norman, Newman, and Detre (2007). However, this model is a model specifically developed to explain retrieval-induced forgetting (RIF), and is not a general model of retrieval processes.

The specific questions this dissertation covers are derived from the first four of these key discrepancies between the two lines of theories. The thesis is an attempt to furnish evidence to the theoretical debate between accounts of remembering and forgetting cue-target associations.

Outline of the dissertation

In an effort to better understand how cue-target associations can be changed, the first part of dissertation investigated the negative mnemonic effect of stopping retrieval. Research on this topic was largely motivated by the works of Sigmund Freud, and is part of an effort to test hypotheses derived from psychoanalytic theory of repression (see Erdelyi, 2006). Therefore,
after a brief review on the literature testing psychoanalytically motivated hypotheses, I will introduce the related concept of inhibition as used in cognitive psychology, narrowing the discussion on inhibition in memory phenomena. I will dwell on describing one of the paradigms, the think/no-think paradigm (Anderson & Green, 2001) most widely used in recent years to understand the inhibitory effects of stopping retrieval, and review the theoretical insights we have gained from both behavioral and neuroimaging results using this paradigm. Then, I will introduce briefly our study that was planned to test a hypothesis proposed by Anderson and Green (2001) stating that humans are capable of stopping retrieval of memories so that these memories become unavailable to any memory cue in their environment (i.e. that forgetting is cue-independent).

The two studies in the second part of the dissertation focused on the negative effect of competitive retrieval. A well-documented finding in the memory inhibition literature is that forgetting of memories related to a cue occurs after repeated retrieval of a target memory that is associated to the same cue. This phenomenon is called retrieval-induced forgetting (RIF) (Anderson, Bjork, & Bjork, 1994). Following an introduction of the phenomena, again, I will present one paradigm that has been most widely used to investigate RIF, the retrieval practice paradigm (Anderson et al., 1994), and review theoretical explanations of RIF in the light of existing data. The second and third study tested hypotheses of the inhibition theory (Anderson, 2003) to further understand the nature of RIF.

The fourth study, in the last part of the dissertation focused on the neural background of the positive effect of repeated retrieval, the testing effect (Roediger & Karpicke, 2006a). Because this phenomenon is somewhat fresh to the study of memory, most of the review at the beginning this chapter will be a detailed description of the effect itself, and will provide only a short overview of the theories (or rather theoretical frameworks) explaining it. The study presented in this part is one of the first attempts to understand the neural correlates of the
testing effect, and as such is a mostly explorative addition to the data available on the testing effect. Nevertheless, its theoretical contributions to our understanding will also be discussed.

Altogether, based on the four studies reported in the dissertation I will suggest that both voluntary and involuntary associative forgetting, just as the positive effect of repeated retrieval, is driven by redistribution, possibly by cognitive control, of activation patterns during the act of retrieval.
Memory inhibition in cued recall

Inhibition theories inspired by the works of Freud

Karl Popper was among the first scientists to criticize the psychoanalytic theory of Freud for being verificationist in its approach to theory (Dienes, 2008). The falsificationist approach advocated by Popper, and the basis of today’s inferential statistics approach in experimental psychology, claimed that Freudian theory is not scientific because it only looks for verifying its statements. In contrast, Grünbaum (1985) argued that psychoanalytic theory does lead to hypotheses about human psychology, but these are rarely tested. One popular target of such tests has been the process of repression or suppression, a term generally coined for mechanisms that decrease conscious access to mental representations (for a review on the use of the term see Erdelyi, 2006).

Two paradigms motivated directly by the possibility to test the theory of Freudian repression (1932/1977) are worth mentioning. In the thought suppression paradigm (Wegner, Schneider, Carter, & White, 1987) participants are instructed to speak out loud everything that comes to their mind during a five minute period. The crucial manipulation is that they are asked to try not to think about a white bear (suppression) during these five minutes. During a subsequent period of five minutes, again instructed to say out loud their thoughts, they are explicitly asked to think about the white bear (expression). The number of thought intrusions decreases during suppression (although it never reaches zero), however during the expression period intrusions are substantially increased as compared to an expression period not preceded by a suppression period. According to Wegner (1994) two opposing processes affect conscious access to mental content. A monitoring process constantly searches for contents that are inconsistent with a desired mental state (i.e., in the suppression period, this would be the white bear), and an operating process searches for contents consistent with the desired mental
state. The monitoring process is proposed to be less demanding in terms cognitive capacity, therefore it is suggested that whenever attentional or cognitive control capacity is reduced, monitoring will have the ironic effect of having the unwanted thought pop up in consciousness. Although the theoretical interpretation and clinical applicability of this rebound effect has been debated, the phenomenon has been replicated by a vast amount of studies (Rassin, Merckelbach, & Muris, 2000). The thought suppression experiment developed by Wegner at al. (1987) might serve as a model to study the return of repressed thoughts (Freud, 1932/1977).

Another task, the think/no-think task developed by Anderson and Green (2001) provides a tool to investigate changes in the accessibility of memories suppressed. In this task, participants learn cue-target associations and then when shown some of the cues, they are instructed not to think about the target associated with the cue. The critical test assesses accessibility of these suppressed target memories. As this task is used in the first study of the dissertation it will be described in detail later in this chapter. In the next sections, I will briefly review how the concept of inhibition has surfaced in the broader context of experimental cognitive psychology. A major argument against the use of inhibition as a cognitive concept will be briefly mentioned here. Later on several other arguments will be addressed in more detail, when introducing theories explaining memory inhibition phenomena investigated in the dissertation.

**Inhibition in cognition**

Inhibition as used in psychology is an elusive and fuzzy concept. It seems that besides the use of the word in naive psychology, its widespread use in scientific theories of cognition also owes a lot to its analogy with the concept of inhibition in physiology (C. M. MacLeod, Dodd, Sheard, Wilson, & Bibi, 2003). The analogy of facilitatory and inhibitory links between neurons is so compelling that it has become commonplace to explain association (and
inhibition) of ideas with association (and inhibition) of physiological entities in the brain (Mitchell et al., 2009). The question of whether the concept of cognitive inhibition is necessary at all to explain human cognition have led to the emergence of inhibitophil and inhibitophobe (C. M. MacLeod, 2007) scientists. The latter claim that inhibitophil researchers, in a circular way, tend to tag below baseline performance as inhibition, and then declare any below baseline performance as evidence for inhibition. Inhibitophobe researchers rather like to replace the word inhibition with interference. These researchers do not fully reject the use of the concept of inhibition, but rather suggest a different definition that does not include inhibition of ideas as such. We will see examples of such argumentations when discussing the logic of research in the field of memory inhibition.

The term inhibition in this field is overgeneral, and refers to a wide range of phenomena observed in studies of attention and cognitive control (Aron, 2007). Indeed, inhibition is generally measured as some negative performance change compared to a baseline condition in stimulus processing.

**Memory inhibition**

It is not surprising then that the common finding in paradigms of inhibition in memory is a decreased accessibility of memories compared to a baseline. A strong inhibitophil view proposed by Anderson and colleagues (e.g., Anderson & Levy, 2007), claims that memory inhibition is a process which actively decreases activation of memories that are irrelevant and are interfering with memory processing at a given time. They suggest that this inhibitory process decreases accessibility of these memories irrespective of later retrieval cues and contexts. From an inhibitophobe view (e.g., C. M. MacLeod, 2007) memory inhibition is simply stopping or overwriting a memory process, partially or fully, with or without conscious intent (p. 5). In the following two chapters I will describe two memory inhibition paradigms widely used to test predictions of both sides.
The think/no-think paradigm

This procedure was devised by Anderson and Green (2001) to measure the effect of voluntary suppression of unwanted memories. To grasp the rationale of the procedure it is important to note that the think/no-think (TNT) procedure was devised as an analogy of the go/no-go paradigm. In this procedure two types of stimuli occur in a rapid sequence, and participants are instructed to make a go response for one type of stimuli and a no-go response (i.e., refrain from responding) for the other. Usually, go stimuli occur more often than the no-go stimuli, in order to increase readiness to respond. The ability to withhold the response when a no-go stimulus is presented is usually taken as an index of inhibitory abilities.

The procedure of Anderson and Green (2001) started with learning 40 moderately associated cue-target word pairs until a 50% criterion level. In this learning to criterion, participants were first presented with all 40 word pairs (each pair presented once on a computer screen for five seconds), and then performed a cued recall on all 40 pairs once (with the cue being presented and participants being instructed to recall the target associated with it). This cued recall phase was repeated until criterion level was reached. Then, during a critical manipulation (the think/no-think phase), 30 cues were presented again, with three different repetition times: ten cues were presented once, ten eight times, and another ten 16 times. Importantly for five cues in each repetition condition, participants were instructed to recall the target (think condition), whereas for the five other cues in each repetition condition, participants were instructed not to think about the target associated to the cue (no-think condition). Importantly, cues always remained on the screen for the entire 4 sec trials, and participants were instructed to fixate on the cue. In the no-think condition they were also explicitly instructed not just to refrain from saying out loud the answer, but to avoid thinking about the target word.
The question was whether a final test would detect a significant forgetting of targets that had been suppressed (no-think items) during the critical manipulation. Indeed, cued recall involving all 40 cues after the think/no-think phase revealed that with repeated suppression attempts, cues became less effective in reminding participants of target words as compared to cues that did not appear during the think/no-think phase, and therefore served as baselines in the experiment. The critical statistical contrast, cued recall of no-think word pairs suppressed 16 times versus cued recall of baseline pairs (suppressed 0 times), was significant also in control conditions that ruled out the effect experimenter bias (offering monetary reward for each retrieved target word in Exp 2, and telling a reverse expectation to participants in Exp 3). Crucially, impairment of suppressed memories persisted in conditions when the target was not cued by the cue learnt during the experiment (same probe test), but with a cue, that had not appeared in the experiment until the final test, but was semantically related to the target word (independent probe test). According to the authors, forgetting on independent probe tests provides direct evidence for the active suppression/inhibition of memory representations that decreases the activation of target memories. Note that forgetting on the same probe test could be explained by non-inhibitory mechanisms, such as the association of a novel target to a given cue, or changing the representation of the cue, or of the cue-target associations. These are unable to explain the forgetting effect when tests use independent probes.

**Logic of research using the TNT paradigm**

From a theoretical point of view, two lines of research followed the work of Anderson and Green (2001). One is mostly an explorative research line. Most work in this line tried to understand the limits of this suppression/inhibition effect. How far the analogy with repression can be extended? Does the effect occur for visual and/or emotional material? Are some populations more able to suppress memories than others? Many of these studies, briefly overviewed in the next section, commit the fallacy of circular argumentation mentioned above
(C. M. MacLeod, 2007): Taking the TNT task as a measure of memory inhibition, even without using the control conditions and/or the independent probe test used in the original paper by Anderson and Green, the magnitude of the effects are considered as evidence for or against the ability to inhibit different kinds of memories in different kinds of populations. Interestingly, as we will see, neuroimaging studies provide a great example of such a fallacy. Other researchers have been more concerned with the core conclusions drawn from the suppression effect found in the TNT tasks. Besides replication attempts, work in this line tried to put the inhibitory hypothesis formulated by Anderson and Green (2001) to test. The first study in this dissertation is part of these latter attempts.

**Major behavioral findings**

In this overview I will refer to the effect of interest, the difference in recall of baseline and no-think items as the TNT effect. The TNT effect has been replicated in cued recall tasks (e.g., Depue, Banich, & Curran, 2006; S. Hanslmayr, Leipold, Pastotter, & Bauml, 2009; Hertel & Calcaterra, 2005), and recognition tasks (Kim & Yi, 2013; Marx, Marshall, & Castro, 2008; Waldhauser, Lindgren, & Johansson, 2012), and not just with words, but also with visual material (Depue, Curran, & Banich, 2007; Kim & Yi, 2013). This shows that the TNT effect is robust and generalizes to many types of tests and materials.

For its relevance to the phenomena of repression, it was also important to show that the TNT effect is produced not solely for neutral memories, but also for emotionally loaded material (Depue et al., 2007), and autobiographical memories (Noreen & MacLeod, 2013). Accepting the rationale that the TNT effect is suggestive of difficulties in suppressing unwanted memories, a literature review suggests that this type of memory inhibition is intact among populations with schizophrenia (Salamè & Danion, 2007), in highly dissociative individuals (Wessel, Wetzels, Jelicic, & Merckelbach, 2005), in depression (Joormann, Hertel,
Brozovich, & Gotlib, 2005; but see Hertel & Gerstle, 2003), but impaired in ADHD (Depue, Burgess, Willcutt, Ruzic, & Banich, 2010).

Besides emotion and clinical conditions, other moderating effects and conditions have also been suggested to influence the TNT effect. Importantly, these all relate to the role of cognitive control functions in producing the TNT effect. For instance, the results of Hanslmayr and colleagues (2009) suggest that when people anticipate a no-think cue, the TNT effect can be enhanced. Individual differences in cognitive control also seem to matter. This is indirectly suggested by developmental studies showing that the effect seems to improve from childhood to early adulthood (Paz-Alonso, Ghetti, Matlen, Anderson, & Bunge, 2009), and then decline with age late in adulthood (Anderson, Reinholz, Kuhl, & Mayr, 2011; Murray, Muscatell, & Kensinger, 2011). In a related vein, cognitive control abilities, as measured by the stop-signal task, have been shown to correlate with the TNT effect (Depue et al., 2010), and electrophysiological correlates (see the next section) have been also shown to overlap during the two tasks (Mecklinger, Parra, & Waldhauser, 2009). In contrast, go-no go performance did not predict the TNT effect in a later study (Paz-Alonso, Bunge, Anderson, & Ghetti, 2013). The above mentioned finding of no TNT among ADHD patients in the same Depue et al. (2010) study is also in line with the proposed role of cognitive control in TNT.

**Major neuroimaging and electrophysiological findings**

The first functional magnetic resonance imaging study of TNT (Anderson et al., 2004), had participants scanned during the TNT phase, and found several results supporting the involvement of cognitive control processes in the TNT effect. These authors found that the magnitude of the TNT effect was associated to the magnitude of activations in the dorsolateral prefrontal cortex (DLPFC) bilaterally, the left ventrolateral prefrontal cortex (VLPFC), the pre-supplementary motor area (pre-SMA), the dorsal premotor cortex (PMDr), and the intraparietal sulcus (IPS). Of these, the DLPFC activation was not only related to the TNT
effect detected by same probes, but also with that found with independent probe tests. A subsequent memory analysis revealed that efforts to suppress subsequently forgotten no-think items were associated with greater hippocampal activations than unsuccessful attempts of suppression. The authors argued that this might have reflected intrusions of no-think target memories, which might have then triggered the suppression by the DLPFC. An additional finding supporting this assumption was that the activation difference in bilateral DLPFC between subsequently remembered and subsequently forgotten no-think target memories correlated with activation differences in the right hippocampus between subsequently remembered and subsequently forgotten no-think target memories.

The above study was followed by a similar one (Depue et al., 2007) that used emotional material and included a fixation baseline during the TNT phase. The baseline condition allowed for identifying the specific contribution of thinking and not thinking about target memories to changes in brain activity. (In the Anderson et al. study discussed in the previous paragraph, results were based on think vs. no-think contrasts which does not allow for disentangling the effect of think and no-think attempts). Here, a bunch of right prefrontal regions (BA8, 9/46, 47, and 10) were more active during suppression than during retrieval of think targets during the TNT phase. The difference was due to both larger than baseline activations during processing no-think cues and a smaller than baseline activation during processing think cues. Areas in the visual cortex (BA 17, 18), the fusiform gyrus (BA37) as well as the pulvinar nucleus in the thalamus exhibited decreased activations during no-think trials as compared to think trials. Again, this was further qualified by contrasts to baseline, with activations being smaller than baseline for no-think trials, and being larger than baseline for think trials. In brief, prefrontal areas, related to cognitive control, increased their activations during suppression, whereas posterior areas, related to processing sensory representations of memories, decreased their activations during suppression.
More fascinating was their correlational analyses for suppression trials only, calculating the linear relationship between activations in pairs of areas reported above for the four time quartiles of the TNT phase (each including three suppression trials). This analysis revealed that only the BA10 area showed significant above baseline activations in all four quartiles. Importantly, when its (positive) correlation with activations in the middle frontal gyrus (MFG) (here, BA9/46) reached its peak in the third quartile, in the next quartile (the fourth quartile) the (negative) correlation of activation in the MFG and activation in the amygdala and hippocampus reached its peak as well. Similarly, when the (positive) correlation of BA10 and inferior frontal gyrus (IFG) (here BA47) reached its maximum in the first quartile, this was followed in the second quartile by the peak of (negative correlation) between the IFG and the fusiform gyrus and the pulvinar.

Based on these findings Depue and colleagues suggested that prefrontal areas react to cognitive control demands during processing of no-think cues. This reaction results in the downregulation of the amygdala and the hippocampus, areas known to be involved in processing of emotional memories.

These studies were further corroborated by a study by Benoit and Anderson (2012) that used dynamic causal modeling to compose possible models linking the DLPFC and the hippocampus - to areas where activations were modulated by suppression trials in the same way as in the above studies. A Bayesian model selection showed that their data was by far best explained by a model that involved a modulatory effect of suppression on bidirectional links between the two areas (both top-down modulation from DLPFC to hippocampus and bottom-up modulation from hippocampus to DLPFC). Importantly, according to the model, increasing DLPFC activity was coupled with decreasing hippocampal activation during suppression trial. Furthermore, when the authors analyzed the data median split according to how successful participants were in suppressing no-think targets, effective connectivity was
stronger for those participants who were more successful suppressors. A functional connectiviti
analysis (Paz-Alonso et al., 2013) led to similar results: During the TNT phase, connectivity in a network, comprising the hippocampus, the DLPFC, as well as mid- and posterior cingulate areas was stronger in the brain of successful suppressors for intentionally forgotten (successfully suppressed target memories) than for unintentionally forgotten memories (think target memories).

Electrophysiological studies provided further insight into the timing of successful suppression, although the results seem mixed. Contrasting ERPs during think and no-think trials in the TNT phase of their experiment, Bergström and colleagues (Bergström, de Fockert, & Richardson-Klavehn, 2009) showed that successful suppression of no-think memories was associated with reduced positivity at left frontocentral sites 300 ms after cue onset, and then at parietal sites 500-600 ms after cue onset. The latter shift was also observed in a study using recognition at final test (Waldhauser et al., 2012). Note, that these positivity effects (here reduced by suppression) are usually associated with feeling of familiarity and recollection, respectively (Rugg, 2004). Although this reduction of positivity was not found by Mecklinger et al. (2009), they showed that both suppression during no-think trials and stopping responses in a stop signal task enhanced N2 component of the ERP signals, a result also replicated in the recognition study of Waldhauser et al. (2012). A recent study (Chen et al., 2012) found no decrease in positivity over parietal areas for neutral stimuli. Rather, these authors found that this positivity was increased for emotional, negative stimuli with a peak at 800 ms post onset.

Mainstream interpretation of findings

Published reviews of results from the TNT paradigm (Anderson & Huddleston, 2012; Bäuml, Pastötter, & Hanslmayr, 2010) offer a coherent picture that also seems to be supported by findings reported since the publication of these reviews, touched upon above: Whenever cues
trigger the retrieval of unwanted target memories (e.g., Anderson et al., 2004; Depue et al., 2007), individuals with good cognitive control capacities (Bergström et al., 2009; Mecklinger et al., 2009; Paz-Alonso et al., 2009) are capable of suppressing this retrieval process. This suppression then renders target memory representations less accessible, i.e. inhibited. Suppression is thought to be triggered by hippocampal signals of recollection to the DLPFC (Benoit & Anderson, 2012; Paz-Alonso et al., 2013), which downregulates areas involved in storing memory representations (Anderson et al., 2004; Benoit & Anderson, 2012; Depue et al., 2007). This view is in line with the inhibitory control hypothesis suggesting that executive control processes lead to active inhibition of memory traces (e.g., Levy & Anderson, 2008). It would also follow that the TNT procedure is a potential paradigm suitable to investigate mechanisms of repression.

**Problems with the mainstream interpretation**

While many have criticized the use of the analogy between the direct memory suppression as interpreted above and Freudian repression (Garry & Loftus, 2004; Kihlstrom, 2002), the discussion that follows here is rather focused on evaluating the available evidence in favor of the inhibitory control hypothesis (Levy & Anderson, 2008).

**The issue of cue-independent memory inhibition**

Although the data supporting the inhibitory control hypothesis seems to be extensive, it must be pointed out that the direct suppression of memory traces would be supported only by TNT effects observed using independent probe tests (Anderson, 2003). Behavioral replications of the independent probe TNT effect have been scarce (Anderson et al., 2004, 2011; Benoit & Anderson, 2012; Bergström et al., 2009; Paz-Alonso et al., 2009), and most of the results reported above and in the two aforementioned reviews come from studies using same probe tests. Importantly, only three neuroimaging studies (Anderson et al., 2004; Benoit & Anderson, 2012; Bergström et al., 2009) reported a TNT effect with independent probes.
Others did not use independent probes, or did not get a significant effect. Interestingly, three neuroimaging studies (Bergström, Velmans, de Fockert, & Richardson-Klavehn, 2007; A. J. Butler & James, 2010; Mecklinger et al., 2009) analyzed correlates of suppression in experiments that did not produce a behavioral TNT effect at all. These studies based their analyses on subsequent memory differences (remembered vs. forgotten) in no-think target memories that were successfully learned during the initial learning phase.

As pointed out by Anderson (2003), TNT effects observed in paradigms using same probe tests might be explained by mechanisms other than inhibition, such as unlearning the cue-target association, or substituting a new target in the place of the no-think target that needs to be suppressed. Indeed, the fact that thought substitution can increase the TNT effect, has been noted also by proponents of the inhibitory account (Anderson & Huddleston, 2012). However, these authors argue that this increase in the TNT effect observed in studies that instruct participants to substitute no-think memories might not necessarily mean that inhibition is not at play. They argue that (1) participants need to select a strategy suitable to suppress memories, which takes time and effort, until a good strategy is selected. When they are prompted to use a predefined strategy, no matter what this strategy is, the consistent use of it might enhance the effect. (2) Substitution is a simpler instruction than suppression, and therefore participants instructed to suppress are more likely not to comply with experimental instructions than participants instructed to substitute. (3) Substitution may lead to retrieval-induced forgetting (RIF) of the substituted memory (see the chapters on RIF). Nevertheless, the initial argument (Anderson, 2003) still holds that substituting B targets in an A-B association with C targets might simply cause later forgetting of B items due retroactive interference and/or output interference at test.

The absence of independent probes renders all findings with clinical populations inconclusive. Because none used independent probes at the final test, the effects found in those studies
might be explained by output interference at test. This suggestion is supported by the findings of Joorman et al. (Joormann, Hertel, LeMoult, & Gotlib, 2009) who showed that whereas depressed participants (contrary to their earlier finding (Joormann et al., 2005)) did not show the TNT effect, 'training' them with thought substitution instruction (to substitute a novel target word in the place of the previously learnt target word) lead to a significant TNT effect.

**Getting the TNT effect with independent probes without suppression**

The strongest argument for a concept of memory inhibition that involves decreased activation of memory traces at the item level is the few instances of cue-independent forgetting that have been found. The argument states that in case a target memory is less accessible than baseline when tested with a novel cue that was not experimentally associated to the target, then this decrease cannot be explained by changes occurring at any level of the cue-target association. A potential blow to this argument comes from Tomlinson, Huber, Rieth, and Davelaar (2009) who showed that associating a simple behavioral response to no-think cues might lead to the TNT effect using independent cues. In their experiment, instructing participants to press the ENTER button as fast as possible whenever they saw a no-think cue led to a significant cue-independent TNT effect. This effect was no different from the TNT effect in a regular no-think condition included in their design, where participants were instructed not to think about no-think targets. In the enter pressing condition, of course, one would not expect memory inhibition of item level target memory representations to occur. Tomlinson et al. argued that in these cases, as well as in instances where participants are instructed to suppress target memories, a two stage model of memory retrieval, such as the search of associative memory model (SAM) (Shiffrin & Raaijmakers, 1992), including a sampling and a recovery phase of memory traces, can explain cue-independent forgetting with reference to interference processes only. This explanation suits well the cue-independent forgetting found in one study using the thought substitution procedure (Benoit & Anderson, 2012).
The issue of effect size

Besides failed or missing replications of the TNT effect with independent probes, the TNT effect with same probes also seems to be small and vulnerable to several factors that vary between individuals, a fact acknowledged by several authors (e.g., Anderson & Huddleston, 2012; Bergström et al., 2007; Mecklinger et al., 2009; Waldhauser, Johansson, Bäckström, & Mecklinger, 2011), and the effect was often not found (Bergström et al., 2007; Bulevich, Roediger, Balota, & Butler, 2006; A. J. Butler & James, 2010; Mecklinger et al., 2009; Waldhauser et al., 2011).

The issue of the missing baseline in neuroimaging studies.

All but one neuroimaging study (Depue et al., 2007) used a baseline condition during imaging in the TNT phase. All other studies contrasted trials that involved recollection (successful think trials or not successful no-think trials) with trials that involved something else (unsuccessful think trials or successful no-think trials). Because these contrasts are highly similar to remember-forget contrasts, or recall-no recall contrasts, it is not surprising that they led to finding decreased hippocampal activations.

Explaining the DLPFC-hippocampus inhibitory orchestra with default network processing

The most interesting finding emerging from neuroimaging studies of the TNT effect seems to be the increased DLPFC activations coupled with decreased hippocampal activations. However, this might simply mean that during no-think trials participants engage in effortful processing which might be finding a novel target or searching for strategies, or trying to shift attention to something else, etc. One interesting, and highly plausible explanation for the increased DLPFC activation coupled with decreased hippocampal activations, unrelated to inhibition, was offered by studies that found correlations between hippocampal and DLPFC activations in recall tasks. One highly relevant study (Israel, Seibert, Black, & Brewer, 2010)
showed that during difficult retrieval trials the DLPFC increases its activation during monitoring of the retrieved responses, and this is correlated with suppression of activity in the hippocampus. A subsequent study from the same lab (Reas, Gimbel, Hales, & Brewer, 2011) showed that hippocampal activation was suppressed during effortful search for targets during cued recall, and this suppression was most pronounced for unsuccessful retrieval attempts. These authors discussed their findings in the light of previous work showing that the default mode network is disengaged during any task that requires cognitive control or effort (Raichle et al., 2001). These networks include the mediotemporal lobe (MTL), and do not include the DLPFC. In an earlier paper, Aron (2007) had already made the suggestion that decreases in the activation in the hippocampus during no-think trials might simply mean that the no-think task is more difficult than the think task, and thus activation patterns resulting from contrasting no-think versus think trials might be similar to hard task versus easy task contrast which will show activations in the resting state default mode network. These findings fit well to the negative coupling of DLPFC and hippocampus in neuroimaging studies of the TNT effect, and are able to explain all the data without reference to any memory inhibition phenomena.

**A cautious interpretation**

What seems to be corroborated from all these findings, is that some people are seemingly able to stop cued recall of a wide range of material, resulting in decreased accessibility of target memories associated to the cue. It also seems fair to state that this ability depends to some extent on cognitive control mechanism. What cannot be stated however, is that forgetting in these instances is caused by inhibition of target memory representations.
A new perspective on the TNT effect

However compelling the analogy with neurophysiological inhibition is, it seems that target memory representations do not get inhibited in the TNT paradigm. In an early study of retrieval cues, Tulving and Thomson (1973) showed that participants might not even be aware of having learnt a certain target, even though they readily recall it when cued with the cue paired together with it during previous learning. As Baddeley (1999) noted, this might be simply because what we learn in such situations is a specific cue-target association. Mitchel et al. (2009) argue similarly that human learning in general is about learning propositions. In the case of cued recall learning this would mean that a certain cue elicits retrieval of a certain target memory, and not that we learn to link the cue and the target. Accordingly, a mathematical analyses of distributed memory models (Kahana, 2002) provided support for the associative symmetry hypothesis over the independent association hypothesis. The former posits that learning cue-target associations involves creating new composit memory representations, whereas the latter conceives cue-target associations as two independent unidirectional links between two memory items.

From this perspective, it seems clear that cue-target associations can be changed by stopping associative retrieval. Study 1 aimed at answering whether the changes in the cue-target association induced by no-think trials lead to symmetrical changes in the accessibility of the cue and the target.

As independent probe forgetting is usually undetectable (even if it existed), another goal of Study 1 was to use a novel approach to test the inhibitory control hypothesis, by testing cue-target associations and reversing these tests so that the target served as a cue to retrieve the cue, which in turn became a target. A general inhibitory account would predict that as the target is inhibited by memory inhibition, even if it is used as a cue, it should lead to decreased memory performance.
This chapter described a phenomenon whereby stopping retrieval of cue-target associations lead to forgetting of target memories: the TNT effect. It also addressed theories explaining the effect with a major point of reference being the inhibitory control theory. The next chapter describes another phenomenon where modifying cue-target relationships leads to forgetting. In this following chapter, the effect investigated is going to be the forgetting of cue-competitor associations occurring due to repeated encounters with cue-target associations.
The retrieval practice paradigm

Retrieval-induced forgetting (RIF)

In our everyday life, we often face situations where a cue initiates the retrieval of a well learned memory which is no longer appropriate. For instance, the PIN of our old credit card might pop into mind several times before we manage to focus our retrieval on the PIN of our new card. Our memory adapts to this need, and presenting the cue ‘PIN’ on the ATM screen will eventually lead to recall of the appropriate digits. Some time later the old PIN might be completely forgotten. Forgetting of the old target memory occurs after repeated retrieval of the new target memory, hence the term retrieval-induced forgetting (RIF) for this phenomenon (Anderson et al., 1994).

The retrieval practice paradigm

Anderson et al. (1994) developed the retrieval practice paradigm to separate the effect of interference at time t, and the effect of the resolution of this interference apparent at time t+1. In this paradigm participants usually study category–exemplar pairs (e.g. fruits – prune, drinks – gin, fruits – banana), then practise retrieval of half of the exemplars from half of the categories with a specific category plus word-stem cue (e.g. fruits – ba______). After a short delay (usually 5-20 minutes) all exemplars from all categories are tested with a category plus word-stem cue (e.g. fruits – p______, fruits – b______, drinks – g______). Final recall of practised items (i.e., banana) is typically enhanced while recall of non-practised items from practised categories (i.e., prune) is worse compared with the baseline (i.e., gin). These findings are usually referred to as practice effect and RIF effect, respectively.
Theories explaining RIF

Two major opposing theoretical lines compete for modeling RIF. Noninhibitory models that incorporate facilitatory processes only argue that successful retrieval of a memory changes the associative network between the cue, the target memory, and interfering memories so that the probability of a later recall of a retrieved item is increased. At a later memory output, this increased probability biases the recall processes toward recall of the strengthened memory leading to forgetting of other, related memories (Raaijmakers & Shiffrin, 1981; Shiffrin & Raaijmakers, 1992; for a review of similar accounts see Anderson & Bjork, 1994). A contrasting view states that there is an executive control mechanism that should resolve memory interference on the first place (i.e. resolve interference when we have to recall for instance the new PIN instead of the old one for the first few times). This view posits an active executive control mechanism that acts directly to inhibit the representation of the memory trace that is no more appropriate to recall (Anderson & Spellman, 1995; Anderson, 2003). In comparison with the above two theoretical lines, a third, less specified inhibitory framework was proposed by Racsmány and Conway (2006). Using the concept of episodic inhibition, they suggested that retrieval practice establishes a pattern of activation/inhibition over the contents of a given episodic memory. When an episodic memory is accessed, the same pattern that was initially encoded over its contents is reactivated. Accessibility of the contents of the episode will vary according to the reactivated pattern. This implies that if a pattern of activation/inhibition is created during the study phase of a retrieval practice experiment, it is this pattern that will be reactivated and modified during retrieval practice. Based on this framework, Conway (2009) suggested that these patterns of activation/inhibition are inherent to long term episodic memories.

In the following sections I will take the most elaborated theory, the inhibitory control theory, highlight its main hypothesis, and the logic of research in the field that is based on testing
these hypotheses. Then, I will briefly describe the vast amount of findings supporting and those not supporting these hypotheses. Again, I will finish this chapter with a critical evaluation of the mainstream theory based on the evidence that has emerged.

**Testable hypotheses of the inhibitory control theory**

Inhibition in this theoretical approach is a process that operates when a relatively strong competing item interferes during retrieval of a target memory. This approach involves four testable properties of RIF. First, RIF is interference-dependent, i.e. only items interfering with the retrieval of a target memory would suffer inhibition, and more strongly competing items cause more inhibition. Second, RIF is retrieval specific, i.e. manipulating target strength without retrieval of the target does not induce competitor forgetting. Third, RIF is strength-independent, i.e. even when targets are retrieved, target strength does not influence RIF. Together, these first three assumptions imply that RIF is the product of competitive retrieval. The fourth property of RIF is cue-independence, i.e. because it is the memory representation itself that inhibition weakens to resolve interference, the representation should be less accessible with all types of cues tapping that representation (e.g. Anderson, 2003). When these four hypotheses are corroborated than one can refute an account of RIF that does not involve any inhibition.

Although Raaijmakers and Jakab (2012a) have already noted that the properties of interference-dependence, retrieval dependence and strength independence are not independent from each other, and all rest on the "assumption that it is the competition during retrieval practice that determines the amount of inhibition observed" (p. 102), in treaties on RIF, these four properties are considered to be of equal value.

Here, we treat the first three hypotheses and the latter one separately, and try to argue that cue-independence is the crucial hypothesis for the inhibition theory of Anderson. From a methodological point of view the first three hypotheses are weaker hypotheses than the fourth
one. One can provide negative evidence for a certain range of interference theories by finding null results (i.e. showing no RIF) in experiments where targets are not retrieval practiced in a competitive retrieval task. However the cue-independence hypothesis is more difficult to corroborate since one needs positive results. To be more specific, when testing the first three properties, the null hypothesis is that there is an (RIF) effect, whereas in tests of the cue-independence the null is that there is no (RIF) effect. In statistical hypothesis testing, it is way less probable to refute the latter null. Therefore providing corroborating evidence for cue-independence is more challenging than providing corroborating evidence for the first three hypotheses. Approached from another angle, interference- and retrieval dependence are necessary conditions for RIF to occur, while strength-independence is a property of retrieval induced forgetting that contradicts accounts based on interference theories of RIF. However if one leans on hypothesis testing, then using independent-cues is the only proper test of inhibition theory itself.

In Anderson's (2003) approach retrieval is a necessary condition to induce RIF, however it is not sufficient. A retrieval must be a competitive one. Even if a target memory is retrieved, it is only interfering items that will get inhibited. In another publication, Anderson points out that cue-independence is the crucial property out of the magic four (Anderson, 2006). In sum, we take the standpoint made by Camp, Pecher & Schmidt (2007): "To distinguish between interference and inhibition, the final memory test should use independent cues" (p. 951.)

**Logic of research using the retrieval practice paradigm**

The convention in the literature on RIF (Raaijmakers & Jakab, 2012a) is to treat the four hypothesis of the inhibitory control theory separately. Most experimental work tried to factorially manipulate variables in order to test one hypothesis at a time. For instance, manipulating practice conditions by repeatedly studying associations versus repeatedly retrieving them to test retrieval-specificity; manipulating target strength by using high versus
low frequency targets to test strength-independence, manipulating competitor strength by using high versus low frequency competitors to test interference dependence, and manipulating final test procedures using independent probes versus same probes at final tests to test cue-independence (for examples, see the sections below).

Again, as with the TNT paradigm, there are two approaches to investigating the effect, one is more exploratory, trying to find the boundaries of when RIF occurs and when it does not, and the other is more theoretical, focusing on the evidence supporting possible theories of RIF. Similarly to the TNT literature, there is a bias in the field when researchers take the former approach towards taking RIF as a measure of inhibition, and considering every significant RIF finding as evidence for inhibition. This bias, again, is the strongest with neuroimaging findings.

Only one attempt has been made to model mechanisms of RIF with the inclusion of a range of parameters derived from all four hypotheses of the inhibitory control hypothesis (Norman et al., 2007). This neural network model proved to be successful in explaining all available evidence with a model that involved the interaction of the need to resolve interference and a fluctuating baseline inhibitory process.

**Behavioral findings**

In this section, following a brief review of behavioral findings, I will consider evidence that has accumulated in favor of the four critical hypotheses of the inhibitory control hypothesis. Since thorough and accessible reviews on these findings are available (Anderson, 2003; Bäuml et al., 2010; Storm, 2011), besides the summary of the relevant findings in these papers, I will only discuss experimental studies that appeared since the publication of the last review.

Retrieval-induced forgetting has been replicated in tests of category cued free recall (Anderson et al., 1994; Jonker, Seli, & MacLeod, 2012; Moulin et al., 2002; Racsmány &
Conway, 2006; Williams & Zacks, 2001), category-plus word stem cued recall (Anderson et al., 1994; Camp, Pecher, & Schmidt, 2007; Storm, E. L Bjork, & Bjork, 2007), recognition (Hicks & Starns, 2004; Román, Soriano, Gómez-Ariza, & Bajo, 2009; Verde & Perfect, 2011), lexical decision (Racsmány & Conway, 2006), and other implicit tests (Camp, Pecher, & Schmidt, 2005; Perfect, Moulin, Conway, & Perry, 2002), as well as in situations of socially shared remembering (Barber & Mather, 2012; Coman, Manier, & Hirst, 2009; Cuc, Koppel, & Hirst, 2007). Besides neutral (Perfect et al., 2004) and emotional (Barber & Mather, 2012) verbal material, including remembering personality traits (M. D. MacLeod & Macrae, 2001), eye-witness memories (Shaw, Bjork, & Handal, 1995), and narratives (Saunders & MacLeod, 2006), it has been replicated using visual memories (Ciranni & Shimamura, 1999; Waldhauser, Johansson, & Hanslmayr, 2012) as well as autobiographical event memories (Barnier, Hung, & Conway, 2004; Storm & Jobe, 2012), and (!) movements (Tempel & Frings, 2012). These replications suggest that RIF is a robust phenomenon independent of the form and modality of memories.

The application of the retrieval practice paradigm in clinical settings is hard to cover in its entirety. A general guideline in these studies was to see whether the different forms of inhibitory deficits observed in these populations also generalize to inhibitory control processes related to memory. Amongst others it has been shown to be intact in schizophrenia (Racsmány et al., 2008), deficient in posttraumatic stress disorder (PTSD) (Amir, Badour, & Freese, 2009) and depression (Groome & Sterkaj, 2010). To capitalize on the most relevant findings, the lack of RIF in Alzheimer's disease (Moulin et al., 2002) has been interpreted such that the high intrusion rate made by Alzheimer patients is not due to potential deficits of inhibitory functions in retrieval. Another study (Storm & White, 2010) showed that ADHD patients exhibit an impairment in RIF, suggesting that their inhibitory deficits generalize to memory processes. Similarly, one might expect adaptive forgetting to be deficient in
obsessive-compulsive disorder (OCD), where the syndrome involves the inability to flexibly inhibit several cognitive processes (Chamberlain, Blackwell, Fineberg, Robbins, & Sahakian, 2005). Here, a study by Jelinek et al. (Jelinek, Rietschel, Kellner, Muhtz, & Moritz, 2012) found intact RIF among OCD, however, a recent study using a better controlled design, did find evidence for a lack of RIF in OCD (Demeter, Keresztes, Harsányi, Csigó, & Racsmány, 2013).

In their review, Bäuml et al. (2010), summarize studies using emotional material in RIF studies, and conclude that although RIF is comparable for emotional and neutral material, RIF does not occur when retrieval practice is performed in negative moods (Bäuml & Kuhbandner, 2007). They suggest that this is possibly because negative emotions induce distinctive encoding of memories which in turn reduces the interference that memory inhibition should overcome. In line with this reasoning, a recent experiment (Koessler, Steidle, Engler, & Kissler, 2013) showed that RIF was not induced while participants performed the retrieval practice paradigm under stress.

Because the inhibitory control theory assumes that an active executive process is involved in inhibiting interfering memory representations, manipulations targeting the involvement of executive functions are highly relevant. In two experiments, Román et al. (2009) made their participants perform an updating task that put a continuous load on executive attention during retrieval practice, and found that the executive load eliminated the RIF effect they found in a control condition. Aslan and Bäuml (Aslan & Bäuml, 2011) measured individual differences in working memory (WM) capacity using an operation span task (OSPAN), and observed a significant correlation between WM capacity and RIF. In contrast, in our own study (Demeter et al., 2013), we did not find significant correlations between the performance on an n-back task and the amount of RIF. However, this might be a power issue, since the correlation found in the Aslan and Bäuml study was based on data from 168 healthy participants, compared to
the 25 healthy participants and 25 participants with OCD in our sample. Another difference might be the complexity of the OSPAN and the n-back task, and that, as suggested by a recent metaanalysis (Redick & Lindsey, 2013), the two task seem to measure different constructs. Here again, as with the TNT effect, studying elderly people is interesting, because the decline in general executive abilities with age might suggest that RIF might not be as large as among younger adults. Studying RIF in people aged between 60-79 years old, Aslan, Bäuml, and Pastötter (2007) in a well controlled study, found that older participants produced a RIF effect comparable to that of younger participants sampled from a university student population. To our knowledge, no studies have yet investigated RIF among children. Further systematic investigations are needed to establish what exact executive functions are related to RIF, and how they vary across the life span. Such studies would inform theoretical debates on mechanisms underlying RIF in establishing the exact contribution of executive functions to RIF.

An interesting issue still debated is the durability of the RIF effect. The two major theories (the inhibitory control theory, and the interference theory) have no clear assumptions about how long the effect should persist, however one of the clear claims of the episodic inhibition theory (Racsmány & Conway, 2006) is that it should be observable irrespective of the delay period between retrieval practice and the final test, given that the episodes formed during retrieval practice are the ones accessed during final test. Support for this hypothesis has also been mixed (Abel & Bäuml, 2013; Baran, Wilson, & Spencer, 2010; M. D. MacLeod & Macrae, 2001; Racsmány, Conway, & Demeter, 2010). It seems that the RIF effect may persist at delays longer than a 12 hour period in cases when sleep induces consolidation of patterns of activation/inhibition (Racsmány et al., 2010). In a fully contrasting account, explaining the negative results of Baran et al. (2010), Macleod and Hulbert (2011) argue that sleep, REM sleep in particular, may dampen the inhibition of previously interfering
memories, and allow for their recovery. Clearly, this issue needs further experiments to answer the question of stability of RIF.

**Neuroimaging and electrophysiological findings**

Bäuml, Pastötter, and Hanslmayr (2010) gives a short review of the three EEG studies and the three fMRI studies that had appeared until 2010. A coherent picture emerging from these studies is that retrieval practice leads to increased frontal activity compared to conditions when participants simply relearn cue-target associations, and importantly, the amount of this change in positivity is correlated with RIF (Johansson, Aslan, Bauml, Gabel, & Mecklinger, 2007; Wimber et al., 2008). This difference in memory (DM) effect has been shown to be coupled with a decrease during each additional retrieval attempt in activations in prefrontal areas as well as in the anterior cingulate cortex (Kuhl, Dudukovic, Kahn, & Wagner, 2007). The fact that this decrease was correlated with the RIF effect was supposed to be the result of successful inhibition of the interfering memory representations during retrieval practice.

During the final test then, these interfering memories, even if retrieved, have been shown to increase ventrolateral prefrontal (VLPFC) activations in one study (Wimber, Rutschmann, Greenlee, & Bäuml, 2009; but not in another: Kuhl, Kahn, Dudukovic, & Wagner, 2008), and decrease evoked potential in the theta and gamma frequency (B. Spitzer, Hanslmayr, Opitz, Mecklinger, & Bäuml, 2009). The authors of the review suggested that these test-based effects show a decreased activation of inhibited memory traces.

Four new studies have been published since the publication of the Bäuml et al. review. Two of these were replications of earlier findings, showing increased theta activity for competitive retrieval compared to non competitive retrieval conditions, and observing correlations between RIF and increase in theta during competitive retrieval (Simon Hanslmayr, Staudigl, Aslan, & Bäuml, 2010; Staudigl, Hanslmayr, & Bäuml, 2010). In the third study, Waldhauser and colleagues (Waldhauser, Johansson, et al., 2012) used a clever design to differentiate the
effects of non-interfering and interfering competitors during retrieval practice of target memories. In each trial they flashed a cue in the middle of the screen and flashed as target a colored box (for 200 ms in order to avoid saccades to targets) to the left or to the right of the cue. Using this procedure they had participants associate two targets with each cue. Importantly, the two targets always differed in position (left or right) but for half of the items they also differed in color (interfering condition) whereas for the other half they did not (non-interfering condition). During retrieval practice, they had participants recall the color for one of the targets for half of the cues. In each trial during the retrieval practice phase, a cue was presented together with a white box to the left or to the right of the cue, indicating to the participant the exact box the color of which they should recall. This design also allowed the authors to contrast ipsi- and contralateral activity in ERPs during events when participants recalled interfering versus non interfering memories. Crucially, their hypothesis, based on earlier studies of oscillatory signatures of inhibition, was that interfering memories would elicit higher alpha/beta power over the hemisphere where the competitor is processed if these memories should be inhibited during retrieval practice. This is exactly what they found, with a nice addition that this increase was also correlated with the amount of RIF they found behaviorally. Similarly to this and other neuroimaging studies reviewed above, using a manipulation of competitor strength, a recent study (Hellerstedt & Johansson, 2013) also found that increased changes in ERP induced by competition during retrieval practice were associated with higher magnitudes of RIF. Again, the results were interpreted as supporting the inhibitory control theory.

Evidence for and against the inhibitory control hypothesis

Retrieval specificity, strength independence and interference dependence

As noted earlier, these properties are not completely independent from one another. Basically they all refer to the conditions when interference should be resolved, and inhibition should
play its role of decreasing the activations of interfering memory traces. Evidence in favor of and against these properties is reviewed elsewhere, with some reviews being biased towards an interpretation favoring the inhibitory control theory (Anderson & Levy, 2011; Anderson & Spellman, 1995; Anderson, 2003, 2005; Storm, 2011), whereas others being biased in favor of the interference account of RIF (Raaijmakers & Jakab, 2012a).

What should be noted is that, in all cases, experiments have used a factorial manipulation of one variable that informed the authors about the relevant hypothesis. For instance, in the case of interference dependence, Anderson and Levy (2011) reviewed experiments that had varied the interference induced by a competitor during retrieval by (1) using high vs. low frequency competitors, (2) using the dominant vs. the subordinate meaning of homonyms, (3) in picture naming tasks, instructing participants to name pictures in a second language vs. their mother tongue, (4) varying the competitors' associations to cues, or varying demands on interference resolution by varying the type of practice used, such as (5) using the standard retrieval practice conditions vs. conditions where participants are simply reexposed to the cue-target pairs, or (6) asking participants to perform a category plus word stem cued recall vs. having them retrieve the cue when presented with the target, or as an extreme (7) having participants perform impossible vs. possible retrievals.

The expectation is that output interference accounts cannot explain results when RIF is not found in the low interference condition of these manipulations. Also, the pattern that RIF is found only in the high interference conditions implies that inhibition only occurs when interference resolution becomes necessary during retrieval.

Only a few behavioral studies attempting to test theories of RIF appeared since the publication of the last two of these reviews. Two (Raaijmakers & Jakab, 2012b; Verde, 2013) have provided evidence against the inhibitory control account by showing that practicing cue-target associations without cued retrieval of targets can induce RIF, i.e. they provided
evidence against retrieval specificity. A further study by Raaijmakers and Jakab (2013) has provided evidence against the strength independence account of RIF. In addition, the behavioral results of two neuroimaging studies described in more detail above have provided evidence for the interference dependence in RIF, by showing that only interfering memories were subject to RIF (Waldhauser, Johansson, et al., 2012), and items with higher interference caused larger RIF than items with lower interference (Hellerstedt & Johansson, 2013).

Cue independence

As discussed above, both for the TNT and the RIF paradigm, cue independent tests of forgetting are the critical tests for the inhibitory control account, as each significant cue-independent forgetting is a falsifier of any theory that does not posit decreases in activation of memory representations at the item level.

The independent probe technique of Anderson and Spellman (1995) was used in several experiments to provide support for the idea of cue-independent forgetting (Anderson & Spellman, 1995; e.g. M. D. MacLeod & Saunders, 2005; Saunders & MacLeod, 2006). In this technique, half of the members of a category are associated to another category, which serves at test as an independent cue. However, this other category might be already associated to the target items during the study phase. Because of this association Camp and colleagues (Camp et al., 2005, 2007) and others (Perfect et al., 2004) argued that the independent probe-technique is not fully independent. Camp et al. (2007) suggested that only item-specific cues that do not appear in the experiment until the test phase could be used as independent cues that validly test the inhibitory hypothesis.

One could argue that if forgetting is cue-independent than showing that RIF generalises to all types of tests would be evidence for cue-independence. Following this line of reasoning, several studies have shown that RIF occurs on item recognition tests (Dehli & Brennen, 2009; Gómez-Ariza, Lechuga, Pelegrina, & Bajo, 2005; Hicks & Starns, 2004; Koutstaal, Schacter,
Johnson, & Galluccio, 1999; Veling & Knippenberg, 2004) and implicit memory tests where the category is not present at test (Bajo, Gómez-Ariza, Fernandez, & Marful, 2006, Exp 2. Veling & Knippenberg, 2004, Exp 3. ). However, Veling and van Knippenberg (2004) argued that recognition tests cannot be independent of the original cue of a target memory because participants can rely on studied categories when making their old-new decisions. Similarly, Camp et al. (2005) and Bajo et al. (2006) argue that implicit tests measuring RIF could be plagued by explicit contamination. Camp et al. (2005) tested this suggestion and showed that RIF occurred in an implicit test only if participants were aware of the relationship between study, practice, and test items. These results cast doubt on whether implicit or recognition tests that found RIF could be considered as instances of cue-independent forgetting.

Nevertheless, the independence of item-specific independent cues is hard to question by the aforementioned arguments (although see: Camp, Pecher, Schmidt, & Zeelenberg, 2009).

Reviewing the literature for corroborative evidence for the cue-independent nature of RIF we found five experiments showing RIF with item-specific independent cues in RIF (Aslan et al., 2007, experiment 2; Johnson & Anderson, 2004, experiment 2; Levy, McVeigh, Marful, & Anderson, 2007, experiment 1, 2; Saunders & MacLeod, 2006 experiment 1) and seven that did not (Camp et al., 2007, experiments 2, 3, and 4A; Johnson & Anderson, 2004, experiment 1; Perfect et al., 2004, experiment 1-3).

Looking at these mixed results, one can either infer that there is no genuine cue-independent forgetting or say that the methods used to detect it are not sensitive enough. Indeed, in an attempt to defend the assumption of cue-independence, based on the results of two experiments, Huddleston and Anderson (2012) argued that some studies might not find cue-independent forgetting because they do not control for or inefficiently control for the semantic relatedness of independent cues and the original cues used during the study phase of these experiments.
A coherent picture involving inhibition and the challenge of interference accounts

On the basis of a vast amount of data, also reviewed above, a coherent picture on the mechanisms underlying RIF emerges, also echoed by recent reviews on RIF (Anderson, 2003; Bäuml et al., 2010; Storm, 2011): Whenever interference occurs during memory retrieval, and is detected (basically supported by the anterior cingulate cortex), this kicks in an interference resolving process which inhibits the interfering non target representation via control mechanisms mediated by prefrontal areas (see for instance Bäuml et al., 2010). It seems that many clinical populations that are characterized by reduced inhibitory control abilities also manifest deficits in this higher order memory inhibition process, failing to produce RIF. However, and interestingly, older people still profit from this adaptive forgetting mechanism. A few theorists have drawn a somewhat different picture (e.g., Camp et al., 2007; Raaijmakers & Jakab, 2012a). As with the TNT paradigm, an inhibitory theory of RIF hinges on RIF found with cue-independent test. These, as seen above, are scarce, and need further support. An especially intriguing problem is whether independent cues can be indeed independent. As seen above, Huddleston and Anderson (2012) and Anderson (2003) suggested that RIF might be attenuated by covert cuing in tests of RIF using independent cues, and urged for a better control of independent cues. However, when Camp et al. (2009) demonstrated that independent cues might not be independent from cues used in the study phase, they argued convincingly that covert cuing might lead to RIF that is explainable by output interference effects, and not that it might attenuate RIF. In brief, it is hard to see, how true independence of cues (if it can be achieved at all), should influence the RIF effect. Indeed, when interpreting the neuroimaging results of RIF, none of which used independent cues, Bäuml et al. (2010, p. 1050) ran into an interesting contradiction. Based on earlier results showing that VLPFC is active during retrieval of 'weakly bound memory traces', they interpreted increased activation in VLPFC during the retrieval of competitors at test 'as
evidence for the item suppression account of RIF. Clearly, when binding of target memories is weakened, it is by definition a cue-target association that is weakened, not an item level representation.

Note that the very concept of a true independent cue implies that some cue representation might be associated to a target that is associated to another cue, and yet the two cues might not share associations. The logic inherent in this suggestion is similar to that of positing a mechanism that affects target memory representations independently of any other associations. This seems to be a dead end.

Importantly, as clinical studies using the TNT paradigm, none of the clinical studies used independent cues to test RIF, therefore finding RIF in one population and no RIF in another does not mean impaired or intact memory inhibition, respectively. Most clinical studies have not controlled for output interference either (for exceptions, see Demeter et al., submitted; Soriano, Jiménez, Román, & Bajo, 2009; Storm & White, 2010), i.e. used category cued free recall tests. In these tests, only a category cue is presented, without specifying the exact target to be retrieved. Therefore, practiced targets usually come to mind first during the test, and these block access to non practiced targets. The RIF observed in these tests can then be explained with interference that occurred during the final test; no inhibition is required at all.

**A potential cautious interpretation**

Although evidence for cue-independent forgetting is weak, and the evidence for strength independence, retrieval specificity, interference dependence is somewhat mixed, we suggest that the studies reviewed above provide converging evidence that cognitive control mechanisms do act to resolve interference during competitive retrieval, and that this interference resolution has some detrimental effect on later accessibility of interfering memories. We would have to discard all the neuroimaging results to suggest that forgetting on a final test in a retrieval practice paradigm is solely due to interference occurring during the
final test. However, the interference resolving process during competitive retrieval that leads to at least part of the RIF effect, might not necessarily involve active inhibition of memory items. Rather, as will Study 2 and 3 further exemplify, it might involve other cognitive control mechanisms, either facilitatory or inhibitory, that contribute to the redistribution of associative strengths and activation resources.

**Understanding the mechanisms of interference resolution**

Of specific interest to us here, is the inconclusiveness of results in favor of the inhibitory control theory or the interference theory, and the need for clarifying the role of latent variables not targeted by experimental manipulations thus far. Study 2 proceeded in elucidating the mixed results concerning the role of interference in RIF.

A key assumption of inhibitory theories is that when a cue-target association is repeatedly encountered, this encounter will cause inhibition if and only if the encounter involves retrieval and interference induced by competitors related to the same cue. This assumption of the inhibitory control theory is termed interference dependence. It entails, that only retrieval competition will initiate cognitive control mechanisms that will inhibit competitor representations. Experiments testing this assumption have provided both positive and negative results, therefore the interference dependence of RIF is still debated. These mixed results might be due to either a week effect, or the fact that RIF is not a monotone function of competitor interference. Study 2 investigated the latter possibility by using a continuous post-hoc measure of interference to plot the relationship between interference and forgetting.

Rather than factorially manipulating interference as seen in studies reviewed by Anderson and Levy (2011), and summed up briefly above, this study allowed for assessing interference dependence in a direct manner. Basic findings suggest that RIF is indeed interference dependent, but the function relating competitor forgetting and competitor interference is non-
monotonic. This fact could have led to the fact that some studies failed to find interference dependence.

Two further assumptions of the inhibitory account are strength independence (i.e., cue-target association strength does not matter for the amount of RIF) and retrieval specificity (i.e. retrieval is necessary for RIF to occur). The most widely used way to test these assumptions has been to manipulate the type of reencounter with cue-target associations (see Anderson, 2003 and Anderson & Levy, 2011, for reviews). The study event is followed by repeated encounters that in one condition involve the presentation of the full cue-target pair (repeated study), whereas in another condition, they involve cued recall of the targets (repeated retrieval). The basic finding here is that both conditions lead to the same amount of strengthening of target memories, but only repeated retrieval leads to RIF. It is concluded that first, repeated retrieval is necessary for RIF to occur, and second, although the strength of target memories after reencounter is the same in the two condition this does not affect competitor recall to the same extent. According to interference accounts, the same level of cue-target association strength should lead to the same amount of cue-competitor forgetting. However, experiments in an other line of research, discussed in the next chapter, have clearly shown that repeated study and repeated retrieval lead to different amount of memory strengthening. Nevertheless, this difference only becomes evident after a delay at least a day long, and becomes fairly robust at delays of at least a week.

In case this difference in strengthening occurs in the retrieval practice paradigm too, it would mean that previous tests of the retrieval specificity and strength independence assumption were based on a false auxiliary hypothesis.

Study 3 looked at this question by using a standard manipulation of repeated encounters (i.e., repeated study versus repeated test), and testing all targets and competitors at a longer delay interval as well. This study showed that, as reflected by tests after a long retention interval,
repeated study and repeated test indeed leads to different levels of memory strength even in a retrieval practice paradigm. These results show that although RIF is retrieval specific, it is not necessarily strength independent.
The positive effect of retrieval practice - The testing effect

All theories explaining the TNT and the RIF effect make predictions about how cue-target and cue-competitor associations change. However, the key concern of research with the retrieval practice paradigm has been to show how cue-competitor relationships might be changed due to stopping retrieval or target recall to a given cue. Similarly, the key concern of research with TNT paradigm has been a negative effect: the forgetting of cue-target associations caused by stopping retrieval. Another currently burgeoning research line has rather focused on how cue-target associations can be changed by different types of reencounters. The key concern of this line has been a positive effect of retrieval: why repeated retrieval of cue-target associations leads to better long-term retention than repeated study? These two research line have started to converge only very recently. This is interesting, given that both lines can feed massive empirical data for one another. For instance, RIF studies testing interference dependence often used manipulation of how associations are reencountered (e.g., repeated study versus repeated retrieval) to examine the behavioral and brain mechanisms correlated with RIF. These studies measured ERPs (e.g. Johansson et al., 2007; Wimber et al., 2008) or BOLD signals (e.g., Kuhl et al., 2007) during repeated study and/or repeated retrieval and correlated the changes occurring during repetitions with later RIF. Data pertaining to the correlations between these signals and the positive effect of cue-target retrieval (not reported in any of these studies) could inform theories on how the positive changes to cue-target associations occur. This chapter describes these positive changes, and briefly introduces current theoretical frameworks explaining these positive effects. It also introduces Study 4 that contributed to the investigation of the positive effects of cue-target retrieval.
The testing effect

What is the capital of Hungary? What brain areas are involved in semantic retrieval? These questions might be fun when watching a quiz show on TV or preparing slides for a memory course. But for students who face these questions at exams, tests are mere assessments of what they know, something bad to get over with, a feeling probably most of us have about tests. Conceiving tests as measurements of knowledge and the stress tests induce might have lead to the rareness of testing in educational settings.

However, it has been shown that testing is a strong memory enhancer, and that testing is more beneficial for long-term retention than restudying (Roediger & Karpicke, 2006a; Tulving, 1967). Reviewing the history of the testing effect, Roediger and Karpicke (2006) suggested that research on testing had been deemed by the rising interest in forgetting in the 60’s. It had been known that testing can stop forgetting, therefore reexposures and tests of material before a final test was considered to be just a confounding factor in forgetting experiments.

A renewed interest has lead to many behavioral replications of the testing effect, i.e. that short term recall is similar for restudied (repeatedly presented for study) and retested (repeatedly tested) memories, but long-term recall is better for retested memories than restudied ones. The benefit of repeatedly retrieving memories over repeatedly restudying them is evident even without giving feedback (Roediger & Karpicke, 2006b), an effect called direct by Roediger and Karpicke (2006a), as opposed to other effects that are mediated by the whole process of testing, including judgment of learning or feedback for instance in decisions to allocate resources to less learned memories during subsequent tests (although, as discussed later, feedback does enhance the effect). The testing effect has been shown to produce larger mnemonic benefits than studying with effective elaborative learning mechanism, such as concept mapping (Karpicke & Blunt, 2011a; however see the ensuing debate Mintzes et al., 2011; Karpicke & Blunt, 2011b)
Judgments of learning, i.e. learners' estimation of the effectiveness of their study strategy is usually found to be uncorrelated with the effect of testing (Karpicke & Roediger, 2008), but under certain condition, these estimates can be improved (Tullis, Finley, & Benjamin, 2012). The critical interaction of recall performance on immediate versus delayed final tests of restudied versus retested memories remains significant even when time of presentation and potential confounds in differential elaboration of restudied and retested material are controlled for (Toppino & Cohen, 2009).

Testing can have additional positive effects, such as enhancing the effectiveness of study (Grimaldi & Karpicke, 2012; Halamish, Goldsmith, & Jacoby, 2011; Szpunar, McDermott, & Roediger III, 2008; Weinstein, McDermott, & Szpunar, 2011; Wissman, Rawson, & Pyc, 2011), leading to the formulation of the concept of test-potentiated learning. For instance, free recall test sessions performed between study sessions reduced proactive interference that was building up with accumulating study sessions compared to a condition without interpolated test sessions (Szpunar et al., 2008). It has been suggested that tests might insulate against the buildup of proactive interference, as shown in word list learning (Szpunar et al., 2008) and face-name pair learning (Weinstein et al., 2011). However interesting, the following sections will not consider test-potentiated learning, and will rather focus on results that further the understanding of the testing effect proper.

**Behavioral findings**

The testing effect, has been replicated using different kinds of materials, such as prose passages (A. C. Butler & Roediger, 2008; Karpicke & Roediger, 2010; Wissman et al., 2011), face-name pairs (Weinstein et al., 2011), word lists (Szpunar et al., 2008), word pairs (Karpicke & Zaromb, 2010), and visuospatial information (Carpenter & Pashler, 2007). Most extensively, due to the availability of a material set standardized for recall probabilities, and generally excluding any preexperimental semantic or episodic memory for the target
memories in the set (Nelson & Dunlosky, 1994), Swahili-English/Dutch/etc. word pairs have been used (Eriksson, Kalpouzos, & Nyberg, 2011; Karpicke & Roediger, 2008; Pyc & Rawson, 2010). The effect of testing has been demonstrated after delays of 9 months (Carpenter, Pashler, & Cepeda, 2009), one week (A. C. Butler & Roediger, 2008), 48 hours (Toppino & Cohen, 2009), five minutes (Karpicke & Zaromb, 2010), using several types of final tests, like multiple choice testing (A. C. Butler & Roediger, 2008), free recall (Szpunar et al., 2008), recognition (Karpicke & Zaromb, 2010), and cued recall (Karpicke & Zaromb, 2010).

A focused review of the testing effect literature showed that the positive effects of retesting versus restudying memories is transferred from one test format to another, and from one knowledge domain to another (Carpenter, 2012). The latter means, for instance (A. C. Butler, 2010) performing better on inference tests that required the use of previously retested information, as compared to when the information had been previously restudied. Two recent findings have corroborated the transfer effects summarized by Carpenter. One study (Putnam & Roediger, 2012) varied mode of retesting (typing, speaking, or thinking about responses) and mode of final test (speaking or typing) and found that testing effects emerged even when the mode of retesting and final test were different. Another study (Verkoeijen, Bouwmeester, & Camp, 2012) showed that the positive mnemonic effect of testing appears on short term tests as well under conditions that require transfer of the acquired memories. In their study, English-Dutch bilingual individuals learnt Swahili-Dutch word pairs, and on a final test administered shortly after the learning phase, using retesting versus restudying, they only produced a testing effect when the Swahili words were cued with their English meaning, not when the original Dutch cues were used.

No clinical studies of the effect have yet been published, and developmental studies are also scarce. Besides an early study (H. F. Spitzer, 1939) now hard to replicate, the two testing
effect experiments, ran with children aged between 7-13 (Bouwmeester & Verkoeijen, 2011), and teenagers (Carpenter et al., 2009) showed that testing is more beneficial than restudying also among children.

**Moderating factors**

Feedback has been shown to enhance the testing effect (A. C. Butler & Roediger, 2008; Carpenter & Vul, 2011; Roediger & Butler, 2011; Thomas & McDaniel, 2013). For instance, giving immediate feedback after taking each multiple choice tests enhanced performance on a cued recall test administered one week after taking the initial multiple choice test, whereas giving delayed feedback (at the end of each test session) improved later recall even more (A. C. Butler & Roediger, 2008).

Arousal also seems to enhance the testing effect. In two studies (Finn, Roediger, & Rosenzweig, 2012; Finn & Roediger III, 2011) participants were presented with negative emotional or neutral pictures from the International Affective Picture System (IAPS) after successful retrieval of some associations. Arousing but not neutral pictures presented three seconds after successful retest trials enhanced the effect of testing significantly.

Also important for the upcoming discussion of theoretical frameworks offered is the effect of the strength of cue-target associations on the testing effect. Varying factorially this association, using weakly versus strongly associated pairs, Carpenter (2009) showed that although initial retrieval of weakly related pairs was slower and less successful than that of strongly related pairs, a test administered only five minutes later showed a reverse effect, with more weakly related pairs being recalled than strongly related ones.

**Theoretical explanations - yet without a coherent theory**

Several accounts have been proposed to account for the testing effect, although none of these can be regarded as a theory proposing to explain all aspects of the behavioral findings. In this
section, I will briefly go through these accounts.

At first glance, the testing effect might simply be due to transfer appropriate processing (Morris, Bransford, & Franks, 1977), assuming that task demands and processing of memories during retest occasions matches final test demands more closely than restudying. This account fails in instances of the testing effect where transfer occurs (see above). The concept of elaborative retrieval has provided another explanation (for a short description see 2009). Elaborative retrieval processing is assumed to happen when a presented cue-target association is encoded with the use of information missing from the actually presented association. For instance, when faced with the task of memorizing the 'grass-drink' pair, you might want to add information like, park, friends, or football, sweat, etc., information that helps you bind the cue with the target. This elaboration is supposed to be boosted through spreading activation (Collins & Loftus, 1975) when you are asked to recall 'drink' when 'grass' is presented. Also, as the target is available during restudy occasions, it seems unlikely that such elaboration is occurring. This account, although supported by some data, does not specify why elaboration should occur during retrieval, more than during study. After all, it starts with the assumption that elaborative encoding occurs during initial study. Therefore based on the premises of spreading activation, one could argue in the opposite direction, and say that elaboration should be greater when the cue and the target are repeatedly presented.

A somewhat similar explanation is offered by Bouwmeester and Verkoeijen (2011) who suggest that the fuzzy trace theory (Reyna & Brainerd, 1995) can also explain the testing effect. The relevant feature of the fuzzy trace theory given by the authors here is that when lists of words are learnt, alongside the verbatim, item level memory traces, gist traces are formed that incorporate the common features of the members of the list. These gist traces are thought to be strengthened through spreading activation during retests, whereas verbatim traces are thought to be strengthened through repeated study. Since restudy does not lead to
gist trace formation, the recall of restudied words produces much less false alarms at a final test administered after a short delay. Also, since verbatim traces are strengthened for restudied materials, these memories will be highly accessible on the short run. However, and this is an additional prediction of the fuzzy trace theory, verbatim traces decay much faster then gist traces, and this then leads to the interaction of learning strategy and retention interval. The problem with this theoretical suggestion is that the results it was based on come from a study with DRM lists which are composed of highly associated words. It is easy to see how prose passages (that are also composed of highly integrated traces) might provide another good instance of the testing effect that the fuzzy trace theory can explain. It is unclear however, how this viewpoint should explain the testing effect found with lists of unrelated words, and testing effects on final tests using cued recall, or recognition.

Predictions of both the above theoretical frameworks can be generated by a more parsimonious one, the mediator effectiveness hypothesis (Pyc & Rawson, 2010). The hypothesis is simply that people generate mediators during encoding and testing. Pyc and Rawson define a mediator as a 'word, phrase, or concept that links a cue to a target', and suggest that mediators are becoming more effective with repeated tests than with repeated study. They offer two explanations for why this might occur. First, recall of mediators during testing might increase the memory strength of mediators, and second, mediators might be changed to more effective ones with subsequent testings. Of course, this hypothesis is very similar to the previous two, and one might even question what differences there are between the elaborative retrieval processing and this one. Essentially, a mediator is one memory used to elaborately encode and retrieve associations. What is definitely not involved in the above two hypotheses is that mediators might become effective during testing thanks to the possible shifts from less to more effective mediators, a suggestion also supported by data from the Pyc and Rawson study (2010).
Another quite parsimonious model, the distribution-based bifurcation model (Halamish & Bjork, 2011; Kornell, Bjork, & Garcia, 2011) suggests that the interaction of retention interval and learning strategy is due to a bifurcated distribution of memory strength of retested memories. Basically it states that the memory strength of initially studied memories is normally distributed. When memories are retested, memory strength of successfully recalled memories is shifted to the right of distribution, i.e. their memory strength is increased, whereas the memory strength of retested memories that are not recalled during the test phase remain unchanged. In contrast, memory strength of restudied memories shifts to the right but to a lesser extent than that of successfully retested memories. Again, although sounding very mathematical, and therefore very theorish like, the distribution-based bifurcation model is in fact descriptive (a fact also acknowledged by the authors), and offers no explanation of why retested memories (or here only successfully retrieved retested memories) are strengthened more than restudied ones.

In order to investigate the role of retrieval mode in the testing effect, one study compared a generation condition (incidental retrieval) with a retrieval condition (intentional retrieval) and found that retrieval enhanced both recall and recognition to a greater extent than generation (Karpicke & Zaromb, 2010). According to the authors it is the retrieval mode that distinguishes the two conditions from one another. Although one can argue that generation is rather a form of semantic retrieval, whereas retrieval in this experiment was an episodic memory retrieval, this still fits with the idea that the difference between these two tasks is the involvement or not of retrieval mode processing (Düzel et al., 1999). One important problem arises though when trying to draw conclusions from the Karpicke and Zaromb paper. Usually the learning strategy in testing effect experiments interact with the retention interval so that restudy produces similar or sometimes larger benefits than retesting after a short delay, and the enhanced retention arising due to retests becomes apparent only after longer delay.
Because the authors here used a delay of five minutes before final test in all four experiments reported, their results do not inform us about the effectiveness of the retrieval mode in producing the testing effect. Follow-up work is needed to determine the exact contribution of retrieval mode to the effect.

Extending this theoretical point of view, later on Karpicke (Karpicke & Smith, 2012; Karpicke & Zaromb, 2010; Karpicke, 2012) suggested that retrieval enhances later recall by constraining the search set activated by cues during testing. This theory in a sense incorporates the elaborative retrieval hypothesis and findings related to retrieval mode involved during retesting memories, and, as will be seen, is highly consistent with the results of Study 4.

**Neuroimaging results contributing to a future theory of the testing effect**

Up to the time of the submission of Study 4 (2012 September), only one neuroimaging result appeared in the literature (see the discussion part of the dissertation for two further studies that are in press now¹). This study, run by Eriksson and colleagues (2011), involved teaching participant word pairs in study-test cycles to 80% criterion, then scanning them while a final cued recall test was administered. During this scanning phase ACC activity was positively, whereas VLPFC activity was negatively correlated with number of successful retest trials during the study-test cycle. They interpreted the negative correlation with VLPFC as an effect of decreased demands on cognitive control mechanisms as during subsequent retest trials retrieval became easier. Without further inspection, the positive correlation with ACC seemed puzzling. Crucially, the study did not end after the final test, they then measured participants' memory for the target words 5 months later as well. And importantly, the increase in ACC activity during scanning correlated with memory performance after 5 months, whereas the decrease in PFC activity did not. Based on rodent studies, showing similar activation patterns

¹ as of 24, July, 2013
during long-term system consolidation of memories Eriksson et al. went on to suggest that the ACC involvement in long-term consolidation of memories underlies the testing effect. Importantly, their study did not involve a restudy condition, and also, their results were not informative of the processes going on in most of the testing effect studies using shorter delays (i.e. max. a week). For instance, the time-scale of changes induced during systems consolidation of memories might be different for PFC and ACC. As they also had a test one week after the scanning, this might have elucidated the role of PFC in consolidation, however as their behavioral results were still at ceiling after a week, they could not calculate correlations of memory performance and brain activity.

Interestingly fMRI and EEG investigations of RIF that contrasted study vs. retrieval practice might be highly informative here (see studies reviewed in the section on the neuroimaging findings of RIF), however these studies did not report any results concerning correlations of brain activity during repeated retrievals and final test performance of retrieved targets and/or contrasts of difference in memory effects between repeatedly retrieved versus repeatedly restudied memories. It might be highly interesting to reanalyze their data from this perspective.

Study 4 in the dissertation was an explorative study that aimed to extend neuroimaging findings using a design that closely matched that of behavioral studies of the testing effect to inform theories of the testing effect.
Main objectives and thesis points

The goal of this dissertation was to understand mechanisms modifying cued recall processes in human episodic memory. These changes can lead to a less effective retrieval process and forgetting, or a more effective retrieval process and better memory. The section below describes the major objectives of the thesis, and sums up the results in four thesis points. These thesis points are based on the four studies that are presented in the next chapter.

Main objectives

As can be seen from the introduction, key points of disagreement between the two major lines of theories (inhibitory and interference accounts) explaining the changes pertaining to retrieval of cue-target associations, concerns the locus of the effect, the cue-dependence of the changes occurring to the accessibility of memories, the retrieval dynamics, and the role of cognitive control. First, inhibitory theories posit that item level representations get inhibited, whereas interference based accounts assume that associations, or cues and contexts are the loci of effect. Second, items' accessibility can be varied independently of cues in the inhibitory models, but not according to interference theories. Third, the inhibitory control theory suggests that memories static representational state affects their accessibility, whereas in interference theories accessibility is determined by the retrieval process guided by strength of associations. The fourth major point of disagreement is whether cognitive control has any role in producing negative and/or positive changes in the retrieval of cue-target associations. Study 1-3 tried to shed light on this debate. In the following paragraphs, I will describe how each study contributed to answering these issues.

The inhibitory control theory of memory inhibition (Anderson, 2003, 2005; Levy & Anderson, 2008) posits that target memory representations can be weakened by intentionally stopping target retrieval in the presence of a cue that was previously associated episodically to
that target (Anderson & Green, 2001). This implies that memory inhibition affects target memory representations, rather than cue-target associations. Our first objective was to test this prediction by testing both cued recall of targets and recall of cues cued by the targets using the think/no-think (TNT) manipulation.

The same theory also posits that the representation of non target memories can also be weakened when these memories interfere with retrieval during cued recall of a target memory, and this interference needs to be resolved by inhibitory executive control, leading to retrieval-induced forgetting (RIF) (Anderson et al., 1994; Anderson & Spellman, 1995). Since competitive retrieval is a necessary condition for this type of memory inhibition to occur (Anderson, 2003), but the results of experiments factorially manipulating competition during retrieval are mixed (Anderson & Levy, 2011; Raaijmakers & Jakab, 2012b; Verde, 2013), our objective in Study 2 was to understand the exact function relating interference and RIF. This knowledge might be crucial for future investigations testing further predictions of the inhibitory control theory of RIF (such as cue independence). If this function is known, future studies could focus on instances of retrieval where competitors do cause the sufficient magnitude of interference to produce RIF.

A further assumption of the above theory is that it is not strengthening of target memories but competitive retrieval that leads to RIF (i.e., the assumptions of strength independence and retrieval specificity). A general logic to test this prediction has been to strengthen memories to the same degree by either repeatedly exposing participants to target memories or have participants repeatedly retrieve them, and see whether it is only the competitive strengthening (retrieval) that causes RIF. However, results of these tests have also been mixed (Anderson, 2003; Jakab & Raaijmakers, 2009; Raaijmakers & Jakab, 2012a). It is important to note that the entire testing effect research is about differences between memory strengthening induced by repeated retrieval versus repeated study events (Roediger & Butler, 2011; Roediger &
Karpicke, 2006a). Hence, the objective of Study 3 was to test the auxiliary hypothesis inherent in tests of strength independence and retrieval specificity, namely that restudying material in the retrieval practice paradigm leads to the same amount of memory strengthening as retesting material.

The fourth objective of the dissertation was to look at the neural correlates of the testing effect. To our knowledge, this study was the first to assess the differential neural correlates of retesting versus restudying where all scanning parameters, and instructions were identical during testing of memories previously retested and testing of memories previously restudied.

**Thesis points**

**Thesis I. Suppressing target memory retrieval leads to changes in cue-target associations, but not in target memory representations. (Study 1)**

The finding that individuals can stop cued recall, and that this leads to forgetting of cue-target associations is well established (Anderson & Huddleston, 2012). It has been unclear whether such a forgetting is due to active inhibition of target memories or substitution of cue-target associations with novel cue-substitute associations. In the former case, forgetting of unwanted target memories would occur due to inhibition through cognitive control (Anderson & Green, 2001), whereas in the latter the cause would be interference caused by substitutes. To test these assumptions in one experiment, Study 1 used a modification of the standard TNT procedure. When presented with cues to unwanted memories, participants were either instructed to suppress or to find substitutes to the unwanted targets. This manipulation allowed us to test whether the effect of direct suppression is indeed the same as that of substitution. A second novel manipulation was that for half of the participants, the final test consisted of a standard cued recall test, whereas for the other half, the test involved a reverse cued recall test. In this reverse test, targets were given as cues, and participants had to respond with the cue word.
Importantly, both suppression and substitution lead to forgetting of targets when participants were cued in the standard way. Crucially, when cued with targets, suppression lead to facilitation of cues, contrary to substitution which did not have such a facilitatory effect. This Study is informative in two ways: First, it shows that both interference and inhibitory accounts can explain the forgetting in TNT. Second, it clearly shows that the forgetting effect is due to changes in cue-target associations, and not changes in memory items.

**Thesis II. Retrieval-induced forgetting is a non-monotonic function of competitor interference during retrieval. (Study 2)**

Retrieval of cue-target association induces forgetting of related cue-competitor associations (Anderson et al., 1994). Interference theories assume that this is due to the fact that the cue-target association is strengthened or biased (see Anderson, 2003). This change will lead to a decreased probability of competitor recall at later time points. Note that this account does not involve any cognitive control mechanisms. In contrast, inhibitory theories suggest that when a competitor interferes with recall of a target related to the same cue, interference is resolved through cognitive control, inhibiting the competitor. In this account, forgetting occurs due to active inhibition through cognitive control. A critical test to decide between the two accounts would be to show that retrieval-induced forgetting (RIF) is interference dependent.

Indeed, several experiments have used factorial manipulations of competitor interference (e.g., by manipulating taxonomic frequency of competitors) to test this assumption (Anderson & Levy, 2011). In case interference dependence were true, these experiments would predict that highly frequent competitors (causing large interference) lead to RIF but low frequency competitors (causing low levels of interference) do not. The results have been inconclusive (Anderson et al., 1994; Williams & Zacks, 2001; Anderson & Levy, 2011). This led to the suggestion that some null results may occur despite the fact that RIF is interference dependent *if* the relationship between interference and RIF is not a linear one. Study 2 aimed at
understanding the exact relationship between interference and RIF, and found that RIF is interference dependent. More importantly, it found that RIF is a non-monotonic function of interference caused by competitors.

**Thesis III. Retrieval-induced forgetting depends on competitive retrieval, but manipulations of learning strategy cannot provide evidence for strength independence of RIF. (Study 2 and 3)**

The same critical difference, as described by Thesis II, between inhibition and interference theories leads to another potential decisive test. Interference theories suggest that it is the amount of cue-target strengthening that determines the cue-competitor recall probability. It follows that the way these associations are strengthened is irrelevant for RIF. Contrary to this suggestion, inhibitory theories assume that RIF can occur independently of cue-target strengthening (i.e. RIF is strength independent). What matters here, is the amount of competition competitors induce at the time of cue-target retrievals.

There is abundant evidence (see Anderson & Levy, 2011, Storm, 2011) that repeated encounters involving repeated study of cue-target associations and those involving repeated retrieval of targets to cues have dissociable effects on cue-target and cue-competitor recall. They lead to the same amount of cue-target strengthening, but only repeated retrieval leads to cue-competitor forgetting. These results clearly favor the inhibitory account over the interference account. However, an auxiliary hypothesis in these tests, that repeated retrieval and repeated study lead to the same amount of strengthening is in sharp contrast with results on the testing effect (e.g. Karpicke & Roediger, 2008). These studies rather show that repeated retrieval has a mnemonic advantage that becomes evident only after longer delays (usually at least around two to seven days). Study 3 tested the auxiliary hypothesis (repeated retrieval and repeated study strengthen memories to the same extent). It found that indeed, in the retrieval practice paradigm as well, repeated retrieval leads to greater memory strength
than repeated study. This advantage was only evident after a delay of five days, and was coupled with a RIF effect at short delays. This suggests that although RIF is dependent on competitive retrieval, as proposed by the inhibitory theories, it is also strength dependent, as suggested by the interference theories.

**Thesis IV. Repeated retrieval enhances later retrieval by stabilizing activity patterns in a network of brain areas related to cognitive control. (Study 4)**

Different manipulations on cue-target associations, e.g. repeated study (restudy) versus repeated retrieval (retest) have been shown to have differential effect not only on cue-competitor recall (Anderson et al., 1994) but also on cue-target recall (Roediger & Karpicke, 2006a). The positive effect on cue-target recall after retest compared to restudy (i.e., the testing effect) becomes evident only after delays of at least a few days. Previous neuroimaging studies on RIF (Kuhl et al., 2007, Kuhl et al., 2008) have shown that retesting cue-target associations led to changes in cognitive control related areas that were correlated with RIF (i.e., cue-competitor recall). Other ERP studies have added to this by showing that restudying cue-target pairs does not lead to similar changes, and compared to retest trials, activity during restudy trials was not correlated with RIF (e.g. Johansson et al., 2007; Wimber et al., 2008). However, neither of these studies has related neural activity during retests to changes in cue-target associations that are observable over long delays. Theories of the testing effect are yet scarce, and neuroimaging support for any of these theories is lacking.

Study 4 was one of the first attempts to investigate neural correlates of the testing effect. To address this issue, we measured brain activity using event related fMRI during a final cued recall test in brain areas related to cognitive control (localized with an n-back task).

Previously to this final test of cue-target associations, participants had learnt cue-target associations in either retest or restudy trials. This pretest learning occurred either twenty minutes or a week before the final test. We found that cognitive control related brain activity
was larger for restudied associations than retested ones on the short term. Importantly though, activity during cued recall of restudied associations decreased to a great extent on the longer run, whereas activity during cued recall of retested associations remained the same. We suggested that the testing effect, i.e., the long-term advantage of retest over restudy, is due to the stabilization of activity patterns in areas of the brain related to cognitive control.
Studies

Study 1

Inhibition and interference in the think/no-think task

Mihály Racsmány · Martin A. Conway · Attila Keresztes · Attila Krajcsi

© Psychonomic Society, Inc. 2011

Abstract Five experiments using the think/no-think (TNT) procedure investigated the effect of the no-think and substitute instructions on cued recall. In Experiment 1, when unrelated A–B paired associates were studied and cued for recall with A items, recall rates were reliably enhanced in the think condition and reliably impaired below baseline in the no-think condition. In Experiments 2 and 5, final recall was cued with B items, leading to reliably higher recall rates, as compared with baseline, in both the think and no-think conditions. This pattern indicates backward priming of no-think items. In Experiments 3 and 4, the no-think instruction was replaced with a thought substitution instruction, and participants were asked to think of another word instead of the studied one when they saw the no-think cued items. As in Experiments 1 and 2, the same amount of forgetting of B items was observed when A items were the cues, but in contrast to Experiment 2, there was no increase in the recall performance of A items when B items were the cues. These results suggest that not thinking of studied items or, alternatively, thinking of a substitute item to avoid a target item may involve different processes: the former featuring inhibition and the latter interference.

Keywords Inhibition · Episodic memory content · Backward facilitation · Priming · Executive control · Interference/inhibition in memory retrieval · Memory · Recall

Rememering is driven, channeled, or controlled by cues that feature in the retrieval process. This has been extensively explored in, arguably, one of its simplest forms, the cued recall of paired associates. A person who learns a list of unrelated A–B terms, such as bread–hat, when cued with the A term, bread is often able to recall the B term with which it was originally paired—that is, hat in this example (for reviews, see Baddeley, 1976; Crowder, 1976; Murdock, 1974; for a contemporary overview, see Kahana, Howard, & Polyn, 2008). Indeed, the principle that retrieval is based on specific cue–target associations—the cue being an item in the retrieval environment and the target a sought-for item in long-term memory—is so fundamental that it is virtually axiomatic to our understanding of retrieval processes (Thomson & Tulving, 1970; Tulving & Osler, 1968). Some recent and intriguing experiments have, however, demonstrated that cues might also be used to avoid, rather than access, items in memory with which they are associated (Anderson & Green, 2001; Anderson et al., 2004; Depue, Banich, & Curran, 2006; Depue, Curran, & Banich, 2007; Hanslmayr, Leipold, & Bauml, 2010).

In the think/no-think (TNT) procedure introduced by Anderson and Green (2001), a list of paired associates were first learned to a criterion such that participants could readily recall B terms when presented with A terms. Following acquisition, there then followed a practice phase in which an A term was presented and either its corresponding B term was thought about (the think condition) or participants were cued not to think about the previously paired B term (the no-
think condition). These TNT trials were repeated a number of times so that thinking and not thinking about associated B terms were practiced. There was also a subset of baseline control items that were neither thought about nor not thought about. The important finding in the subsequent cued recall test, in which A terms acted as cues to B terms, was that recall of B terms that had been thought about was high, recall of baseline items was intermediate, and recall of no-think items was reliably lower than baseline, suggesting inhibition of these items and showing how cues might be shaped to either promote remembering or hinder it.

These controversial results prompted a lively debate about the reliability and the possible explanations of TNT. Anderson and colleagues demonstrated a reliable amount of forgetting in the TNT procedure (see Anderson & Green, 2001; Anderson et al., 2004; for a summary of results from 687 participants, see Levy & Anderson, 2008), while Bulevich, Roediger, Balota, and Butler (2006), Hertel and Calcaterra (2005), Mecklinger, Parra, and Waldhauser (2009), and Bergström, Velmans, de Fockert, and Richardson-Klavehn (2007) were not able to reproduce the TNT effect. There are two alternative explanations for TNT and the memory effects observed in it (when present). According to an inhibitory explanation, favored by Anderson and colleagues, intentionally avoiding and practicing avoiding the recall of a specific target memory (no-think B items in the TNT task) inhibit the representation of the B item and so reduce access to the items in the test phase.

Additional support for this suggestion has come from the attenuation of recall performance for no-think items when cued with a so-called independent cue. If, in recall, a no-think B term such as hat is cued with clothing, a previously unpresented item, memory for the B term is still reliably lower than baseline (Anderson & Green, 2001; see also Bergström, de Fockert, & Richardson-Klavehn, 2009). However, in several studies in which the TNT procedure has been used, only tests where no-think items were tested with the original learning cues have been reported (Hertel & Gerstle, 2003); or, when independent probes have been used, the TNT effect has been absent (Algarabel, Luciano, & Martínez, 2006; Bulevich et al., 2006; Wessel, Wetzel, Jelicic, & Merckelbach, 2005). In our own laboratory, in unpublished studies, we have not been able to obtain the TNT effect using independent probes. One problem with the notion of independent probes is the assumption that they were not, in fact, present in the original study and/or practice phases. It is possible that these “independent probes,” which are always semantically related to the B terms they cue, were in fact activated when the B terms were processed at study and/or practice and have become part of the resulting memory representation of the list. If so, they might provide an alternative route to the “inhibited” item and so facilitate, rather than inhibit, recall (see Racsmány & Conway, 2006).

A second and alternative explanation of the TNT phenomenon is based on interference theory, which argues that the accessibility of items in memory can be reduced if there are other related or associated items in memory that compete and so interfere with access to and retrieval of a target item. Thus, it may be the case that following the no-think instruction during the practice phase, participants adopt a strategy of thinking of some other item—for example, another word (see Bulevich et al., 2006). Thinking about an alternative will create interference for the cue–target relationship similar to the interference seen in the well-established A–B, A–C procedure. Thus, learning bread–hat and then bread–lamp reduces the efficiency with which the A terms elicit the target B term. It is this interference that will, not surprisingly, cause attenuated recall performance for target items on the final test, and concepts such as inhibition need not then be invoked.

In the present experiments, we investigated both the inhibition and interference accounts of TNT. The inhibition account proposes that the effect of not thinking about selected B items in the practice phase leads to the inhibition of those items (Anderson & Green, 2001; Anderson et al., 2004). A strong prediction that follows is that in addition to being poorly recalled to A item cues, inhibited B items should themselves also be relatively ineffective cues to recalling A items. Experiments 2 and 5 tested this prediction. According to the interference account, the effects of thinking about alternative items (C items) to no-think B cues in the practice phase should lead to the poorer recall of B items in the test phase. In other words, the effects of not thinking or thinking about another item should produce identical effects in later recall. In Experiments 3 and 4, we tested this prediction too.

Experiment 1

The aim of this experiment was to replicate the original result of Anderson and Green (2001) and produce a reliable decrease in recall performance, relative to baseline, following eight cycles of suppression (not thinking). Pilot work indicated that, at least among our participants, eight cycles of suppression were sufficient to produce a robust no-think effect. We note that the TNT procedure has not always proved effective in attenuating later memory in the no-think condition (Bulevich et al., 2006), and for this reason, we wanted to establish that we could, in fact, obtain the effect.

Method

Participants Data were obtained from 31 native Hungarian speakers. We ran the experiment until we had data from 30 participants who reached the 51% learning criterion in five
cycles. (One participant did not reach the criterion and was not used for this reason.)

**Procedure and materials** Participants first took part in a learning task in which they were asked to learn 40 semantically unrelated word pairs. The stimuli consisted of 80 unrelated Hungarian words with a moderate word frequency, as measured by Szószablya, a Hungarian Web Corpus (Halácsy et al., 2004). The items were randomly paired and then inspected. Any related pairs were re-paired to produce the 40 unrelated paired associates (PAs). The PAs were randomly allocated to four sets of 10 assigned to the think, no-think, baseline, and filler conditions. All items were presented on a computer screen, and order of presentation in each phase of study, practice, and cued recall was random. The PAs were displayed individually in white uppercase letters for 5 s in the center of the screen. In the study phase, participants attempted to learn all the word pairs. Test–feedback cycles followed in which participants recalled the word pairs in a cued recall task. One cycle consisted of 40 cued recall trial cue–target pairs. On each trial, after the cue appeared on the screen, there were 5 s in which to recall the target word aloud. When a response was omitted or when the 5 s had passed, the target word appeared on the screen to the right of the cue word. The next trial followed with a 1-s intertrial interval. After all 40 cues had been presented, another test–feedback cycle followed, until a minimum of 51% of the word pairs had been correctly recalled.

After the learning phase, participants took part in the TNT practice phase and were given the following instructions: “You are going to see the left-hand side members of the previously presented word pairs in different colors on the computer screen. If you see a word in ‘Green,’ try to recall the other word previously seen together with this word and say it out loud. If you see a word in ‘Red,’ try not to think of the other word previously seen together with this word and do not say it out loud.” Participants first practiced this instruction with the filler words. There were eight cycles of this task.

After the learning phase, participants took part in the TNT practice phase and were given the following instructions: “You are going to see the left-hand side members of the previously presented word pairs in different colors on the computer screen. If you see a word in ‘Green,’ try to recall the other word previously seen together with this word and say it out loud. If you see a word in ‘Red,’ try not to think of the other word previously seen together with this word and do not say it out loud.” Participants first practiced this instruction with the filler words. There were eight cycles of this task. After the learning phase, participants took part in the TNT practice phase and were given the following instructions: “You are going to see the left-hand side members of the previously presented word pairs in different colors on the computer screen. If you see a word in ‘Green,’ try to recall the other word previously seen together with this word and say it out loud. If you see a word in ‘Red,’ try not to think of the other word previously seen together with this word and do not say it out loud.” Participants first practiced this instruction with the filler words. There were eight cycles of this task.

Results

The 30 participants who finished the experiment reached the learning criterion in 3.1 cycles ($SD = 1.15$). A one-factor ANOVA showed a main effect of item type, $F(2, 58) = 6.5$, $p < .01$. As can be seen in Table 1, row 1, the recall percentage for the no-think items was lower than that for the baseline items, and this effect was reliable, $F(1, 29) = 6.99$, $p < .01$. This finding shows attenuation and, possibly, inhibition of no-think items. The percentage of recalled baseline items was significantly lower than the percentage of recalled think items $F(1, 29) = 15.59$, $p < .01$, showing the benefits of rehearsal. These results are highly consistent with those of Anderson and Green (2001) and show a robust TNT effect.

**Table 1** Mean cued recall from five cued recall experiments using the think/no-think (TNT) task

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Cues at Test</th>
<th>Think</th>
<th>No-Think</th>
<th>Baseline</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experiment 1</td>
<td>A cues</td>
<td>.77 (.14)*</td>
<td>.61 (.15)</td>
<td>.70 (.17)</td>
</tr>
<tr>
<td>Experiment 2</td>
<td>B cues</td>
<td>.76 (.21)</td>
<td>.74 (.20)</td>
<td>.63 (.17)</td>
</tr>
<tr>
<td>TNT With Substitution: Experiments 3 and 4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experiment 3</td>
<td>A cues</td>
<td>.83 (.11)</td>
<td>.57 (.21)</td>
<td>.70 (.14)</td>
</tr>
<tr>
<td>Experiment 4</td>
<td>B cues</td>
<td>.84 (.17)</td>
<td>.66 (.21)</td>
<td>.71 (.14)</td>
</tr>
<tr>
<td>Standard TNT: Experiment 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experiment 5</td>
<td>B cues</td>
<td>.83 (.13)</td>
<td>.78 (.19)</td>
<td>.70 (.15)</td>
</tr>
</tbody>
</table>

*Standard deviations are shown in parentheses. ---

**Experiment 2**

This experiment used the same procedure and analysis as in Experiment 1, with the following single modification: In the final cued recall test, target words (B items) served as cues, and cue words (A items) were to be recalled. Thirty new right-handed native Hungarian speakers with normal or corrected-to-normal vision were recruited for this experiment. The mean age was 22 years (range, 19–26), and there were 20 women and 10 men. In all other respects, the design and analyses were identical to those in Experiment 1. Data were obtained from 30 native Hungarian speakers. All participants reached the 51% learning criterion in 5 cycles ($M = 2.9$, $SD = 1.29$).

The one-factor ANOVA again showed a reliable main effect of item type, $F(2, 58) = 9.2$, $p < .01$. However, as can be seen from Table 1 (second row), the recall percentage of the no-think items was significantly higher than that of the baseline items $F(1, 29) = 5.9$, $p < .05$. The percentage of recalled baseline items was significantly lower than the percentage of recalled think items, $F(1, 29) = 6.99$, $p < .01$. There was no reliable difference between no-think and think items. Thus, B items in the no-think condition can be effective cues to the recall of A items, as effective as A items are to B items in the standard procedure.
In Experiment 1, the association $A \rightarrow B$ in previously learned unrelated paired associates was attenuated by not thinking about $B$, given $A$, being repeatedly practiced. Later cued recall, using $A$ items as cues, showed recall of $B$ items to be reliably lower in the no-think condition, as compared with recall of baseline and practiced items (Table 1). In contrast, in the practice or think trials, the association $A \rightarrow B$ was strengthened, and recall was found to be higher than baseline following the think trials. One explanation of this pattern of recall is that the effect of the no-think trials is to inhibit the previously acquired $B$ items—hence, the lower recall. The effect of the think trials, on the other hand, is to strengthen, by rehearsal, the representation in memory of the $A$–$B$ pairs and so enhance their later recall (Anderson & Green, 2001). However, in Experiment 2, it was found that when $B$ items were used as cues, recall of $A$ items in the no-think condition was as high as recall of $A$ items in the think condition, and both were reliably higher than baseline. Assuming that the effect of the no-think trials was the same in both experiments, and given that they were identical in other respects, it cannot be the case that $B$ items are themselves inhibited. Indeed, the level of recall of items in the no-think condition suggests that the association of $B$ items to $A$ items is, in fact, primed.

An alternative explanation, and our original hypothesis, is that it is the relation between the word pairs that is affected by the no-think trials and, in particular, the unidirectional relationship $A \rightarrow B$ (Hertel & Calcaterra, 2005). Thus, when $A$ items are used as cues to memory for no-think items, they are comparatively ineffective (Experiment 1). In sharp contrast, quite the reverse was found when $A$ items, in the no-think condition, were cued by $B$ items in the present experiment. This suggests that the association $A \leftarrow B$ is primed by the no-think trials but that the relation $A \rightarrow B$ is suppressed by them. Thus, paradoxically, an item can be inhibited and primed at the same time, depending on its association with other items.

However, there is an alternative explanation of these findings that derives from interference theory: In the practice phase, when presented with an $A$ cue (e.g., grass) in a no-think trial, participants may avoid thinking of the target word orange by thinking of a different word (e.g., kiwi). To the extent that this occurs, it constitutes a version of the $A$–$B$, $A$–$D$, interference procedure, and $B$ items, such as orange, become less retrievable to $A$ cues, because the $A$ cues are associated with more $B$ items that compete for and, in the process, interfere with retrieval. Thus, an interference account of the low memory performance following the no-think trials in Experiment 1 is a possibility. How an interference account would apply to the enhanced memory levels following the no-think trials and recall cued with $B$ items in Experiment 2 is, however, not clear (see Table 1). Assuming that participants routinely and covertly generate alternative words to $B$ items on no-think trials, then, according to interference theory, these $B$ cues should be less efficient in accessing $A$ items simply because of the $A$–$B$, $A$–$D$ relations present in memory. Because, in Experiment 2, memory levels for $B$-cued no-think items were significantly above baseline and equivalent to memory levels for think items, it would seem, following the reasoning above, that $B$ items are effective cues of $A$ items despite the interfering effects of covertly generating an alternative word on the no-think practice trials. Given the paradoxical nature of these findings, it was decided to explicitly investigate the suggestion that participants achieve not thinking about or not retrieving a cued word by blocking retrieval with an alternative. In the following two experiments, we tested this idea by replacing the no-think instruction with a thought substitution instruction. This simply required participants to think of another word whenever they saw an item that was cued not to be thought of and spoken about. Note that this procedure was also used by Bergström and colleagues in an event-related potential study (Bergström et al., 2009). They found that their participants with the substitution strategy produced cue-dependent but no cue-independent forgetting, in contrast to participants with the standard no-think instruction (Bergström et al., 2009).

**Experiment 3**

Experiment 3 was identical in all aspects to Experiment 1, with one crucial modification in the instructions given for the think/no-think phase. For words appearing in green, participants were given the same instructions as in Experiment 1. For red words, however, participants were given the following instructions: “When you see a word in red, say out loud the first word that comes to your mind that this red word reminds you of.” So, for example, if the $A$ item was orange, the word apple might be generated. Further instructions emphasized that the word the cue had previously been paired with (the original $B$ item) should not be spoken. Data were obtained from 33 native Hungarian speakers, 30 of whom reached the 51% learning criterion in 5 cycles. Three participants did not reach the criterion and were not used, for this reason. The 30 participants who completed the experiment reached the learning criterion in 2.5 cycles ($SD = 1.13$). Their mean age was 20 years (range, 19–22), and 6 were females.

The ANOVA was the same as that used previously, with item type the single within-subjects variable consisting of three levels: baseline, think, and substitute. Mauchly’s test of sphericity was significant, $\chi^2(2) = 8.78, p < .05$; therefore, we used degrees of freedom corrected with Greenhouse–Geisser estimates of sphericity ($e = .71$). Item type had a significant effect on recall performance, $F(1.57, 4.67, 1.43) = 3.46$, $p = .04$; the differences between think and substitute were not significant. However, the interaction of item type with repetition was significant, $F(2, 58) = 4.26$, $p = .02$; the differences between think and substitute were not significant. The interaction of item type with repetition was also significant, $F(2, 58) = 4.26$, $p = .02$; the differences between think and substitute were not significant.
45.54) = 25.19, \( p < .001 \). As can be seen in Table 1, row 3, the average recall percentage for the think items was higher than that for the baseline items, while recall of substitute words was lower than the baseline. Planned contrasts confirmed that recall in the think condition was significantly higher than baseline, \( F(1, 29) = 28.29, \ p < .001 \), and recall in the substitute condition was significantly lower, \( F(1, 29) = 9.61, \ p = .01 \). These findings then mirror those of Experiment 1 (see Table 1, rows 1 and 3).

Experiment 4

Experiment 4 was the same as Experiment 2, but with the no-think instruction replaced with the same generate-a-substitute instruction as in Experiment 3. Data were obtained from 32 native Hungarian speakers. Two participants did not reach the learning criterion in 5 cycles and so took no further part. The 30 participants who finished the experiment reached the learning criterion in 2.27 cycles (SD = 1.33). Their mean age was 21.4 years (range, 19–26), and 6 were female. A reliable effect of item type was observed, \( F(2, 58) = 17.58, p < .001 \). As can be seen in Table 1, row 4, recall of items in the think condition was higher than baseline, but recall in the substitute condition was not substantially different from baseline. Planned contrasts confirmed that only recall in the think condition differed significantly from baseline, \( F(1, 29) = 24.9, \ p < .001 \). It can be seen in Table 1 that mean recall in the substitute condition was lower than baseline, but this was not found to be a reliable difference.

Experiment 3 demonstrates that virtually exactly the same effect can be produced by thinking about a substitute item as by not thinking about a target item in memory (see Table 1, rows 1 and 3). In contrast, thinking about a substitute item when the item substituted is subsequently used as a cue does hurt memory (Experiment 4), as compared with simply not thinking about an item that is later used as a cue, where recall is facilitated (Experiment 2). This suggests that different processes might underlie not thinking versus thinking about a substitute. The results of Experiments 3 and 4 only partially replicated the results of Bergström et al. (2009), who found that using substitution, instead of a no-think strategy, produced the same cue-dependent effect, but only the no-think strategy produced cue-independent forgetting. The results of Experiments 3 and 4 support our original hypothesis that not thinking of a specific target, when presented with its cue, harms only this specific cue–item relationship and primes all other relationships of this specific target item; in other words, the no-think effect is not independent of the retrieval cue. This is not the case for the substitution strategy, which probably alters the cue–target relationship by generating interference for this cue, and hence, participants will not access and prime the target items during the TNT phase.

Experiment 5

One problem with the findings above, and it is a problem in all TNT studies, is that the baseline levels of performance frequently shift across experiments. So, for instance, the baseline level of performance in Experiment 2, above, was considerably less than the baseline level of performance in the other experiments. If the baseline in Experiment 2 had been similar to the baseline in the other experiments, our main results may not have been reliable, and there would be no significantly higher recall of no-think items when B items are used as cues, relative to baseline. Why baselines vary from experiment to experiment and across studies, too, is not known, but it seems likely that there may be many factors in play relating to participants, environment, slight variations in procedure, time of day, and other uncontrolled chance influences. It is, therefore, possible that in Experiment 2, we observed reliable above-baseline recall of no-think items simply because of a baseline that was low by chance.\(^1\)

To exclude this possibility, in Experiment 5, we repeated Experiment 2. In this control experiment, all aspects of the procedure, design, and analyses were identical to those in Experiment 2, with one single exception: A new set of word stimuli were used. These were a set of word pairs taken from other TNT studies in our laboratory. These word pairs had consistently produced a TNT across several studies. We decided to use a different material because we wanted to show that the effect we found is reliable over different materials, too, (even if we failed to reproduce a baseline similar to that in our other experiments). Also, in this experiment, we used a questionnaire (a Hungarian version) developed by Bulevich et al. (2006), in order to exclude participants who did not follow the TNT instructions. Data were obtained from 46 native Hungarian speakers. The mean age of participants was 21.6 years (range, 18–30), and 13 were female. One participant did not reach the 51% learning criterion in 5 cycles and was not used. The 45 participants who finished the experiment reached the learning criterion in 1.89 cycles (SD = 0.93). More participants were included on the assumption that some would have to be excluded on the basis of their questionnaire responses. On the basis of questionnaire responses, data from 8 participants were excluded. But note that including this excluded data in the analyses did not change the pattern of results. A significant effect of item type was observed, \( F(2, 72) = 17.07, p < .001 \). As can be seen in Table 1, row 5, recall of both think and no-think items was higher than baseline. Planned contrasts confirmed that just as in Experiment 2, these differences were significant \( F(1, 36) = 43.40, \ p < .001 \), for the contrast between think and baseline items, and \( F(1, 36) = 7.86, \ p < .01 \), for the contrast between no-think and

\(^1\) We thank an anonymous reviewer for pointing this out.
baseline items]. Note that this pattern of findings exactly replicates the findings of Experiment 2 with a higher baseline. Baseline performance notwithstanding, then, the critical effect observed in Experiment 2, of B items priming recall of A items, is robust.

Additional analyses

The experiments above were conducted in relatively simple between-subjects designs, with each successive experiment changing a variable of theoretical interest. We adopted this approach in order to ensure compatibility with the original TNT experiments (Anderson & Green, 2001). One drawback to this approach is that of changing baselines from study to study. However, given that the changes between the experiments were in the experimental variables and all other conditions remained the same—for example, different groups of participants in the different experiments were drawn from the same pool of participants, all of similar ages, educational levels, and socioeconomic backgrounds; the experiments were conducted in the same laboratory at the same time of day by the same experimenters; and stimuli were held constant—it seems reasonable to treat Experiment 1–4 as a single experiment. In this analysis, a mixed design $2 \times 3 \times 2$ ANOVA was employed in which instruction (suppress vs. substitute) formed a between-subjects variable and item type (think, no-think, and baseline) and cue type (A cues and B cues) formed within-subjects variables. A strong and highly reliable (observed power of .919) interaction of instruction with item type was found, $F(2, 232) = 6.85, p < .001$, highly consistent with the earlier analyses. Also reliable (observed power .943) was the item type × cue type interaction, $F(2, 232) = 7.57, p < .001$, demonstrating across experiments impaired memory for no-think B items when cued with A items after either suppress or substitute instructions, and the reverse when recall of A items were cued with no-think B items after suppress but not after substitute instructions. Exploring these interactions further with planned contrasts of think and no-think items with baseline, we found that the cue type × item type interaction effect was due to the differential effect of the forward versus backward cue manipulation on the no-think items ($p < .001$; power, .936), and not the think items ($p = .38$). Similarly the interaction × item type interaction was due to the differential effect of the suppress versus substitute instruction manipulation on the no-think items ($p < .05$; power, .71), and not the think items ($p = .32$). This overall analysis confirms that despite changing baselines, the pattern of reliable effects is consistent over analyses.

General discussion

Two important findings emerged in these experiments. The first is that recalling two associated items can be simultaneously attenuated or primed depending on how the association is accessed (Experiments 1, 2 and 5). The second is that not thinking about a target item, as compared with thinking about an alternative, can produce the same decrements in cued recall (Experiments 1 and 3) or, sometimes, differences (Experiment 4). These findings are summarized in Fig. 1, and here we consider each in turn and their implications for the nature of the underlying memory representations that mediate them.

Episodic inhibition and the representation of paired associates

According to our account of episodic inhibition (Racsmány & Conway, 2006) in TNT and procedures like it, partic-
plicants first form an episodic memory of the study phase that contains some of the items activated during study, contextual, and possibly other associated information (Conway, 2009; Kahana et al., 2008). During the practice phase, items represented in the episodic memory of the study phase are accessed or access is resisted, and this establishes a pattern of activation/inhibition over the contents of the memory. In other words, the effects of selectively thinking and not thinking about different items alters their activation levels to render them highly accessible or comparatively inaccessible. This pattern of accessibility subsequently determines performance in the cued recall test phase. Items highly activated (think items) are readily accessible and can be recalled to a high level. Items activated but not so strongly can be recalled to a moderate level (baseline items), and items that are inhibited (no-think items) are difficult to access and, as a consequence, are recalled to the lowest levels. Thus, it is the pattern of activation/inhibition over the contents of the episodic memory of the study phase resulting from the effects of the practice phase that determines the various levels of cued recall.

What is clear from Experiment 2/5 is that this account needs modifying because, when no-think B items are used as cues at test, they lead to high levels of recall of associated A items. In fact, they can lead to levels of recall equivalent to recall of the think items, indicating priming of no-think B items (Experiment 2/5; see Fig. 1). It would be paradoxical to propose that an item in memory could be simultaneously inhibited and primed, and we certainly do not propose this. Rather, we consider how the nature of the underlying representations in memory could support such an apparently contradictory finding. In earlier thinking in PA learning, the A–B relation has been viewed as associatively symmetric (see Asch & Ebenholtz, 1962). In a recent review, Kahana et al. (2008) concluded that although there is some evidence that the A–B relation may be associatively asymmetric, the evidence overwhelmingly favors the symmetric view. In further support of this, a recent study (Carpenter, Pashler, & Vul, 2006)4 found that under certain practice conditions, cuing with either term, A or B, enhanced recall of the other. Thus, a model of the representation of PAs may take the form of A ←→ B. In this model, there is a single bidirectional connection between the representation of the A and B terms in the PA. The present findings suggest, however, that this model, too, requires modification.

The finding that no-think B terms can be inhibited when cued by A terms but facilitate recall of A terms when they themselves are used as cues indicates that the B term’s representation in memory cannot be inhibited. This is a finding and conclusion that runs counter to other accounts of inhibition in the TNT task (e.g., Anderson & Green, 2001) that posit inhibition of no-think items. Instead, it might be proposed that what is inhibited is the bidirectional link between A and B, A ←→/→B, while the representations of the two terms remain at some raised level of activation. But this, too, fails to account for the effectiveness of no-think B items in cuing recall of A items (see Fig. 1). The model that seems to us to account for the findings is one in which the associations A → B and A ← B are both independently represented in an episodic memory of the study phase. It may be that the repeated practice in list learning during the initial study phase facilitates the development of a memory representation in which independent unidirectional links exist among representations of PAs in a specific and detailed episodic memory created during the learning trials (see Conway, 2009, for a recent account of specific episodic memories).5

Assuming that a memory resulting from the study phase contains A → B and A ← B representation of PAs, the effect of the practice phase might be as follows: The think trials raise the activation levels of all items and their various associations, making them more accessible to retrieval processes and, eventually, leading to high levels of recall. The no-think trials decrease activation of the A → B association while increasing activation of the items themselves and of their other associations—for example, A ← B. This may occur because in order to decrease activation of, or inhibit, the relation A → B, both items must be accessed, as must other associations between them that are not targeted by no-think strategies for attenuation.6 If this is the case, recall of B given A will be attenuated, whereas recall of A given B will be facilitated. Essentially, this explanation posits inhibition of the unidirectional association A → B, while all other representations in association with the memory of the A–B pairs remain activated above the activation levels of baseline items (see Grison, Tipper, & Hewitt, 2005, for a similar explanation of negative priming). Furthermore, this model of independent associa-

---

4 Indeed, one interesting manipulation suggested by this would be to have learning trials that alternate between learning B given A and A given B and explicitly foster memory representations in which the two terms are associated by independent unidirectional links that together act as a (virtual) bidirectional link. Selective priming/inhibition following later processing of the list items in memory might be optimized by such a procedure.

5 Note that this may be conscious on some trials, particularly on the first few no-think trials, and on later trials become nonconscious. Interestingly, a pattern like this is seen in the ‘White Bear’ procedure (Wegner, 1994), where not thinking about the concept of a white bear for a 5-min period is marked by strong intrusions in the first 2 to 3 min but by virtually no intrusions in the last 2 or so minutes of the 5-min period.

3 We thank Henry Roediger III for drawing this work to our attention and for a number of other important comments and suggestions that helped develop the present article.
tions, A→ B and A← B, not only explains the effectiveness of “inhibited” no-think B cues in the recall of A items, but also preserves associative asymmetry, since any pair of unidirectional associations can act together as a bidirectional association.

One further feature of the model is that because inhibition is assumed to be directed at associations between representations of A and B terms, it is possible for representations of the terms themselves to remain above some resting level of activation, as can other associations between them not targeted for inhibition. For instance, the PA bread–lamp might be represented with independent associations, as described earlier, but also with other, additional (semantic) associations. Consider the case where, quite spontaneously and as part of processing not controlled in the study phase, the B term lamp has, in memory, the associations lamp → light and lamp ← light. If, at test, the A cue bread were now substituted with the cue light, a so called independent cue (Anderson & Green, 2001), there would be no inhibition and, instead, light would cue recall of lamp. This would occur, according to the independent associations view, because the representation of light in the episodic memory is above a resting level of activation and so are its other associates (to varying degrees). This line of reasoning may explain why it has proved so difficult to produce inhibitory effects with semantically associated “independent cues” (see, e.g., Bulevich et al., 2006).

Inhibition and interference

The two main competing accounts of the TNT effect posit that no-think items are hard to recall because they are inhibited (Anderson & Green, 2001) or because access to them in memory is blocked by substitutes covertly generated during the practice phase (Bulevich et al., 2006; Hertel & Calcaterra, 2005). Experiment 3 in the present series found, definitively, that explicitly generating substitutes can produce a TNT effect that is indistinguishable from that often observed (see Fig. 1). Given that this is the case, it seems reasonable to ask how the two views could ever be distinguished.

One way might be to simply ask participants what they are aware of doing when they encounter a not-to-be-thought-about item. Levy and Anderson (2008) reported some data on this, and we routinely ask our own participants. The predominant reply is that they “just go blank”; importantly, very few participants ever report thinking about other words. Indeed, thinking about substitutes in the practice phase is a difficult task, as participants in Experiments 3 and 4 all reported. Also relevant here are the findings of Experiment 4, in which a substitution strategy did not produce effects that paralleled those of Experiments 2. In Experiment 4, generating substitutes and then being cued to recall A items to (blocked) B cues did not lead to the striking and reliable increase in recall observed in Experiments 2 and 5 (see Table 1). Experiment 4 found that using substitution rather than no-think, B-cued recall of no-think (substituted) items did not reach the level of think items; indeed, it was reliably lower but did not differ from baseline. Perhaps, what is occurring in the substitution task is an attenuation of B items, rather than an inhibitory dysfacilitation/weakening of the representation of the AB associations. In the substitution task, B items become associated in memory with their substitute, and during cued recall, the substitute competes for recall with the B items, causing interference and attenuating access to A items. Interestingly, however, this interference is not sufficiently strong to reliably depress B-cued recall of A items below baseline. On the other hand, the interference was strong enough to reduce A-cued B substitute items below baseline (Experiment 3; see Table 1). Why this is so and why this pattern is so strikingly different from that in
Experiments 1 and 2 are unclear. One possibility is that when B is the cue, accessing B representations in memory is not as attenuated as when A is the cue. This may be because, when A is the cue, a more complex discrimination must be made during retrieval.

Whatever the case, the patterns of cued recall seen in Experiments 1 and 2/5 are determined by the nature of activation/inhibition over the contents of an episodic memory of the study list, as described earlier, whereas the patterns of cued recall observed in Experiments 3 and 4 are a product of interference in access caused by representations of substitute items and their associations in memory with representations of B items. In other words, the comparatively poor performance observed in the no-think conditions can be caused by either inhibition or interference, with interference somewhat less effective in depressing recall than is inhibition, at least in the present experiments. Furthermore, it may be possible to distinguish inhibition and interference by examining the processing that inhibited versus blocked items can differentially contribute to—that is, in acting as cues to associated items (Experiment 2). The positive effects of B items in the recall of A items are not as strong when other items and associations are represented with B items.

In conclusion, the present findings suggest that the locus of inhibition in the TNT task is not the representation of the items themselves in memory but, rather, the associations between them and, in particular, the A→B association. Using a substitute rather than a no-think task can produce identical effects (Fig. 1), but a substitute task produces different effects from a no-think task when B items are used as cues. Taken together, the latter findings suggest that both inhibition and interference can hurt memory in similar ways but differ in their wider effects.

Author’s Note Financial support for the research was provided by OTKA (Hungarian National Science Foundation) K84019. Martin A. Conway is supported by a Professoral Fellowship, RES-051-27-0127, from the Economic and Social Research Council of Great Britain.

References


Study 2

Interference resolution in retrieval-induced forgetting: Behavioral evidence for a nonmonotonic relationship between interference and forgetting

Attila Keresztes · Mihály Racsmány

Published online: 21 December 2012 © Psychonomic Society, Inc. 2012

Abstract Retrieving memories renders related memories less accessible. This phenomenon, termed retrieval-induced forgetting (RIF), is thought to be the result of processes that resolve interference during competitive retrieval. In several studies, researchers have manipulated the level of interference to test different theoretical accounts of RIF (e.g., inhibitory vs. non-inhibitory). However, the nature of how interference and RIF are related has not been systematically investigated. Here, we introduce a design that allows for assessing interference during competitive retrieval by measuring the recall RTs associated with target recall. Using such a design, we found that RIF occurred only when interference during competitive retrieval reached moderate levels, but not when it was too low or too high. This finding might indicate that low levels of interference do not trigger interference resolution, whereas interference resolution might fail when the interference reaches extremely high levels.

Keywords Retrieval · Forgetting · Retrieval-induced forgetting · Interference · Inhibition

Interference as a cause of forgetting has long captured the interest of scholars of memory. One specific question that has resurfaced in scientific discussions has concerned the way that interference during memory retrieval is resolved (for a review of interference theories from this perspective, see, e.g., Anderson & Bjork, 1994). The focus in these discussions was not solely on how interference causes memory failures during retrieval. Rather, it centered around the consequences of interference resolution. What happens to memory representations when a target memory is to be retrieved in the face of competing memories? How do we achieve retrieval of the correct target memory, and what happens to competing memories? In a seminal study, Anderson (2003) suggested an executive process—analogous to response override—that resolves interference by weakening memory representations that interfere with target memories at the time of recall. This weakened representation would be evident in the decreased probability of recall of the interfering memory when it is tested at a later time. This model was the first to hypothesize an active executive process that can act to weaken memory representations so that those memories become less accessible for retrieval.

Early on, Anderson, Bjork, and Bjork (1994) devised a procedure, the retrieval practice paradigm, that can separate a recall phase at time $t$, when interference from competing memories has to be resolved, from a recall phase at time $t+1$, when the accessibility of these competing memory representations can be measured. In this paradigm, participants are shown several category–member word pairs (e.g., animal–tiger, furniture–couch, and animal–chicken) and then practice retrieval of half of the members from half of the categories with category-plus-stem cues (e.g., animal–ti…?). Anderson et al. (1994) reasoned that, during retrieval practice, nonpracticed members of practiced categories (e.g., chicken) would interfere with the recall of practiced members, and therefore have to be inhibited. This should be evident from later testing, and that was exactly what they found: Participants’ recall of nonpracticed members of practiced categories was worse than their recall of members of categories that had not appeared in the retrieval practice phase. Anderson et al. (1994) termed this effect retrieval-induced forgetting (RIF).

Inhibition, in this theoretical approach, is a process that operates when a relatively strong competing item interferes with the retrieval of a target memory. This approach involves three testable properties of RIF that are relevant for our study. First, RIF is interference dependent; that is, only items interfering with the retrieval of a target memory would suffer inhibition. Second, RIF is retrieval specific; that is, manipulating target strength without retrieval of the target would not induce competition-based forgetting. Third, RIF is strength independent; that is, even when targets are
The relationship between interference and interference resolution

Although the dependence of RIF on interference seems to be well established, we know little about how interference-resolving processes operate in the face of increasing interference. Proponents of different inhibitory accounts of RIF have conceived this relationship quite differently. Anderson (2003, p. 421) suggested that “The more strongly associated to the category an unpracticed competitor was, the more impairment was found.” This implies a linear function between interference and the result of inhibition: The more that an item interferes with retrieval, the more inhibition it suffers.

Bäuml, Pastötter, and Hanslmayr (2010, p. 1049) suggested that “very low levels of interference may not trigger inhibitory processes when competing material is retrieval practiced,” but otherwise the strength of the interference that a competing item causes during retrieval plays only a minor role in defining the level of forgetting. According to their interpretation, inhibition would be a process that kicks in only when interference reaches a certain threshold. They also theorized that, over this threshold level, the effect of inhibition would not change significantly with increasing interference.

Norman, Newman, and Detre (2007) programmed a neural network model of RIF in which increasing competitor strength increased the effect of inhibition, but over a certain point, this could lead to a decrease in its success (i.e., a decrease in RIF). Similarly, Anderson and Levy (2010) suggested that a positive linear relationship exists between the level of interference and inhibition demands, and a negative linear relationship between the level of interference and inhibition success. Together, these opposing relationships lead to a demand–success trade-off in which very low levels of interference do not lead to RIF, because the level of inhibition demand remains low, whereas interference can reach a level over which inhibition cannot be effective, resulting in above-baseline facilitation of competitors. The carryover assumption put forward by these authors states that RIF should be seen only for items that induce moderate levels of interference.

The lack of knowledge about the function relating interference to forgetting makes it hard to design tests that try to tap properties of RIF. Take a study that tries to provide evidence for interference dependence by including a group of items with strong taxonomic frequencies (supposedly inducing great interference) and another group with low taxonomic frequencies (supposedly inducing little interference; e.g., Anderson et al., 1994; Williams & Zacks, 2001). This study would reliably provide significant RIF differences between these two groups if the relationship between interference and forgetting was a simple linear one, as suggested by Anderson (2003). In the case of a threshold-like interference resolution process, as suggested by Bäuml et al. (2010), differences would only be found if the low-taxonomic-frequency words did not achieve a certain threshold at which inhibition kicks in. Moreover, if inhibition causes forgetting to decrease over a certain level of interference, as predicted by Norman et al. (2007) and Anderson and Levy (2010), one might see no differences between the two groups, because one of them could cause no interference at all, while the other one could cause too much interference. Studies using factorial designs might obtain contradictory results (see, e.g., Anderson et al. [1994] vs. Williams & Zacks [2001]) simply because the groups of words chosen to cause great or little interference are chosen on an arbitrary basis and without any knowledge of the underlying relationship between interference and the effect of interference resolution.

Another advantage of understanding how the effect of interference resolution changes as interference increases would be to design tests that are more sensitive to detect RIF. Such tests could focus only on memories that truly caused interference during memory retrieval, and thus that would be expected to suffer the results of interference resolution. Such sensitive tests would be very useful in settling some hot debates about the nature of interference-resolving processes in memory—for instance, to clarify whether RIF generalizes to novel, independent cues (for positive evidence, see, e.g., Aslan, Bäuml, & Pastötter, 2007; Levy, McVeigh, Marful, & Anderson, 2007; Saunders & MacLeod, 2006; for negative evidence, see, e.g., Camp, Pecher, & Schmidt, 2007; Perfect, Stark, Tree, Moulin, Ahmed and Hutter 2004).

Item-by-item RIF

Our goal in this study was to develop a test that could give an indication of how RIF changes as a function of competition during retrieval. Therefore, we needed a design that could provide data on how retrieval of each memory item was affected by interference during the retrieval practice phase. We set two objectives to achieve this goal. First, the design should be such that each item had an individual competitor that interfered with it. Second, we needed to collect data that at least indirectly would inform us about the amount of interference that a memory item suffers during its retrieval in the retrieval practice phase. For this second purpose, we chose to measure the reaction times (RTs) of target memory retrieval during the retrieval practice phase.

RTs have been used to measure the levels of interference caused by competing representations or processes in a number of paradigms—among others, negative priming (Tipper,
variability of these potential factors.

In the method section, we will discuss how we tried to control the interference. They are influenced by several other factors as well, such as target strength and number of competitors. However, RTs were slower if the participants had first generated other exemplars from the same category.

RT data have rarely been collected in RIF studies. Anderson (2003, p. 439) suggested that “when the measure of interference is reaction time, the presence of multiple competitors or a single strong competitor should slow the recall of a target.” Indeed, RTs have been used in RIF studies to measure the magnitude of the RIF effect (e.g., Racsmanő & Conway, 2006; Veling & van Knippenberg, 2004; Verde & Perfect, 2011).

In a similar vein, RTs have been used to measure interference during retrieval practice. In one study (Levy et al., 2007, Exp. 2), the participants were split into two groups according to the interference that a memory suffered during retrieval practice. In this study, participants had to name pictures in their second language and were tested later using the same pictures in their first language. Levy et al. performed a median split of their sample based on the overall RT differences between the participants’ performance in the first and second languages. The authors suggested that slower naming performance in the second than in the first language indicates poorer knowledge of the second language. On this basis, they hypothesized that participants with larger RT differences would need to resolve greater interference from the first language when naming pictures in their second language than would participants who have better knowledge of their second language. This would lead to greater RIF among poorer speakers than in the other group, and this is exactly what was found.

Kuhl, Dudukovic, Kahn, and Wagner (2007), measured RTs and activation in prefrontal areas during retrieval practice and correlated the amount of RIF with the decrease of these measures from the first to the third practice cycle. They found that the decreases in prefrontal activation, but not RTs, correlated positively with forgetting of the interfering memories. It is important to note that such a reduction is more a measure of successful interference resolution than of interference per se.

Here, we used a variation of the retrieval practice paradigm introduced by Anderson et al. (1994), in which only two items share the same category cue (and compete for retrieval) in every category. We did not manipulate interference in a factorial design, but rather used the retrieval practice RTs as an independent variable to assess the magnitude of interference. Of course, we do not assume that retrieval RTs only reflect interference. They are influenced by several other factors as well, such as target strength and number of competitors. In the Method section, we will discuss how we tried to control the variability of these potential factors.

Using such item-by-item RIF, we intended to reproduce findings supporting the interference dependence of RIF and to better understand how interference and the forgetting effect caused by interference resolution are related.

Method

Participants

A group of 64 students (age: $M = 21.81$ years, $SD = 2.12$; 32 women, 32 men) participated in the experiment for credit in partial fulfillment of an introductory psychology course requirement at Budapest University of Technology and Economics. The participants were tested individually in a quiet room in sessions that lasted for a maximum of 30 min. Due to a computer error, one participant could not complete the test phase. This participant’s data were excluded from the analyses.

Materials

We used 22 categories with two members in each category, making a list of 44 word pairs. To induce the competitive retrieval that is supposed to be necessary to produce RIF, and to avoid moderation of the RIF effect (see Anderson, 2003), we followed strict selection criteria. To produce any RIF effect, it would be essential to have items in a category that would interfere with each other. Integration has been shown to counteract the RIF effect robustly (Anderson & McCulloch, 1999), and reducing the number of elements in a category increases the chances of integration (e.g., Camp et al., 2007). Since we used only two members per category, we had to take care to reduce the chances of integration.

Frequency and association data were drawn from the open-source Frequency Dictionary of the Hungarian Webcorpus, developed by BME Média és Szociológia Tanszék–Media Oktatási és Kutató Központ (Media Research Centre at the Department of Sociology and Communications of Budapest University of Technology and Economics; BME-MOKK, 2003). For a full description of the database, see Halácsy et al. (2004) and Komai et al. (2006). We included categories that were not associated with each other (either semantically or phonetically) and that were themselves of moderate frequency. The category labels and targets were neutral words. Category members were moderate-frequency words, and within their category they had a moderate to high relative frequency. Category members that were either too typical or too rare were excluded. No member from a given category was associated with another member in another category, nor was it associated to another category cue. We made an effort to choose the two members of one category from different subcategories. To avoid cues that would uniquely refer to one
target in semantic memory during the test phase, the first letter of each target was shared with at least one other low- or moderate-frequency category member that did not appear in the experiment. In contrast, to avoid extraexperimental interference during retrieval practice, we excluded words whose first two letters could be completed to create another category member not seen in the experiment. The first two letters had a moderate versatility; that is, a moderate number of words could be generated from the same two letters from semantic memory. We made an effort to reduce the number of words in which the first two letters made up or contained a syllable of the word.

After filtering possible materials through these selection criteria, we had a list of 88 words, including four words belonging to each of the 22 categories. In order to reduce item-based confounds in the RT data, we wanted to create a final list that would produce the least variation in baseline retrieval RTs. To this end, we ran a pilot study in which participants learned all 88 category-member pairs and then performed retrieval practice on all of the items once. To obtain the final list to be used in our experiment, we excluded two items per category on the basis of the retrieval practice results in this study. Using recall RTs, we excluded words that produced RTs that either were more than one standard deviation away from the group mean or differed substantially (more than 1,000 ms) from the group mean RT of their category. Using recall accuracy, we excluded both words that were recalled by every participant in the pilot and words that were recalled by less than 33 percent of the participants (around the lower and upper deciles of the 88 words; see the Table 2 for the final list of word pairs selected.)

We used Presentation® software (Version 14.1, Build 09.21.09) for presentation of the stimuli and preanalysis of the data.

Design

Out of the 22 categories, two were used to provide filler items, and ten were categories from which no members were presented in retrieval practice (i.e., Nrp categories and targets). From the other ten (Rp) categories, one member (Rp+) was practiced during retrieval practice, leaving the other member nonpracticed (Rp−). Members of the Nrp categories were divided into Nrp+ and Nrp− items, which served as baselines for the Rp+ and Rp− items, respectively. For each participant, the categories (except filler categories) were randomly assigned to category types (Nrp vs. Rp), and members of each category (except filler items) were then randomly assigned to item types (+ vs. −). Fillers were from the same categories throughout the experiment.

Procedure

The participants went through four phases of the experiment; a study, a retrieval practice, a delay, and a test phase. In the study phase, participants were shown all 44 category–member pairs once on a computer screen and were asked to remember the members with the help of the category cues. In each trial, a category word appeared to the left of the middle of the screen, together with one of the words from that category to the right of the middle of the screen. The word pair was shown for 3,000 ms, followed by a 500-ms intertrial interval (blank screen). We opted for such a short presentation of the word pairs in order to further decrease the possibility of integration of items from the same category. The study list was pseudorandomly shuffled for each participant, with the constraint that the same category could not appear within five consecutive trials. Presentation of the study list started and ended with two of the filler category–member pairs.

When the study phase was finished, participants immediately received the instructions for the retrieval practice phase. This phase consisted of three cycles. In every cycle, all Rp+ items were presented for retrieval practice once, in a random order. In each trial, the participants saw a category cue to the left of the middle of the screen and the two-letter stem of the Rp+ member of that category. The instructions were to try to recall and report the correct answer. Participants were asked to press the response button (the Enter key on the keyboard) as soon as they had the answer in mind. In order to have a valid measure of how fast an item came to mind (and not just a measure of category familiarity or feeling of knowing), we told the participants that we were curious about how fast they could recall memories, and instructed them to act as if they were on a TV quiz show, where they could lose points if they pressed the response button but could not come up with an answer immediately. After pressing the button, they were asked to type in the answer. They had 8 s to do this. If they pressed the response button or exceeded the time limit, they were shown the subsequent trial. In the first cycle, participants had 6 s to report that they knew the answer, and in the following two cycles they had 4 s. If participants did not press the response button, the next trial was introduced. The retrieval phases also started and ended with two filler trials.

After retrieval practice, the participants engaged in a 5-min two-back task, which served as a delay before the test.

The test phase consisted of 44 trials that tested memory for all of the category members. This phase also started and ended with two of the filler items. Trials were presented in the same way as in the first retrieval practice cycle, except that the category-plus-word-stem cue contained only a first-letter stem of the category member. In order to avoid output interference effects (Anderson, 2003), the test phase involved two blocks. Rp− items and their controls were tested in the first block, followed by Rp+ items and their controls in the second block. Items were randomly intermixed within both blocks. The use of different control items for the Rp+ and Rp− items was necessary in order to circumvent baseline deflation (Anderson, 2003).
Results

During analysis, we used alpha set to .05 and corrected for multiple comparisons using Bonferroni correction. The retrieval practice success rates were 85%, 86%, and 89% in the three practice cycles, respectively. The final recall performance can be seen in Table 1.

To test whether our retrieval practice manipulation was successful, we performed a one-way repeated measures analysis of variance (ANOVA) on the final recall data, with four levels of item type: Rp+, Nrp+, Rp−, and Nrp−. Item type had a significant effect on final recall, $F(3, 186) = 88.11, p < .001$. To test for beneficial effects of practice on the practiced items and a detrimental effect of practice on the recall of competitors, we performed two post-hoc tests. Recall of Rp+ items was significantly better than recall of their Nrp+ baseline, $t(62) = 15.31, p < .001, r = .89$, and recall of Rp− items was significantly lower than recall of their Nrp− baseline, $t(62) = 2.46, p = .034$ (Bonferroni corrected), $r = .30$. This shows that our item type manipulation was successful and that it provided a strong practice effect and a medium-size RIF effect.

The primary target of our investigation was the relationship between the recall RTs of Rp+ items during practice and later recall of their Rp− competitors. We analyzed first-cycle RTs only because we assumed that variance in interference, and thus in the RT data, would be greatest in the first practice cycle and would be smoothed out during further practice.

In order to rule out cheating (i.e., pressing the button when the participant did not yet know the answer), we analyzed typing time (the time that elapsed between two Enter presses: the first indicating that participants knew the response, and the second indicating that they had finished typing) for each participant. This analysis showed that no participant had individual outliers in typing times, and therefore all successfully recalled Rp+ items were included in the analysis.

Within each participant, we ranked Rp+ items by their RTs measured during the first practice cycle. Then, on the basis of this rank, we split Rp+ items into tertiles with fast, moderate, and slow RTs. For each tertile, we calculated the recall rates of the corresponding Rp− items at test (see Fig. 1).

To test which of the Rp− tertiles contributed to the RIF, we ran an ANOVA on the final recall data with four levels of item type (Rp−1.tert, Rp−2.tert, Rp−3.tert, and Nrp−). In this analysis, Mauchly’s test of sphericity was significant, Mauchly’s

| Table 1 Mean recall percentages at test for the four item types |
|-----------------|-----------------|-----------------|-----------------|
| Item Type       | Rp+             | Nrp+            | Rp−             | Nrp−            |
| M               | .82             | .45             | .46             | .53             |
| SD              | .17             | .19             | .19             | .20             |

Fig. 1 Recall of Rp− items at test as a function of recall RTs for the associated Rp+ items during the first cycle of the retrieval practice phase. The empty rectangle on the right represents average Nrp− (baseline for the Rp− items) recall. The data are grouped into three tertiles according to the Rp+ RTs during the first practice cycle. Rp− recall was significantly below baseline only for second-tertile items, associated with moderate practice RTs. Error bars represent the standard errors of the means.

$W = .81$; therefore, degrees of freedom were corrected using Greenhouse–Geisser estimates of sphericity. Item type had a significant effect on final recall, $F(3, 186) = 2.85, p = .042$. Planned contrasts (Bonferroni corrected) showed that RIF was significant only for items corresponding to second-tertile Rp− items (corresponding to Rp+ items with moderate RTs), $F(1, 62) = 9.73, p = .006$. Rp− items corresponding to Rp+ items with fast and slow RTs also showed lower recall than baseline, but these differences were not significant: $F(1, 62) = 0.57, p = .99$, for Rp− items corresponding to Rp+ items with fast RTs, and $F(1, 62) = 2.28, p = .11$, for Rp− items corresponding to Rp+ items with slow RTs.¹

To assess the nature of the relationship between interference and the results of interference resolution, we conducted a repeated measures ANOVA on Rp− recall, with First-Cycle Rp+ Recall RT (fast vs. moderate vs. slow) as a within-subjects factor. We found a trend toward an effect of Rp+ RTs on recall of Rp− items, $F(2, 124) = 2.17, p = .118$, which was due to a tendency toward a quadratic trend in the final recall data, $F(1, 62) = 3.78, p = .057$, .99, for Rp− items corresponding to Rp+ items with fast RTs, and $F(1, 62) = 2.28, p = .11$, for Rp− items corresponding to Rp+ items with slow RTs.¹

¹ Originally, we ran the experiment with 32 participants. In this original experiment, the final recall percentages (with standard errors in parentheses) for Rp−1.tert, Rp−2.tert, Rp−3.tert, and Nrp− items were .46 (.05), .38 (.06), .44 (.05), and .51 (.04), respectively. Rp− recall was significantly below baseline only for second-tertile items with moderate practice RTs, $t(31) = 2.35, p = .038$, one-tailed (Bonferroni corrected). Because this experiment was, in essence, exploratory, in order to see that this result was not a Type I error, we extended the experiment with the inclusion of another 32 participants. Logically, this was an extension rather than a replication of the original experiment (same materials, same university population, same lab, same assistant). The pattern of results obtained from this extension replicated the results of the original experiment, and the extended experiment provided greater power in detecting the same effect: Only second-tertile Rp− items were recalled below baseline, $t(62) = 3.16, p = .006$ (Bonferroni corrected). The data presented here are pooled from all 63 participants (as described above; one of the participant’s data were excluded from the analyses).
indicating that there was one change in the direction of the relationship between Rp+ RTs during retrieval practice and the recall rate of the corresponding Rp− items at final test.

Discussion

We found practice and RIF effects with a variant of the retrieval practice paradigm that involved only two members per category. This was a novel finding, showing that the materials and design adopted here did not allow for integration of the two category members, an effect that might have masked RIF (e.g., Anderson, 2003).

Retrieval of target memories induced forgetting of competing items only when the targets were recalled with moderate retrieval RTs; RIF did not occur for competitors of memories that were recalled either too fast or too slow. This suggests that processes resolving interference during recall lead to forgetting when retrieval attempts produce moderate levels of interference.

Crucially, we showed that only a subsample of memories contribute to the RIF effect. Choosing the right sample of items to be included in the analysis might be critical for detecting the RIF effect. This can guide further investigations of the boundary conditions of RIF—for instance, when designing studies that test RIF’s cue independence.

As for the exact nature of the relationship between target recall RTs and later recall of competing memories, our study is not conclusive. Our data suggest that the direction of the relationship between interference and the recall of interfering memories changes at one point from negative to positive. This would then support the suggestion that RIF is an inverted-U-shaped function of interference (Anderson & Levy, 2010; Norman et al., 2007). However, since this was supported only by a statistical tendency, strong conclusions are not warranted.

One weakness of our study is that it is hard to find three data points that would lead to rejection of a U-shaped function. A better test of this type of relationship would be to analyze final recall data binned into quartiles instead of tertiles. However, the number of items in our study was too low to provide enough power to detect such an effect if the data were split into more than three bins.2

Future studies could clarify several further issues raised by our results. For instance, retrieval RTs are affected not only by the magnitude of the interference that has to be resolved during retrieval, but are influenced by a range of factors, such as target strength and the strength of the associations between category cues and targets, or the relative strengths of targets and competitors. To assess the differential contribution of these factors to interference during retrieval would require new methodologies.

Another interesting issue is that the use of RT data made it impossible to analyze the effect of interference during unsuccessful retrieval attempts. Storm, Bjork, Bjork, and Nestojko (2006) showed that even unsuccessful Rp+ retrieval contributes to Rp− forgetting. Therefore, an experiment based on a measure of interference that can be collected for both retrieved and nonretrieved Rp+ items might be a significant addendum to the pattern of results presented here.

We have provided converging evidence for the interference dependence of RIF, and suggest that interference-resolving processes cause forgetting of interfering memories at moderate levels of interference. This might provide evidence for both a theoretical model based on the carryover assumption of Anderson and Levy (2010) and the computational model of Norman et al. (2007), both of which suggest that the supposed nonmonotonic function relating interference to forgetting is the sum of two linear monotonic functions: one positive, relating interference and inhibition demand, and one negative, relating interference and the success of inhibition. Although our results seem to be in line with these theories, two caveats should be mentioned here. First, as noted earlier, converging evidence will be necessary to refute either of the models describing the relationship between interference and forgetting. Second, nothing in our data suggests that the interference-resolving process involves inhibition at all. Replicating our findings with independent cues would be a strong indication of the role that inhibition plays in resolving interference.

The leap of thought introduced in the seminal article of Anderson and Bjork (1994) was the shift of attention from interference as a cause of forgetting to the consequences of interference resolution (Anderson, 2003). Our results support the view that the amount of interference plays a role in how the retrieval probabilities of related memories are shaped for later retrieval. Our study also highlights the fact that using factorial designs alone might not be sufficient to fully understand the mechanisms of interference resolution. In recent years, we have gained considerable knowledge about how the brain implements interference resolution at the systems level (e.g., Kuhl et al., 2007; Wimber, Rutschmann, Greenlee, & Bäuml, 2009). The approach and results presented here may contribute to a better understanding of interference resolution at the level of cognitive processes.

Author note This work is connected to the scientific program of the “Development of Quality-Oriented and Harmonized R+D+I Strategy and Functional Model at BME” project. This project is supported by the New Hungary Development Plan (Project ID: TÁMOP-4.2.1/B-09/1/KMR-2010-0002) by OTKA (the Hungarian National Science Foundation) Grant No. K84019. We thank Péter Pajkossy for his useful comments on a previous version of the manuscript.
Appendix

Table 2  English translations of the categories and target words used in the experiment

<table>
<thead>
<tr>
<th>Category Cue</th>
<th>Target Member</th>
</tr>
</thead>
<tbody>
<tr>
<td>bird</td>
<td>gull</td>
</tr>
<tr>
<td>bird</td>
<td>pelican</td>
</tr>
<tr>
<td>body part</td>
<td>elbow</td>
</tr>
<tr>
<td>body part</td>
<td>front</td>
</tr>
<tr>
<td>cloth</td>
<td>gloves</td>
</tr>
<tr>
<td>cloth</td>
<td>pajamas</td>
</tr>
<tr>
<td>drink</td>
<td>hot chocolate</td>
</tr>
<tr>
<td>drink</td>
<td>lemonade</td>
</tr>
<tr>
<td>fish</td>
<td>herring</td>
</tr>
<tr>
<td>fish</td>
<td>trout</td>
</tr>
<tr>
<td>flower</td>
<td>geranium</td>
</tr>
<tr>
<td>flower</td>
<td>lily</td>
</tr>
<tr>
<td>fruit</td>
<td>ananas</td>
</tr>
<tr>
<td>fruit</td>
<td>prune</td>
</tr>
<tr>
<td>game</td>
<td>dominoes</td>
</tr>
<tr>
<td>game</td>
<td>hide-and-seek</td>
</tr>
<tr>
<td>illness</td>
<td>allergy</td>
</tr>
<tr>
<td>illness</td>
<td>cold</td>
</tr>
<tr>
<td>insect</td>
<td>butterfly</td>
</tr>
<tr>
<td>insect</td>
<td>tick</td>
</tr>
<tr>
<td>job</td>
<td>model</td>
</tr>
<tr>
<td>job</td>
<td>soldier</td>
</tr>
<tr>
<td>kitchen utensils</td>
<td>microwave</td>
</tr>
<tr>
<td>kitchen utensils</td>
<td>whisk</td>
</tr>
<tr>
<td>mammal</td>
<td>bear</td>
</tr>
<tr>
<td>mammal</td>
<td>elephant</td>
</tr>
<tr>
<td>material</td>
<td>aluminium</td>
</tr>
<tr>
<td>material</td>
<td>concrete</td>
</tr>
<tr>
<td>musical instrument</td>
<td>harp</td>
</tr>
<tr>
<td>musical instrument</td>
<td>synthesizer</td>
</tr>
<tr>
<td>nature</td>
<td>cliff</td>
</tr>
<tr>
<td>nature</td>
<td>swamp</td>
</tr>
<tr>
<td>office utensil</td>
<td>calculator</td>
</tr>
<tr>
<td>office utensil</td>
<td>xerox</td>
</tr>
<tr>
<td>spice</td>
<td>marjoram</td>
</tr>
<tr>
<td>spice</td>
<td>parsley</td>
</tr>
<tr>
<td>sports</td>
<td>horse riding</td>
</tr>
<tr>
<td>sports</td>
<td>triathlon</td>
</tr>
<tr>
<td>weapon</td>
<td>rifle</td>
</tr>
<tr>
<td>weapon</td>
<td>spear</td>
</tr>
<tr>
<td>country*</td>
<td>Argentina*</td>
</tr>
<tr>
<td>country*</td>
<td>Bulgaria*</td>
</tr>
<tr>
<td>tools*</td>
<td>pliers*</td>
</tr>
<tr>
<td>tools*</td>
<td>screwdriver*</td>
</tr>
</tbody>
</table>

The original experimental materials were in Hungarian. * Filler categories and filler items.

References


Study 3

Differential Long-Term Effects of Selective Retrieval Practice and Selective Restudy:
Common Processes Behind Retrieval-Induced Forgetting and the Testing Effect

Attila Keresztes and Mihály Racsmány
Budapest University of Technology and Economics, Budapest, Hungary

Author Note

Attila Keresztes, Department of Cognitive Science, Budapest University of Technology and Economics; Mihály Racsmány, Department of Cognitive Science, Budapest University of Technology and Economics.

This research was supported by the Hungarian National Science Foundation (OTKA), project id. K84019, and the New Hungary Development Plan, project id.: TÁMOP-4.2.1/B-09/1/KMR-2010-0002. We would like to thank Péter Pajkossy, Péter Simor, and Gyula Demeter for their useful comments.

Correspondence concerning this article should be addressed to Attila Keresztes, Department of Cognitive Science, Budapest University of Technology and Economics, Egry J. u. 1, T511, Budapest, 1111-Hungary
Phone: +361 463 3525, E-mail: akeresztes@cogsci.bme.hu
Abstract

Testing is very efficient at establishing long-term memories but, besides being painful, it might also induce forgetting of related memories. Research on this forgetting effect, known as retrieval-induced forgetting (RIF), has shown that retesting and restudying memories lead to similar amount of memory strengthening, but only retesting leads to RIF. These results have lent support to accounts that explain RIF by inhibitory mechanisms that act to suppress competing memories during testing. However, research on the testing effect has suggested that apparently equal memory strengthening after retest and restudy opportunities can in fact be different when memories are tested after a long delay. We tested long-term memory strengthening and RIF in one experiment. Participants studied cue-target associations, some of which were then either restudied or retested. A final test on all associations was administered after a delay of either five minutes or five days. Retrieval-induced forgetting occurred only when memories were retested and not when they were restudied. In addition, over a five-day delay, massive forgetting occurred for restudied memories, but no forgetting was observed for retested ones. This suggests that memory strengthening was larger for retested than for restudied memories. It seems that short-term costs and long-term benefits of testing result from processes that strengthen memories through reconstructing memory-sets during retrieval.

*Keywords:* forgetting, memory strength, retrieval-induced forgetting, testing effect
Differential Long-Term Effects of Selective Retrieval Practice and Selective Restudy:

Common Processes Behind Retrieval-Induced Forgetting and the Testing Effect

Testing is a strong memory enhancer, and is more beneficial for long-term retention than restudying (e.g. Karpicke & Roediger, 2008; Roediger & Karpicke, 2006; Tulving, 1967). Interestingly, under certain circumstances testing also induces forgetting of related non-tested material (for a review, see Anderson, 2003), a phenomenon termed retrieval-induced forgetting (Anderson, Bjork, & Bjork, 1994).

Theories explaining retrieval-induced forgetting (RIF) have focused on the potential role of executive processes in resolving interference during competitive retrieval (Anderson, 2003). For instance, the inhibitory control hypothesis (Levy & Anderson, 2008) proposes that interference is resolved through active inhibition of competing memory representations via executive processes. In contrast, non-inhibitory accounts (e.g. Shiffrin & Raaijmakers, 1992) assume that RIF is not the result of executive processes, but is simply due to the fact that the strength of memories accessed during retrieval increase and this renders them more accessible compared to competing memories.

In an effort to settle the debate, Levy and Anderson (2008) suggested that if one could show that strengthening of memories can occur without impairment of related memories, whereas retrieval practice always leads to such impairment, then it would suggest that retrieval practice, not strengthening is the cause of impairment, i.e. that RIF is retrieval-specific, and strength-independent. Indeed, several experiments have shown that restudying and retesting memories strengthen these memories to the same degree, but only retesting leads to RIF (for a summary of findings, see Anderson, 2003). Basing their study on the same rationale, Raaijmakers and Jakab (2012) provided evidence against the retrieval specificity and the strength-independence of RIF.
The common ground behind the interpretation of these opposing findings was that varying restudy vs. retest as the mode of practicing memories is a manipulation that only affects competition and retrieval mode during memory formation, and does not affect the magnitude of memory strengthening.

Research on the testing effect (e.g. Roediger & Karpicke, 2006), however, has shown that although recall after a short delay is similar for restudied and retested memories, recall after several days is better for retested memories than for restudied ones. This implies that although both restudying and retesting strengthen memories to some extent, they might have different effects on long-term memory representations.

The main question of our study was whether the testing effect generalizes to an experimental manipulation that is widely used to investigate processes leading to RIF, the retrieval-practice paradigm (Anderson et al., 1994). If it does, it would clearly imply that manipulating memory practice by introducing a restudy vs a retest condition in one experiment does not allow for dissociating memory strengthening and memory impairment, because choosing one type of practice would affect both strengthening and impairment. Such a result might question the validity of results that have lent support for or against the strength-independent account of RIF based on differential memory strength manipulations.

We used a variation of the retrieval-practice paradigm (Anderson et al., 1994) in which participants study several category–member pairs (e.g. animal–tiger, furniture–couch, animal–chicken, etc.) and then take repeated tests on half of the members of half of the categories (e.g. animal - t...?). Typically, a final test administered after a short delay reveals that retesting leads to forgetting of related material (e.g. 'animal - c...?') compared to baseline (e.g. furniture - c...?).

We introduced two modifications to this paradigm. First, we added a condition where participants, instead of taking repeated tests (retest condition), repeatedly studied the same
material (restudy condition). Second, final test was administered after either a short (five-minute) or a long (five-day) delay.

We expected restudy and retest to produce similar recall after a short delay, and that differences in memory strength, as reflected in greater forgetting of restudied than retested memories, will become apparent only after a long delay. Crucially, we hypothesized that memory strength, measured this way, would be related to RIF. This was a fairly straightforward hypothesis derived from the testing effect literature that nevertheless has not yet been tested.

Method

Participants
Since our hypothesis concerned finding no RIF in the restudy condition, we performed preliminary power calculations with G*Power (Faul, Erdfelder, Buchner, & Lang, 2009) using earlier published (Racsmány, Conway, & Demeter, 2010) and unpublished data from retrieval practice paradigms from our lab. These included data from ninety-four participants in Experiment 2 of Racsmány et al., 2010) and data from sixty participants from one condition of an unpublished experiment. These data were chosen because the basic procedure and materials used were similar to the one used in this study. To achieve a $\beta = .9$ to detect RIF we needed at least 26 participants for each of the four conditions (see the design section), therefore we settled on 30 (a priori $\beta = .93$). One hundred twenty participants were recruited at the Budapest University of Technology and Economics. One participant did not show up at the final test. Data was screened for outliers in the four experimental groups separately, for recall performance in the four within subject conditions (see the Design section). Outliers were defined as data points falling more than three standard deviations from the group mean. Two participants were identified as outliers and were excluded from further analyses.
Therefore, the results section shows the data for 117 participants (53 men and 64 women), aged between 19 and 26 years, \((M = 20.62, SD = 1.54)\).

**Design**

We varied delay (five minutes or five days) and practice type (retest or restudy) between subjects. Item type was varied within subjects. We used ten categories and six words from each category, a total of 60 category-word pairs. Members of two categories were used as filler items. The remaining 48 words from the remaining eight categories were assigned to one of the four item types. Four categories were selected randomly for each participant to be practiced categories. The four others were to be unpracticed categories. Words within each category were split randomly into two groups. One half of the words \((Rp^+)\) in each practiced category was to be practiced during the practice phase, the other half \((Rp^-)\) was not. Words in the unpracticed categories were used as baseline items. One half of the words \((Nrp^+)\) in each unpracticed category was randomly assigned to be the baseline for \(Rp^+\) words, the other half \((Nrp^-)\) served as baseline for \(Rp^-\) words.

**Procedure**

In the study phase, participants were presented all 60 words paired with their category label. Each pair was shown once for 5000 ms in the center of the screen with the category label on the left and the category member on the right. Participants were instructed to memorize the words with the help of the category label. Presentation of the pairs was pseudo-randomized with the constraint that two words belonging to the same category could not appear consecutively.

The practice phase consisted of three cycles, each containing a practice block with 18 trials followed by a reexposure block with 18 trials. Practice trials were different for the retest and restudy conditions, reexposure trials were the same. Reexposure trials served as a
feedback in the retest condition, and were introduced in the restudy condition as well to equal
the time on study in the two conditions.

In both the restudy and retest condition, practice and reexposure blocks each consisted
of 12 trials with Rp+ items and six trials with filler items. The use of filler items was
necessary to reduce primacy and recency effects. The first and the last two items in each block
were filler items. The order of the rest of the items was pseudorandomized with the constraint
that two consecutive trials never involved members of the same category. Practice trials in the
restest condition were cued recall trials. In each trial, the category label of the target word plus
a two-letter stem cue for the target word appeared in the middle of the screen. Participants
were instructed to complete the stem to the corresponding target. They had 6000 ms in the
first cycle and 4000 ms in the second and third cycle to type the answer. Practice trials in the
restudy condition were the same as trials in the study phase, except that restudy trials lasted
6000 ms in the first cycle and 4000 ms in the second and third cycle. Reexposure trials were
the same as trials in the study phase, except that reexposure trials lasted 1000 ms. Participants
were told that they would see some words again in a rapid sequence as a memory enhancer.

The three practice cycles followed each other in a repeated spaced retrieval schedule
in order to enhance the effect of testing (see Karpicke & Bauernschmidt, 2011). We
introduced one, three, and six minutes of delay filled with a two-back task, before the first,
second, and third practice cycle, respectively.

After the practice phase participants performed a five-minute long two-back task.
Then, participants in the short delay condition were introduced to the final test phase.
Participants in the long delay condition were mislead to believe that there was an error in the
script, and that their experiment was over. They had been previously told that they should
return to the lab five days later for a different test. This ‘cheating’ was necessary because
participants in the restudy condition had not been tested up until this point in the experiment,
and would have expected a test. This expectation would have been a variable that would have
taken different values in the retest and restudy condition, thus adding unwanted systematic
variation in final recall performance. For instance, participants in the restudy condition could
have prepared for a final test by practicing retrieval category-member pairs.

The final test consisted of two blocks. In order to avoid output interference (see
Anderson, 2003) Rp- items and their controls (Nrp- items) were tested in the first block,
followed by Rp+ items and their controls (Nrp+ items) in the second block. Items were
randomly intermixed within blocks (Camp, Pecher, & Schmidt, 2007). The use of different
control items for Rp+ and Rp- items was necessary to circumvent baseline deflation (see
Anderson, 2003). Both blocks started and ended with two filler items. Trials were the same as
in the first retrieval practice block except that the category-plus-word-stem cue contained only
a first-letter stem of the category member.

Results

Mean recall performances are shown in Table 1. We conducted a mixed design ANOVA on
recall data with item type (Rp+, Rp-, Nrp+, Nrp-) as a repeated measures variable, and
practice type (retest vs. restudy) and delay (five minutes vs. five days) as between subject
variables. Item type had a significant main effect on final recall, $F(3,339) = 122.08, p < .001$,
which was due to better recall of Rp+ ($M = .75, SD = .08$) items than any other item type
(Rp-: $M = .44, SD = .10$, Nrp+: $M = .43, SD = .09$, Nrp-: $M = .46, SD = .08$). Final recall was
better after a five-minute retention interval than after a five-day retention interval, $F(1,113) =
32.14, p < .001$, and retesting led to better overall recall than restudying, $F(1,113) = 5.12, p
< .05$. Importantly, item type interacted significantly with practice type, $F(3,339) = 10.38, p
< .001$, and delay $F(3,339) = 3.72, p < .05$.

(Table 1 about here)

Testing-effect

102
Final recall performance for practiced items (Rp+) is shown in Figure 1. To test the hypothesis that retesting Rp+ items is more beneficial for long-term retention than restudying them, we conducted a 2 (practice type: retest vs. restudy) X 2 (delay: five minutes vs. five days) factorial ANOVA on Rp+ recall rates. Both practice type, \( F(1,113) = 37.29, p < .001 \), and delay, \( F(1,113) = 16.86, p < .001 \), had a significant main effect, and we found a significant interaction of the two variables, \( F(1,113) = 8.85, p < .001 \). As can be seen in Figure 1, this interaction term is due to the difference in forgetting of practiced items over five days in the two conditions. Independent \( t \)-tests confirmed that there is a large and significant forgetting over five days (\( M = 20.46\%, SD = 4.82\% \)) in the restudy condition, \( t(59) = 4.24, p < .001, r = .48 \), but no forgetting (\( M = 3.27\%, SD = 2.95\% \)) in the retest condition, \( t(54) = 1.11, ns \). Interestingly, we found a medium-size effect of testing in the five-minute delay condition, \( t(57) = 2.64, p < .05, r = .33 \), indicating that retests were more beneficial for retention than restudy even after a short delay.

(Figure 1 about here)

**Retrieval-induced forgetting**

To detect a RIF effect, we performed paired-samples \( t \)-tests for each of the four between subject conditions, contrasting Rp- recall with Nrp- recall. The RIF effect was only significant in the retest condition after a five minute delay, \( t(28) = -3.13, p = .004, r = .51, \) and \( p > .16 \) in all other conditions. In brief, testing induced forgetting only in the short delay condition and only when participants were retested during the practice phase, and not when they restudied the same material.

**Discussion**

We investigated positive and negative effects of testing in one experiment. We demonstrated that retesting but not restudying target memories led to forgetting of related memories. This forgetting occurred only after a short delay, and not after a long delay. We also demonstrated
pronounced forgetting for restudied memories over a five-day delay, and almost no forgetting for retested memories.

The fact that participants tested after a short retention interval showed similar enhancement of memories in the retest and restudy conditions has been taken as evidence that memories undergo equal strengthening after retesting and restudying (e.g. Anderson, 2003; Hulbert, Shivde, & Anderson, 2012; J. G. W. Raaijmakers & Jakab, 2012).

However, our results, in line with the testing effect literature, show that despite the apparently equal strengthening¹, memory representations change differently when being tested as compared to when being restudied. Clearly, as shown by the delayed test in our study, retesting led to larger memory strength increments than restudying. It follows then, that the lack of (or the presence of) RIF after restudy may not provide evidence for (or against) the strength-independence of RIF as it was suggested by previous research (Anderson, 2003). Testing the strength-independence assumption of RIF would need novel methods that do not confound the effect of memory strengthening and memory impairment.

Our results lend support to the suggestion that common processes underlie both the testing effect and RIF. One major theory that can, and has already been used to account for both effects, is the new theory of disuse (R. A. Bjork & Bjork, 1992). This theory differentiates between two types of memory strengths, retrieval strength and storage strength, that can be described by ease of access and learnedness, respectively. The probability of successful retrieval is determined by current retrieval strength. Storage strength indirectly mediates the dynamics of retrieval strength fluctuation. Specifically, the theory predicts that retrieval strength decays over time due to learning of other memories, but forgetting can be slowed down by accumulating storage strength. Successful retrieval of memories with low retrieval strength leads to larger increments in storage strength than successful retrieval of

¹ Note that we found a medium-size effect benefit of retesting over restudying on the short run as well.
memories with high retrieval strength. That is, a desirably difficult learning event (R. A. Bjork, 1994; R. A. Bjork & Bjork, 1992; see also Pyc & Rawson, 2009) is more powerful than learning events involving fluent processing. A speculative interpretation of our results in this framework is that retesting is a more demanding learning event then restudying, therefore it leads to more pronounced changes in memory strength. Strengthening of memories would lead to forgetting only in instances when it is powerful enough to induce changes in the storage strength of a memory.

Both RIF and the testing effect have been explained by modification of reactivated memory patterns (for RIF see R. A. Bjork & Bjork, 1992; see also Racsmány & Conway, 2006; Racsmány et al., 2010; for the testing effect see Halamish & Bjork, 2011). These theoretical lines are also in agreement with the reconsolidation approach. In this framework, cue-dependent, retrieval induced reactivation of memories is believed to render cue-target sets labile, a state in which memory sets can be reconstructed (e.g. Hupbach, Gomez, Hardt, & Nadel, 2007; Sara, 2010).

If retests reshape or reconsolidate memory patterns then it is curious to see that the RIF pattern found in our experiment was not retained over a five-day delay. The delay-dependence of RIF is still debated (see e.g. Racsmány et al., 2010 vs. Baran, Wilson, & Spencer, 2010). Below we outline some speculative explanations of the absence of long-term RIF in our study.

According to the theory of disuse it is plausible that competitive retrieval from a given memory set during retrieval practice increases relative strength of Rp+ items and decreases relative retrieval strength of Rp- items. However absolute retrieval strength of the whole memory set (both Rp+ and Rp- items) should increase. In addition, since storage strength is an accumulating function of learning, there is no reason to suppose that storage strength of Rp- items decreases due to retrieval practice.
The theory of disuse offers another possible explanation to why we did not see long-term RIF. As Bjork and Bjork (1996) points it out, memories regress over time. After updating a memory set’s representation, it is the new representation that is most accessible, but with time there is a recovery (both in absolute and relative terms) of the earlier representation. In the retrieval practice paradigm the study phase is the earliest representation of practiced memory sets that creates equal recall probabilities for all Rp items (both Rp+ and Rp- items). This memory set is updated during retrieval practice through shifting relative strength in favor of Rp+ items. However, with time, the earlier representation of the memory set might recover due to disuse of both representations. This means that relative retrieval strength of Rp+ and Rp- items shifts toward equal recall probabilities. This might explain recovery from RIF. The fact that Rp+ items have larger absolute retrieval strength than Rp- items at the long term test is simply explained here by the differential increments in absolute retrieval strength, and storage strength due to practice.

The episodic inhibition framework that Racsmány et al. (2010) used to explain a RIF effect that persisted over 12 hours filled with sleep (but not 12 hours awake) states that once a pattern of activation/inhibition over a memory set is established, the recall pattern of elements of that set will pertain whenever the same episode in which the pattern was established is reinstated. However, Racsmány and colleagues (2010) underline that the pattern of activation/inhibition can or cannot be consolidated depending on what happens after encoding a representational pattern. Based on their results, they suggest that a 12-hour period filled with sleep that follows immediately the retrieval practice can lead to consolidation of a representational pattern, whereas a 12-hour period spent awake might hinder the same consolidation processes. Indeed, they found RIF when 12 hours of sleep followed the retrieval practice, but no RIF when their participants spent 12 hours awake after it. We suggest that in the study presented here, RIF on the long run might have also been absent due to the fact that
the period spent awake after initial learning was not controlled. Although this interpretation was supported by one further study (Abel & Bäuml, 2012), the factors that influence the delay dependence of RIF will need future investigations.

Strength is not a unitary characteristic of memories (R. A. Bjork & Bjork, 1992; Tulving & Pearlstone, 1966). Our results show that some effects of memory strength manipulations on remembering and forgetting might appear only after long delay intervals. The data reported here suggest that testing memories provides long-lasting positive mnemonic effects that are accompanied by the temporarily lower accessibility of other, related memories.
References


doi:10.3389/fnbeh.2010.00185


Table 1

*Recall Performance (In Percent Recalled) at Final Test for the Four Item Types in the Four Between Subject Conditions.*

<table>
<thead>
<tr>
<th>Practice type and delay</th>
<th>N</th>
<th>Rp+</th>
<th>Rp-</th>
<th>Nrp+</th>
<th>Nrp-</th>
</tr>
</thead>
<tbody>
<tr>
<td>Retest</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Five minutes</td>
<td>29</td>
<td>.85 (.12)</td>
<td>.44 (.19)</td>
<td>.49 (.23)</td>
<td>.57 (.13)</td>
</tr>
<tr>
<td>Five days</td>
<td>27</td>
<td>.82 (.09)</td>
<td>.40 (.18)</td>
<td>.37 (.17)</td>
<td>.40 (.14)</td>
</tr>
<tr>
<td>Restudy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Five minutes</td>
<td>30</td>
<td>.76 (.14)</td>
<td>.48 (.19)</td>
<td>.52 (.18)</td>
<td>.53 (.18)</td>
</tr>
<tr>
<td>Five days</td>
<td>31</td>
<td>.56 (.23)</td>
<td>.42 (.20)</td>
<td>.33 (.16)</td>
<td>.37 (.17)</td>
</tr>
</tbody>
</table>

*Note.* Rp+: Practiced words from practiced categories, Rp-: unpracticed words from practiced categories, Nrp+: words from unpracticed categories used as baseline for Rp+ words, Nrp-: words from unpracticed categories used as baseline for Rp- words. Standard deviations are shown in brackets. By mistake, 31 participants were assigned to the Restudy/Five days condition, and 29 in the Retest/Five days condition.
**Figure 1.** Recall Performance of Practiced Items (Rp+) at the Final Test in the Four Between Subject Conditions. Error bars represent 95% confidence intervals.
Study 4

Testing Promotes Long-Term Learning via Stabilizing Activation Patterns in a Large Network of Brain Areas

Attila Keresztes1, Daniel Kaiser2, Gyula Kovács1,2,3,4 and Mihály Racsmány1

1Department of Cognitive Science, Budapest University of Technology and Economics, Budapest, Hungary, 2Institute of Psychology, University of Regensburg, Regensburg, Germany, 3DFG Research Unit Person Perception, Friedrich-Schiller-University of Jena, Jena, Germany, and 4Institute of Psychology, Friedrich-Schiller-University of Jena, Jena, Germany

Address correspondence to M. Racsmány, Budapest University of Technology and Economics, Egyr József u.1., Budapest, Hungary. Email: racsmany@cogsci.bme.hu

Gyula Kovács and Mihály Racsmány contributed equally to this study.

The testing effect refers to the phenomenon that repeated retrieval of memories promotes better long-term retention than repeated study. To investigate the neural correlates of the testing effect, we used event-related functional magnetic resonance imaging methods while participants performed a cued recall task. Prior to the neuroimaging experiment, participants learned Swahili–German word pairs, then half of the word pairs were repeatedly studied, whereas the other half were repeatedly tested. For half of the participants, the neuroimaging experiment was performed immediately after the learning phase; a 1-week retention interval was inserted for the other half of the participants. We found that a large network of areas identified in a separate 2-back functional localizer scan were active during the final recall of the word pair associations. Importantly, the learning strategy (test or restudy) of the word pairs determined the manner in which the retention interval affected the activations within this network. Recall of previously restudied memories was accompanied by reduced activation within this network at long retention intervals, but no reduction was observed for previously retested memories. We suggest that retrieval promotes learning via stabilizing cue-related activation patterns in a network of areas usually associated with cognitive and attentional control functions.

Keywords: fMRI, forgetting, long-term learning, retrieval, testing effect

Introduction

Understanding the neural basis of how we lose access to previously encoded knowledge is a fundamental question of cognitive science as well as the psychology of learning and education. Since the seminal work of Ebbinghaus (1885/1964), the effect of the retention interval on forgetting has been one of the central topics of memory research. Several factors have been identified that can potentially explain aspects of the strong connection between retention interval and forgetting. Two such factors are the negative effect of acquiring new information after encoding the target event and the effect of sleep on memory consolidation (Roediger et al. 2010). Although some core processes of forgetting—such as the failure of memory consolidation and the consequences of interference resolution from competing irrelevant memories during retrieval—have already been identified (Uncapher and Wagner 2009; Wimber et al. 2009; Levy et al. 2010), our knowledge of the neural mechanisms of long-term forgetting is far from comprehensive. Hence, it is not surprising that some of the most remarkable experimental results regarding forgetting are those that demonstrated that even a single factor (an additional retrieval after memory encoding) can significantly reduce the negative influence of retention interval on recall performance (Spitzer 1939; Tulving 1967; Carrier and Pashler 1992; Roediger and Karpicke 2006a).

The finding that additional retrieval practice promotes better long-term retention and a slower forgetting rate than the simple restudy of the same information has been termed the “testing effect,” an effect that is currently attracting considerable attention (Roediger and Butler 2011). This phenomenon contradicts what is typically thought about successful learning and is also in conflict with general educational practice, in which testing is only the checkpoint of consecutive study phases (Roediger and Karpicke 2006b).

Furthermore, recent experiments have demonstrated that the rate of forgetting is influenced by learning strategy. Although retesting had no mnemonic advantage over restudying at short retention intervals, it produced significantly higher learning performance than an equal amount of restudying when the retention interval was longer than 1 day (Wheeler et al. 2003; Karpicke and Roediger 2008; Toppino and Cohen 2009). These results suggest that the efficiency of testing over restudying has a positive correlation with the length of retention interval. Although this interaction between learning strategy and retention interval seems to be an important aspect of human learning, the responsible functional neural networks have not yet been identified.

As a first step in seeking for the neural correlates of the testing effect, we investigated areas of the human brain that are known to be involved in cue-driven episodic retrieval (ER) processes. In previous experiments, ER was typically studied with associative cued recall and recognition tasks (Rugg and Henson 2002). These experiments demonstrated that successful memory retrievals are associated with activations in a large cortical network, including the prefrontal (PFC), posterior parietal (PPL), and medial temporal cortices, and hippocampus (Fletcher and Henson 2001; Rugg 2004; Spaniol et al. 2009; Kim 2011). Importantly, this retrieval-related network has a striking overlap with the network activated by working memory (WM) tasks (Cabeza et al. 2002). This result corresponds to WM theories that assume that WM activation is crucial for enhancing the efficiency of retrieval cues in guiding memory search (Bunting 2006; Unsworth and Engle 2006, 2007). Interestingly, 2 recent neuroimaging studies (Kuhl et al. 2007; Eriksson et al. 2011) demonstrated that when compared with a single retrieval, repeated retrieval practice leads to a reduced activation of a large portion of these regions, including the bilateral ventrolateral PFC, inferior frontal cortices (BA 9/44), the right DLPFC (BA 45/46), the left prefrontal (BA 39), and the bilateral superior parietal lobule (BA 7). These results were considered to be evidence that repeated testing reduces...
cognitive control demands during future ER by making the cue-target link easier to process (Kuhl et al. 2007). Furthermore, as Karpicke (2012) pointed out, each time a person retrieves a piece of information from memory, the future accessibility of this information improves because retrieving enhances the effectiveness of the specific retrieval cue in reconstituting all associated memories. According to the account of Karpicke and Smith (2012), this effectiveness is driven by a mechanism that by each retrieval act refines the search set and renders it smaller. This in turn may reduce the demand on WM to accomplish successful retrieval (Karpicke and Blunt 2011; Karpicke 2012). Altogether, these findings indicate the possible role of a network of areas related to WM in producing the long-term advantage of testing.

The aim of the current study was to investigate the role of cortical areas related to updating information in WM, attentional control, and controlled retrieval in the testing effect. We predicted that retrieval during the test phase promotes long-term memory advantages via efficient retrieval cue processing. Furthermore, we assumed that following repeated successful retrieval attempts, a given retrieval cue can efficiently activate WM and cognitive control-related networks even after long retention intervals. This would be beneficial for all future associative search processes, leading to the positive effect of retrieval (i.e., the testing effect). In contrast, without an initial retrieval attempt during learning, processing of retrieval cues may load heavily on control processes during tests following short retention intervals, and might not be effective following longer retention intervals. Thus, we compared the neural correlates of the associative recall of memories learned with 2 different learning strategies (retesting vs. restudying) after either a short or a long retention interval.

Materials and Methods

Participants
Twenty-nine healthy participants (2 left handed, 20 females, mean ± SD age: 22.93 ± 2.26 years) were recruited at the University of Regensburg. All participants were native German speakers and gave informed written consent to participate in the study, which was approved by the ethics committee of the University of Regensburg. None of the participants had any history of neurological diseases, and all had normal or corrected-to-normal visual acuity. We excluded 5 participants from the final analysis: for 1 person, fMRI data acquisition failed, and the other 2 participants did not follow instructions.

Stimuli and Design
The stimuli were 60 Swahili–German word pairs translated from the Swahili–English normalized data published by Nelson and Dunlosky (1994). We used word pairs with moderate recall probabilities according to the Nelson and Dunlosky (1994) normalized data. Thirty word pairs were randomly assigned to both the restest and the restudy conditions (see below).

Procedure
The full experiment was run in 2 parts. In the first part, participants completed an initial learning phase (learning Swahili–German word pairs). In the second part, participants were scanned in 3 sessions: First, they completed a final test for the material studied during the initial learning phase; second, they were asked to lay still and relax while a structural scan was performed; and third, they performed an n-back task. After these scanning sessions, the second part of the experiment ended with an off-scan test for all the material studied during the initial learning phase.

In the initial learning phase, participants learned the Swahili-German word pairs alone in a quiet room, with the aid of a computer screen (80 Hz, 1280 × 1024 resolution, viewing distance: 65 cm). First, participants were presented with all 60 Swahili–German word pairs subsequently. Each pair was presented randomly for 5000 ms in the center of the screen with the Swahili word on the left and its German meaning on the right. Participants were instructed to memorize all of the word pairs for the later test phase. They were verbally told that they should see the Swahili word during later testing and be asked to recall its German meaning. Next, participants learned the 60 word pairs through 6 learning cycles; each cycle included 1 restet, 1 restudy, and 1 feedback block. Unknown to the participants, half of the word pairs were assigned to the restest strategy condition and half to the restudy strategy condition. The restest–restudy words varied randomly across participants. In the restest blocks, all 30 word pairs assigned to the restest condition were tested once, in random order. During a trial, the Swahili member of the word pair appeared on the left side of the screen, and participants were instructed to recall and type in the German meaning in a box that appeared on the right side of the screen. Participants had 8000 ms to accomplish the task. In the restudy blocks, all 30 word pairs assigned to the restudy condition were presented randomly, each for 5000 ms, with the Swahili word on the left and its German meaning on the right. In each feedback block, all 60 word pairs were presented again, each for 1500 ms. These feedback blocks served to enhance the effect of testing (Rosediger and Butler 2011). In each learning cycle, the order of the restest and restudy blocks was random, and each cycle ended with a feedback block.

Next, half of the participants (n = 15) were assigned to the short retention interval group, while the other half (n = 14) to the long retention interval group. As noted above, 3 participants’ data were excluded from the analyses, leaving n = 13 in both groups. In the short retention interval group, the second part of the experiment (final test of the Swahili-German words in the fMRI scanner) was performed right after the learning phase (on average, there was a 20-min interval between the end of the learning phase and the beginning of the scanning). In the long retention interval group, this final test and the scanning were performed exactly 1 week after the learning phase. In order to avoid self-testing during the retention interval, all participants were told that the fMRI part of the experiment would examine social cognition and that it would be unrelated to the “memory experiment” they had just performed. In both cases, participants were informed about the security issues of the scanning procedure prior to the final test. In the scanner, stimuli were back-projected via an LCD video projector (JVC, DLA-G20, Yokohama, Japan, 72 Hz, 800 × 600 resolution) onto a translucent circular screen (diameter = 30°), placed inside the scanner bore 63 cm from the observer. Stimulus presentation was controlled via Presentation (Version 14.1 Build 09.21.09).

The final test phase consisted of cued recall trials (which were similar to the trials of the restest block during the learning phase) intermixed with fixation trials. Each of the 60 word pairs was tested once. In each trial, the Swahili word appeared in the middle of the screen, and participants were instructed to silently recall its German meaning. Participants were told to press a response button if they knew the answer, but to refrain from saying the word out loud. Each trial lasted 10 s, irrespective of whether the participant responded. Each cued recall trial was preceded by fixation trials (1000, 3000, or 5000 ms) that were used to jitter the cue onset during the test phase. The 3 types of fixation trials appeared equally often and were randomized in order. Participants were told to press the response button as quickly as possible because we were interested in observing how fast they could remember the word. To measure their correct recollection rate, we specifically instructed them that they should press the response button if they would be able to report the German word at a follow-up test immediately after scanning in the laboratory. Participants had a 30-s rest period after the 30th cued recall trial. During the follow-up test right after the scanning sessions, participants were asked to recall the remembered words. In all further analyses, we considered a word pair to be remembered only if the participant signaled during scanning that (s)he remembered it and if (s)he could report the answer correctly in the follow-up test. Incorrect trials (i.e., trials in which the participant had responded that they had known
the response, but could not report the correct target at the follow-up test) were dismissed from further analyses.

**Scanning Parameters and Data Acquisition**

Imaging was performed using a 3-Tesla MR head scanner (Siemens Allegra, Erlangen, Germany). For the functional series, we continuously acquired images (34 slices, 20° tilted relative to axial; $T_2^*$-weighted EPI sequence, TR = 2000 ms, TE = 30 ms; flip angle = 90°; 64 x 64 matrices; in-plane resolution: 3 x 3 mm; slice thickness: 5 mm, 10% gap). High-resolution sagittal $T_1$-weighted images were acquired using a magnetization-prepared rapid gradient-echo sequence (MP-RAGE; TR = 2250 ms; TE = 2.6 ms; 1 mm isotropic voxel size) to obtain a 3D structural scan. Details of preprocessing and statistical analysis are given elsewhere (Kovács et al. 2008, 2012; Cziraki et al. 2010). Briefly, the functional images were corrected for acquisition delay, realigned, normalized to the Montreal Neurological Institute (MNI)-152 space, resampled to 2 x 2 x 2 mm resolution and spatially smoothed with a Gaussian kernel of 8-mm full-width half-maximum (SPM8, Welcome Department of Imaging Neuroscience, London, UK).

Region of interest (ROI) analysis was based on the results of separate functional localizer runs which were 5 runs of the following 2 x 30 s blocks: a 30-s epoch of letters (700-ms exposition time + 300-ms blank for each letter) preceded by an instruction to “respond if the current letter is the same as the 1 presented 2 letters previously (2-back)”, followed by a 4-s blank period and another 30-s period of letters (700-ms exposition time + 300-ms blank for each letter) preceded by the instruction screen “respond if the current letter is a ‘D’ (detect a D)”. This functional localizer was similar to the one used for localizing the cortical network activated by a 2-back task in Drobyshevsky et al. (2006). The data were analyzed using the MARSBAR 0.42 toolbox for SPM (Brett et al. 2002).

The locations of ROI areas were determined individually as areas responding more strongly during the 2-back task than during the detection task in the functional localizer scans ($P_{uncorrected} < 10^{-5}$; $T = 4.86$, df = 273) (Figure 1). The coordinates of the areas are presented in Table 1. The ROIs were selected individually on the single subject level from the thresholded T-maps. Areas lying closest to the corresponding reference cluster (based on the results of the previous literature and the results of the random-effects analysis for differential contrasts; $P_{uncorrected} < 10^{-3}$; $T = 3.12$, df = 241) were considered as their appropriate equivalents at the single subject level. A time series of the mean voxel values within an 8-mm radius sphere around the local maximum of the areas of interest was calculated and extracted from our event-related sessions using finite impulse response models (Ollinger et al. 2001). The convolution of a reference hemodynamic response function (HRF) with boxcars (which represented the onsets and offsets of events) was convolved with the reference HRF, and the resulting impulse response function was deconvolved from the time course of the mean voxel value within a 20-mm radius sphere around the local maximum of the area of interest. A time series of the mean functional ROI signal was then submitted to a repeated-measures ANOVA with the factors BOLD effect (if significant) and time (pre-administration, post-administration, post-rest) as within-subject factors.

![Figure 1](http://cercor.oxfordjournals.org/) Activation maps for the functional localizer task (2-back vs. detection). Regions consistently activated across subjects are color-coded according to $P_{uncorrected} < 0.001$. The z-coordinate in Talairach space is indicated above each section. For anatomical details of the activations, see Table 1.
Neural Correlates of the Testing Effect  Keresztes et al.

durations of the experimental conditions) was used to define the regressors for a general linear model.

Data Analysis

We performed a 2-way mixed design ANOVA on final recall accuracy and final recall RTs with strategy (2, retest, restudy) as the within-subject factor and retention interval (2, short, long) as the between-subject factor. As for the BOLD signal, trials were analyzed and separately modeled at the onset of the stimuli (duration = 10 s). The peak of the event-related averages at 6–8 s poststimulus onset was used as an estimate of the response magnitude and averaged across repetitions for each condition and participant separately. We performed 3-way mixed design ANOVAs on the peaks with strategy (2, retests, restudy) and success (2, remembered, forgotten) as the within-subject factors and retention interval (2, short, long) as the between-subject factor.

Results

Behavior Results

Participants learned on average 75% of Swahili–German associations until the end of the initial learning phase. Recall success (in percentages) for retest items increased from cycle 1 to cycle 6, \(M = 0.15, SE = 0.02\) in cycle 1, \(M = 0.29, SE = 0.03\) in cycle 2, \(M = 0.42, SE = 0.04\) in cycle 3, \(M = 0.57, SE = 0.04\) in cycle 4, \(M = 0.67, SE = 0.04\) in cycle 5, and \(M = 0.75, SE = 0.04\) in cycle 6. Performance in the short versus long retention interval groups did not differ in any of the learning cycles (all \(P > 0.33\)).

The upper panel of Figure 2 presents the performances of the participants at the final test, expressed as the proportion of correctly recalled words for the retest and restudy strategy conditions and for the short and long retention interval groups, separately. As expected, retention interval had a significant main effect on the final recall accuracy \((F_{1,24} = 14.26, P < 0.001)\): participants’ overall recall accuracies were lower after a 1-week retention interval \((M = 44.74\%, SE = 4.56\%)\) than after a 20-min retention interval \((M = 69.1\%, SE = 4.56\%)\). Although strategy had no main effect on recall accuracy, we observed a significant interaction between strategy and retention interval \((F_{1,24} = 5.80, P = 0.024)\). Post hoc tests demonstrated that this result arose because the recall accuracies of the retest condition were significantly higher \((M = 50.26\%, SE = 6.93\%)\) than those of the restudy condition \((M = 39.23\%, SE = 3.25\%)\) in the long retention interval condition \((t_{122} = 2.33, P = 0.038)\). However, there were no differences in the short retention interval condition, \((t_{122} = 0.92, ns, M = 67.44\%, SE = 4.55\%) and \(M = 70.77\%, SE = 4.60\%\) for retest and restudy, respectively). This result confirms previous findings in which repeated retrieval led to better long-term retention than additional study, even though the 2 conditions produce similar performances on short intervals \((Roediger and Karpicke 2006b)\).

Analysis of RTs (Fig. 2, lower panel) revealed a significant main effect of strategy \((F_{1,24} = 8.93, P = 0.006)\), that was due to shorter recall RTs overall in the retest condition \((M = 2411\, ms, SE = 148\, ms)\) compared with the restudy condition \((M = 2859\, ms, SE = 149\, ms)\). Retention interval also had a main effect with shorter RTs in the short retention interval group \((M = 2249\, ms, SE = 148\, ms)\) than the long retention interval group \((M = 3021\, ms, SE = 182\, ms)\). In contrast to the ANOVA on final recall accuracy, the ANOVA on RTs did not reveal any significant interaction between strategy and retention interval.

fMRI Results

Interaction of Learning Strategy and Retention Interval

The main aim of the current study was to determine whether there are cortical areas which show activation patterns that reflect the interaction of learning strategy and retention interval of the task, similarly to previous behavioral results \((Roediger and Karpicke 2006a; Karpicke and Roediger 2008)\). To this end, we performed a 3-way mixed design ANOVA on the extracted BOLD signals. We reasoned that if an area is related to the superior performance observed after repeated retrieval and long retention periods, then the activity of that area should show a significant interaction of learning strategy and retention interval. Table 2 presents main effects and interactions for each area separately. A number of ROIs demonstrated this type of interaction. Figure 3 presents the average (±SE) BOLD signal as a function of time for 4 representative areas as well as the extracted peak activations for all areas with significant interactions. As can be observed in the HRFs, the basis of the interaction between learning strategy and retention interval was that activations in the restudy condition were higher when compared with those in the retest condition after short retention interval, but the opposite effect was observed after long retention interval: retest activations exceeded those of the restudy condition. Post hoc t-tests (see Supplementary Table 1) showed that from short to long retention interval, activation did not decrease significantly for retested items in any of the ROIs \((all\, P > 0.33)\), and only 1 region, the right thalamus showed a significant increase \((P < 0.026,\; all\, other\, P > 0.18)\). In contrast, for restudied items, all areas showed a nominal decrease of activation that was significant in several areas, including the left insula, the anterior cingulate bilaterally, the left...
A 2 × 2 × 2 ANOVA was performed on percent signal changes in regions of interest (see Table 2). Retrieval success (remembered vs. forgotten) and learning strategy (retest vs. restudy) were varied within participants, retention interval (20 min vs. 1 week) was varied between participants.

### Table 2
Significant main effects and interactions affecting percent signal change in regions of interest

<table>
<thead>
<tr>
<th>Effect</th>
<th>Area</th>
<th>F</th>
<th>df</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Main effect of success</td>
<td>Inferior parietal, right (~BA40)</td>
<td>18.82</td>
<td>1.22</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>Thalamus, right</td>
<td>34.41</td>
<td>1.10</td>
<td>0.0001</td>
</tr>
<tr>
<td></td>
<td>DL-PFC, right (~BA9/45)</td>
<td>10.53</td>
<td>1.22</td>
<td>0.004</td>
</tr>
<tr>
<td></td>
<td>Posterior/dorsal prefrontal, right (~BA6)</td>
<td>9.52</td>
<td>1.22</td>
<td>0.005</td>
</tr>
<tr>
<td></td>
<td>Inferior parietal, left (~BA7)</td>
<td>8.05</td>
<td>1.23</td>
<td>0.009</td>
</tr>
<tr>
<td></td>
<td>Anterior cingulate, left (~BA32)</td>
<td>8.57</td>
<td>1.20</td>
<td>0.008</td>
</tr>
<tr>
<td></td>
<td>Thalamus, left</td>
<td>10.21</td>
<td>1.10</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>Midbrain</td>
<td>7.85</td>
<td>1.8</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td>Anterior prefrontal, right (~BA10)</td>
<td>5.55</td>
<td>1.21</td>
<td>0.028</td>
</tr>
<tr>
<td></td>
<td>Fusiform, left</td>
<td>5.5</td>
<td>1.16</td>
<td>0.032</td>
</tr>
<tr>
<td></td>
<td>Insula, left</td>
<td>4.69</td>
<td>1.24</td>
<td>0.041</td>
</tr>
<tr>
<td></td>
<td>Superior parietal, left (~BA7)</td>
<td>3.97</td>
<td>1.24</td>
<td>0.058</td>
</tr>
<tr>
<td></td>
<td>Superior parietal, right (~BA7)</td>
<td>4.23</td>
<td>1.23</td>
<td>0.051</td>
</tr>
<tr>
<td>Main effect of retention interval</td>
<td>Middle orbitofrontal, right (~BA11)</td>
<td>6.3</td>
<td>1.22</td>
<td>0.02</td>
</tr>
<tr>
<td>Interaction: success × strategy</td>
<td>DL-PFC, right (~BA9/45)</td>
<td>5.24</td>
<td>1.22</td>
<td>0.032</td>
</tr>
<tr>
<td></td>
<td>Anterior prefrontal, right (~BA10)</td>
<td>5.28</td>
<td>1.21</td>
<td>0.032</td>
</tr>
<tr>
<td>Interaction: strategy × retention interval</td>
<td>Insula, left (~BA7)</td>
<td>19.86</td>
<td>1.24</td>
<td>0.0001</td>
</tr>
<tr>
<td></td>
<td>Fusiform, left</td>
<td>13.65</td>
<td>1.16</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td>DL-PFC, right (~BA9/45)</td>
<td>9.39</td>
<td>1.22</td>
<td>0.006</td>
</tr>
<tr>
<td></td>
<td>Anterior cingulate, right (~BA9/32)</td>
<td>7.34</td>
<td>1.18</td>
<td>0.014</td>
</tr>
<tr>
<td></td>
<td>Thalamus, right</td>
<td>7.63</td>
<td>1.10</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td>Insula, right</td>
<td>6.55</td>
<td>1.22</td>
<td>0.022</td>
</tr>
<tr>
<td></td>
<td>Inferior parietal, left (~BA7)</td>
<td>5.62</td>
<td>1.23</td>
<td>0.026</td>
</tr>
<tr>
<td></td>
<td>Superior parietal, right (~BA7)</td>
<td>5.14</td>
<td>1.23</td>
<td>0.033</td>
</tr>
<tr>
<td></td>
<td>DL-PFC, left (~BA9/45)</td>
<td>4.85</td>
<td>1.23</td>
<td>0.038</td>
</tr>
<tr>
<td></td>
<td>Inferior parietal, right (~BA40)</td>
<td>4.32</td>
<td>1.22</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td>Middle orbitofrontal, right (~BA11)</td>
<td>4.21</td>
<td>1.22</td>
<td>0.052</td>
</tr>
<tr>
<td></td>
<td>Posterior/dorsal prefrontal, right (~BA6)</td>
<td>4.15</td>
<td>1.22</td>
<td>0.055</td>
</tr>
<tr>
<td>Interaction: success × strategy × retention interval</td>
<td>Insula, left</td>
<td>9.14</td>
<td>1.24</td>
<td>0.006</td>
</tr>
<tr>
<td></td>
<td>Anterior cingulate, left (~BA32)</td>
<td>9.08</td>
<td>1.20</td>
<td>0.007</td>
</tr>
<tr>
<td></td>
<td>Thalamus, right</td>
<td>9.84</td>
<td>1.10</td>
<td>0.011</td>
</tr>
<tr>
<td></td>
<td>Fusiform, left</td>
<td>7.77</td>
<td>1.16</td>
<td>0.013</td>
</tr>
<tr>
<td></td>
<td>Fusiform, right</td>
<td>7.29</td>
<td>1.12</td>
<td>0.019</td>
</tr>
<tr>
<td></td>
<td>Inferior parietal, right (~BA40)</td>
<td>5.35</td>
<td>1.22</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td>Anterior prefrontal, right (~BA10)</td>
<td>4.73</td>
<td>1.21</td>
<td>0.041</td>
</tr>
<tr>
<td></td>
<td>Anterior prefrontal, left (~BA10)</td>
<td>4.27</td>
<td>1.23</td>
<td>0.05</td>
</tr>
</tbody>
</table>

Note: A 2 × 2 × 2 ANOVA was performed on percent signal changes in regions of interest (see Table 1). Retrieval success (remembered vs. forgotten) and learning strategy (retest vs. restudy) were varied within participants, retention interval (20 min vs. 1 week) was varied between participants.

This finding suggests that, when compared with repeated study, repeated retrieval leads to higher activations in a network of areas activated during a WM task after long retention intervals, which, in turn, leads to superior memory performance. Thus, the activity of these areas could serve as the functional basis of the behaviorally observed testing effect.

### Interaction of Learning Strategy, Retention Interval, and Retrieval Success

Interestingly, a subset of the areas (anterior PFC (BA10) bilaterally, the left insula, the left ACC, the right inferior parietal area (BA40), the right thalamus, and fusiform gyrus, bilaterally), the activities of which were modulated by strategy and retention interval also showed modulation according to the success of retrieval. This modulation was manifested in the significant 3-way interaction between strategy, retention interval, and success. Figure 4A shows examples of the HRF for 2 such areas, the left insula (upper panel) and the left ACC (lower panel), while Figure 4B shows the extracted peak activations for all other areas with significant 3-way interactions. At short retention intervals, previously restudied items elicit larger activations than previously retested items, although this appears to be driven by the nearly complete absence of BOLD signal change for the previously restudied and forgotten items.

### Areas Related to Retrieval Success

The main effect of retrieval success was significant in right frontal areas (right DL-PFC, right posterior/dorsal PFC and the right anterior PFC), left ACC, left insula, inferior parietal ROIs bilaterally, thalamus bilaterally, a midbrain area, and the left fusiform gyrus. In addition, we found a tendency for a main effect of success in superior parietal ROIs bilaterally ($P = 0.06$ and $P = 0.05$ for the left and right hemispheres, respectively). Because retrieval success interacted with either strategy or strategy and retention interval in several areas, we ran post hoc paired samples t-tests (see Supplementary Table 2) separately for the retest and restudy items in both the short and long retention interval conditions. This analysis revealed that only a few areas showed an effect of success at short retention interval: the left fusiform gyrus for retested items only, the right thalamus for both type of items, and the midbrain for restudied items only. At long retention intervals, however, the effect of success was significant in all but 2 of the above ROIs for restudied items (no significant effect was found in the left fusiform gyrus, and the midbrain ROI). In contrast, for restudied items, only the right thalamus showed a significant effect. Briefly, most ROIs were activated differently during successful versus unsuccessful retrieval attempts of the restudied items and mainly at long retention intervals. This effect contributed to the main effect of retrieval success.
To test whether any additional areas showed differential activation for the retest and restudy strategies, we performed a whole-brain analysis as well. This analysis revealed no significant activations at the $P_{FWE} < 0.05$ level at an extended threshold of 50 voxels in the short or long retention interval groups, neither for the retest > restudy nor for the restudy > retest contrasts. Similarly, we observed no significant activations for the interaction of retention interval and strategy. To further explore our data, we ran the same analyses at the more liberal $P_{uncorrected} < 0.0001$ level (with an extended threshold of 50 voxels) as well. At the short retention interval, again, no significant activations were found for either the retest > restudy or the restudy > retest contrasts. In contrast, as shown in Figure 5A, at the long retention interval, the retest versus restudy contrast revealed significant activations in a medial frontal/anterior cingulate area (8, 38, 10) and in an area in the occipital lobe at around the early visual cortex (2, −92, 2). Importantly, the interaction of retention interval and strategy, as shown in Figure 5B revealed 2 clusters of activations bilaterally over the inferior frontal gyrus (30, 28, −2; $k = 61$; $Z = 4.89$ and −32, 26, −2; $k = 124$; $Z = 4.23$), corresponding to the bilateral insular cortices in our ROI analyses. The restudy versus retest contrast did not reveal any significant activation at the long retention interval either.
As the comparison of data presented in Table 2 and in Figure 5 indicates, both the voxelwise and the ROI-based approach provide evidence showing that activations in the insular and the cingulate cortices are modulated by the interaction of strategy and retention interval. In addition, the results of the whole-brain analysis revealed only one additional area, the early visual cortex, which has been previously suggested to play a role in both WM and episodic memory retrieval-related tasks (Cabeza et al. 2003; Kim 2011).

Finally, to check the specificity of the results to the previously described areas we applied the same ROI analyses as above to additional areas, using ROIs defined by the complementary contrast of the functional localizer scan (detection > 2-back). This contrast, showing areas that are more active during the cognitively less loaded task, activated a network of areas, very similar to the recently described default network (Shulman et al. 1997; Gusnard and Raichle 2001; Mazoyer et al. 2001; Buckner et al. 2008) and included the medial posterior cingulate (BA 30/31, x: −3, y: −51, z: 28), the orbito-frontal (BA 10/11; −1, 55, −9) and the superior frontal gyrus (BA 9/10; −5, 62, 14). The ROI analyses of these 3 areas, did not show any significant main effect of strategy or delay, nor any interactions (all Ps > 0.15), supporting further the specificity of our results to areas related to cognitive and attentional control functions.
Discussion

The major findings of our study are the following. 1) Parietal and frontal areas, as well as the thalamus, the left fusiform gyrus, and a midbrain area were activated when participants had to recall previously learned memory items. The same areas were also activated during active updating and manipulating of information in WM during a 2-back task. 2) In most ROIs identified by the functional localizer 2-back task, the learning strategy of participants determined how the retention interval affected activations during the final test: repeated study and repeated retrieval of the learning material led to different BOLD signals during final recall after short and long retention intervals. In addition, the effect of learning strategy was different for participants who had to retain the memories for a few minutes versus for a week. 3) For several ROIs identified by the functional localizer 2-back task, the interaction of learning strategy and retention interval was also influenced by retrieval success. Our results show, for the first time, that the long-term behavioral advantage of repeated retrieval over repeated study is due to the differential activation of a large network involving parietal, frontal, and insular cortical areas, as well as the thalamus and the fusiform gyrus.

Memory Retrieval Activates a Network of Areas Activated During Updating and Manipulating Information in Working Memory

The anterior and dorsolateral part of the PFC, the superior and inferior parietal cortex, the anterior cingulum, the thalami bilaterally, an area in the midbrain, the left fusiform gyrus, and the left insula were activated both during the 2-back localizer task and episodic recall of words. This result supports earlier findings of Cabeza et al. (2002) and Ranganath et al. (2003) who showed that these regions, together with the cerebellum, were activated in both ER and WM tasks. The WM task used in our study involves online monitoring, updating, and manipulation of remembered information (Owen et al. 2005), and is therefore assumed to place great demands on a number of key processes within WM. Our findings suggest that participants may have leaned on these cortical areas to effectively process retrieval cues during associative recall. Indeed, theories of ER suggest that WM is necessary for several steps of the recall process, such as the initiation of a search process for a specific target memory or the monitoring of the accessed responses (Fletcher et al. 1998; Henson et al. 1999; Cabeza et al. 2002; Ranganath et al. 2003). Determining whether the currently found activations of areas identified by a 2-back task during the cued recall task are due to any of these steps was beyond the scope of the current study (designed to evaluate the possible effects of repeated retrieval vs. that of repeated study) and requires further neuroimaging studies.

Neural Correlates of Testing Effect: Learning Strategy Affects Long-Term Stability of Activations During Recall in a Network of Areas Activated During Updating and Manipulating Information in Working Memory

Second and more importantly, our behavioral results confirm the existence of testing effect in an fMRI scanner; a long retention interval produced a lower memory performance for previously restudied items compared with the performance on previously retested items. In addition, the analysis of RTs during final recall revealed that repeated retrieval of memories generally increased the effectiveness of retrieval cues; participants could recall the items faster in the retest condition than items in the restudy condition, irrespective of the length of the retention interval.

Furthermore, the imaging data obtained during final cued recall suggests that repeated retrieval of memories might contribute to the long-term stability of memory traces via the activation of retrieval-related areas whereas repeated study does not modulate these activations. In other words, during ER the activation of a network activated by a WM task is largely influenced by the learning strategy of the participants, which is a possible neural correlate of the testing effect. At short retention intervals, there is a significant activation of this network, irrespective of the learning strategy. At long retention intervals, this activation is more pronounced for memories that have been encoded through repeated retrieval compared with memories encoded through repeated study.

The Effect of Learning Strategy Depends on Retrieval Success

Our results indicate that at short retention intervals, retrieval cues activate areas in a network also activated by a WM task,
irrespective of retrieval strategy, and more importantly, irrespective of retrieval success. In other words, the BOLD activation, associated with successfully recalled and forgotten words, were similar for both restested and restudied items. Similarly, after a week-long retention interval, these areas were activated for the previously restudied memories, irrespective of recall success. However, for the previously restudied items, activation at final recall after a week-long retention interval depended largely on recall success, with virtually no BOLD signal change during retrieval attempts of restudied but forgotten items. This result suggests that at short retention intervals, cues related to the restudied memories activate areas of this network (and to a larger extent than cues related to restudied memories). At long retention intervals, however, lower activation of the same areas suggests that the cue processing is not initiated in many trials, which might lead to lower recall accuracy for previously restudied items compared with previously restested items, that is, the emergence of the testing effect.

Our results show that when a target memory of a cue-target association has been repeatedly retrieved during learning, cue processing will activate an overlapping network related to ER and WM tasks, even after a long retention interval. In contrast, for target memories that have been repeatedly studied, the cues might only activate these overlapping networks when the retention interval is short. Our neuroimaging results suggest that some of the restudied memories cannot be recalled after a week-long retention interval, most likely because of the failures of retrieval-related cue processing.

In interpreting our findings, 2 relevant neuroimaging studies should be mentioned. Eriksson et al. (2011), investigating the effect of repeated successful retrieval on changes in brain activity, found that the more times an item had been successfully retrieved during a prescan learning phase the higher the activity level in the ACC and the lower the activity level in the superior parietal and midventrolateral cortex was during a final retrieval phase. According to Eriksson et al.’s (2011) interpretation, decreased activation in the fronto-parietal network reflected reduced demands on cognitive control mechanisms necessary for successful retrieval. In a more recent study, Wiklund-Hörnquist et al. (2012) showed that repeated and successful retrieval during scanning was paralleled by decreases in the activity level of brain areas in orbitofrontal, insular as well as medial frontal regions, and the ACC (BA 47, 45, 6, 32, respectively). These results are in line with our present finding showing that, in the short retention interval condition, activity level of fronto-parietal networks was lower following repeated retrieval than following repeated study cycles.

Presently, there is no widely accepted theoretical account of the testing effect. We discuss 2 possible theories that have been raised in recent discussions. According to the elaborative encoding hypothesis (Carpenter 2009, 2011), attempts to reconstruct target memories during repeated retrieval produce extra information related to the cues which might mediate retrieval during later tests (Pyc and Rawson 2010). At long retention intervals, when target memories become harder to be reconstructed from single cues, it is the use of extra cues that would produce the long-term advantage of repeated retrieval over repeated study. In contrast, the search set constraining theoretical framework (Karpicke and Zaromb 2010; Karpicke 2012; Karpicke and Smith 2012) suggests that retrieval prompts a process, probably through effective temporal context reinstatement, which narrows the cue-related search set, and even a single retrieval can decrease the number of potentially retrievable items in response to a specific retrieval cue (Karpicke and Zaromb 2010; Karpicke and Blunt 2011; Karpicke 2012). In this account, retrieval is a discrimination process, where the effectiveness of a given cue will be determined by its ability to specify a given memory fragment in the context of many similar and interfering memory features.

The aim of the present study was not to contrast experimentally these 2 theoretical frameworks. However, the observed interaction between learning strategy and retention interval (with activations in areas activated during a WM task being higher in the restudy than in the restest condition after short retention interval, and lower after long retention interval) in our study, and results of earlier studies showing that each retrieval act leads to a decrease in fronto-parietal activations that is correlated with memory efficiency (Kuhl et al. 2007; Eriksson et al. 2011; Wiklund-Hörnquist et al. 2012) provide indirect support to the search set constraining framework.

In addition, the fact that restested memories were recalled with shorter RTs than restudied memories during final recall at both short and long retention intervals, also suggests that repeatedly retrieving memories increased the effectiveness of retrieval cues. One possible interpretation of the fMRI results, together with the pattern of RT findings is the following. At short retention intervals, repeated retrieval of associative memories leads to reduced demands on WM compared with restudying the same memories. This may be due to the fact that the search set and potentially activated features are significantly constrained during repeated testing cycles. According to this idea, a network of areas also related to WM, and cognitive and attentional control in general (Yarkoni et al. 2011), is responsible for calibrating the processing of cues to search long-term memories and delimit the search set to the target items. This result suggests that at short retention intervals, cues related to the restudied memories activate areas of this network (and to a larger extent than cues related to restested memories), as a direct consequence of the extended search set and larger amount of activated semantic elements following repeated study. At long retention intervals, however, lower activation of the same areas suggests that the cue processing is not initiated in many trials, which might lead to lower recall accuracy for previously restudied items compared with previously restested items, that is, the emergence of the testing effect. In sum, we suggest that the average RT advantage of the restest condition is the consequence of a smaller search set at short retention intervals, while it is due to the effective and more successful target reconstruction following long retention interval. This interpretation is supported by the fact that the RT advantage was accompanied by higher recall performance only following long retention interval.

In sum, these findings suggest that the retention interval of the first retrieval of a target memory, after learning, will determine the activation of overlapping areas in networks activated in ER and WM tasks. The first retrieval attempt of a cue-target association may trigger cue processing only when the retention interval between initial learning and retrieval is short. In contrast, when the retention interval is long, participants cannot effectively process the cue and a large percentage of retrieval attempts fail. Thus, the testing effect may be a consequence of...
processes that, through each additional retrieval act, conserve the effectiveness of the retrieval cue to access a specific memory. Based on our findings, we suggest that this strengthening arises from an effective and stable response for specific episodic cues in a network of brain areas related to cognitive control functions.

**Supplementary Material**

Supplementary material can be found at: http://www.cercor.oxfordjournals.org/.

**Funding**

This work was supported by the Deutsche Forschungsgemeinschaft (KO 3918/1-1, 1-2, 2-1 to G.K.), by OTKA (Hungarian National Research Fund) (K84019 to M.R.), and by the New Hungary Development Plan (Project ID: TAMOP-4.2.1-B-09/1/KMR-2010-0002).

**Notes**

We thank the help of Krisztina Nagy and Iulia Lavric during the collection of behavioral data. Special thanks to Mártá Zimmer for her help in preparing the figures. Conflict of Interest: None declared.

**References**


Discussion

This dissertation aimed at understanding mechanisms modifying the effectiveness of cued recall processes. In particular, it investigated mechanisms behind the negative effect of stopping retrieval, and the negative and positive effects of cued recall processes on later recall of retrieved memories and of non-retrieved memories related to the retrieved ones.

Study 1 demonstrated that cued recall can be effectively impaired when, in the presence of the cue, people try not to think about the target (no-think condition), but also when people try to come up with substitute targets (substitute condition). Importantly, target-cued recall of cues (recalling cues to targets) was facilitated in the no-think condition but impaired in the substitute condition. This suggests that forgetting in the TNT task can occur due to both not thinking about a target in the presence of a cue, and creating a substitute target (Hertel & Calcaterra, 2005). The fact that not thinking about a target leads to inhibition of cue→target association together with a facilitation of target→cue association clearly speaks against a cue-independent view in which inhibition is posited to act at the item level (Anderson, 2003). Moreover, the differential effects of the two instructions on the target-cued recall of cues suggests that both ways of inducing forgetting leads to changes at the level of unidirectional cue→target and target→cue associations, which might be independent components of episodic memories created in the study phase of these experiments (Conway, 2009).

Using RTs as an estimate of interference during competitive recall, Study 2 found that only moderate levels of interference led to RIF, and RIF was not observed when interference was too low, or too high. Although failing to provide an exact function relating interference to forgetting, this study clearly demonstrated that forgetting in the retrieval practice paradigm was a non-monotonic function of the level of interference during retrieval. This supports earlier theoretical (Anderson, 2003) and computational (Norman et al., 2007) models of RIF
that suggested that RIF is interference dependent. Besides its theoretical implications, this study described an item-by-item retrieval practice paradigm that can provide a test of properties of RIF that is more sensitive in detecting a RIF effect than the current paradigms used. This is important because a key issue in understanding the mechanisms of RIF, whether cue-independent forgetting exists or not, hinges on contrasting a few positive findings of cue-independent forgetting against many null effects. However, these null effects can be easily tackled by arguments that the effect itself is weak and hard to detect (Huddleston & Anderson, 2012), therefore devising more sensitive tests of RIF is one way to further the understanding of this controversy.

Study 3 replicated earlier findings that both repeated study and repeated retrieval of targets can lead to enhancement of recall of these targets, but of the two only repeated retrieval leads to RIF. Importantly, contrary to what has been posited by earlier studies of RIF, but in line with the findings of experiments on the testing effect, this study also showed that the enhancement induced by repeated study and repeated retrieval is far from being equal, the latter leading to superior performance compared to the former. The difference between the effectiveness of the two strategies was striking when a retention interval was five days long. This finding indirectly questions the assumption that RIF is strength independent, by showing that experiments providing evidence for strength independence using manipulations of learning strategy were based on a wrong auxiliary hypothesis. In particular, these experiments showed that repeated study and repeated retrieval lead to similar enhancement (Anderson & Levy, 2011; Raaijmakers & Jakab, 2013) but only the latter leads to RIF. However, study 3 showed that although the enhancement seems similar after a short retention interval, there is a striking difference as shown by tests administered after a longer retention interval.

In study 4, we found that brain areas that were differentially activated by an n-back task in comparison with a simple detection task were more active during cued recall of previously
restudied targets than during recall of previously retested targets after a short retention interval. In contrast, at a one-week retention interval, these same areas were not activated by cued recall of restudied targets, whereas retested targets elicited similar activations as on the short-term test. The results of study 4 showed that the advantage of repeated retrieval over repeated study on cued recall tests after long retention intervals is coupled with a temporally stable activations in a brain network related to cognitive control.

Recently, two further studies investigating the neural correlates of repeated retrieval have been published (van den Broek, Takashima, Segers, Fernández, & Verhoeven, 2013; Wing, Marsh, & Cabeza, 2013). The former found evidence for the involvement of the middle temporal gyrus and the left inferior parietal lobe in the positive effect of testing, and linked their results to the putative role of these areas in the elaboration of semantic associations (Binder, Desai, Graves, & Conant, 2009). They also found increased activity in the inferior frontal gyrus when retesting was contrasted to restudying, suggesting that larger cognitive effort was involved in testing than studying. Interestingly, successful retesting (as compared to unsuccessful retesting, i.e. retested and remembered at final test vs. retested but forgotten at final test) was also associated with ventral striatal and midbrain activations, which the authors interpreted as indicative of the involvement of reward and motivation circuits in the positive effects of retesting. Although this study adds important information to the understanding of the testing effect, it failed to directly contrast processes involved in retesting and restudying. Rather, participants were scanned during retesting and restudying and DM measures were calculated for both processes separately. The same criticism is even more relevant in the case of the Wing et al. study in which retest trials were directly contrasted to restudy trials, yielding, not surprisingly, activations that are generally observed in correlation with episodic retrieval (hippocampal, prefrontal and parietal). Our study, in contrast, had participants either restudy or retest cue-target pairs, and then scanned participants at a final test where the recall
procedure was identical for both retested and restudied targets. Therefore our study made direct comparison of retesting and restudying possible.

In sum, the four studies of this dissertation demonstrated that cue-target representations can be modified by stopping retrieval and active effortful retrieval, either competitive or non-competitive. In all these cases unidirectional cue-target associative links, that are part of episodic memories, are modified.

Reiterating the key issues along which interference and inhibitory accounts differ from one another, the results of the dissertation favor the interference accounts over the inhibitory accounts. 1) Changes in the accessibility of memories was explained better by changes occurring at the level of associations than by changes occurring at the level of item memory representations. 2) Although not directly tested in the dissertation this suggests that changes are not cue-independent. 3) Recall success was determined by cue driven retrieval processes in a dynamic way, rather than by static representational states of memories.

However, two results of the dissertation are not easily accommodated by interference theories. First, that RIF was interference dependent, and second, that it was shown to be retrieval specific. The results of Study 2, showing interference dependence, suggest that cognitive control might play a role in changing cue-target associations. In fact, cognitive control might have played a role in at least two other studies of the dissertation. In Study 1, the role of cognitive control is simply suggested by the fact that the task required conscious effort to suppress or substitute targets in cue-target associations. In Study 4, cognitive control mechanisms also played a key role in the positive mnemonic modulation of cue-target associations by repeated retrievals. The role of cognitive control played in these effects would support inhibitory theories over interference theories. However, as stated earlier, several results of our studies are inconsistent with the inhibitory account. Together, these results suggest that cognitive control might mediate changes that affect cue-target associations.
In addition, the results of Study 3 and 4 suggest that retrieval is a process more powerful in changing cue-target associations than other, less effortful, types of reencounters with memories. On the one hand, the positive mnemonic effect of retrieval is not surprising given the vast amount of data on the testing effect (Carpenter, 2012; Karpicke, 2012; Roediger & A. C. Butler, 2011; Roediger & Karpicke, 2006). On the other hand, the negative mnemonic effect of retrieving cue-target associations, compared to other types of reencounters, does not fit most interference accounts. Importantly again, this does not necessarily provide support for inhibitory control theory.

There are two current theories that might accommodate retrieval specificity, without reference to any inhibitory mechanism that acts on item level representations, independently of cues. One is the latest upgrade of the SAM retrieval model (Raaijmakers & Shiffrin, 1981), the SAM-REM theory (Malmberg & Shiffrin, 2005). Besides incorporating time factors, as the context fluctuation model (Mensink & Raaijmakers, 1988) also derived from the SAM model, this model makes one additional important assumption critical for explaining some learning and forgetting phenomena, including retrieval specificity. The SAM-REM model assumes that retrieval leads to a larger strengthening of context-target associations than restudy, and this can lead to no RIF in instances involving no retrieval prior to final test. According to the differential encoding hypothesis postulated by Verde (2013), this model assumes that different practice types (e.g. retrieving versus restudying) have differential effects on later sampling and recovery of target memories. This hypothesis has been successfully tested in three behavioral experiments by Verde in the same paper. A framework closely related to the SAM-REM model is the episodic inhibition proposed by Racsmány and Conway (2006). This framework assumes that cue-target associations form episodic representations independent of the memories' representational activation in the semantic network. Episodic patterns of activation and inhibition on these cue-target associations are encoded and reencoded each
time retrieval occurs. The notion of retrieval specificity is inherent in this framework, because it clearly dissociates the effect of activations in the semantic network and pattern shaping during episodic retrieval.

Another general framework that might accommodate both effects is the new theory of disuse (Bjork & Bjork, 1992). This theory differentiates between two types of memory strengths, retrieval strength and storage strength, that can be described by ease of access and learnedness, respectively. The probability of successful retrieval is determined by current retrieval strength. Storage strength indirectly mediates the dynamics of retrieval strength fluctuation. Specifically, the theory predicts that retrieval strength decays over time due to learning of other memories, but forgetting can be slowed down by accumulating storage strength. Successful retrieval of memories with low retrieval strength leads to larger increments in storage strength than successful retrieval of memories with high retrieval strength. That is, a desirably difficult learning event (Bjork, 1994; Bjork & Bjork, 1992; see also Pyc & Rawson, 2009) is more powerful than learning events involving fluent processing. We argue that in this framework, where memory retrieval is a cue-dependent process that alters the state of the memory system (Bjork, 2011), changes in cue-target associations would be retrieval-specific because retrieval provides a desirable difficulty for reshaping patterns of cue-target associations, whereas other forms of reencountering memories do not. However, this theory is a descriptive one, with no specification of underlying mechanisms.

Although it seems that both the SAM-REM model (Malmberg & Shiffrin, 2005) and the episodic inhibition framework (Raesmány & Conway, 2006) can accommodate most of the results reported in the literature, and this dissertation, it is not clear how the interference dependence reported in this dissertation could be accommodated by these theories. It would be interesting to see for instance, how incorporating the role of cognitive control in mediating the modulation of cue-target retrievals could potentially expand the explanatory power of
these theories. A future question is what further specifications are needed for both the SAM-REM model (Malmberg & Shiffrin, 2005), and the episodic inhibition framework (Racsmány & Conway, 2006) to accommodate the present findings.

One line of research that might inform these future quests is research on human reconsolidation. In this framework, cue-dependent, retrieval induced reactivation of memories is believed to render cue-target sets labile, a state in which memory sets can be reconstructed (Hupbach, Gomez, Hardt, & Nadel, 2007a; Nader & Einarsson, 2010; Sara, 2010; Schiller & Phelps, 2011). Importantly, several modulating factors of reconsolidation have been described in the animal, and some already in the human domain (Nader, 2009). The RIF and the testing effect have already been explained by modifications of reactivated memory patterns (Bjork & Bjork, 1992; Finn et al., 2012; Halamish & Bjork, 2011; Racsmány et al., 2010; Racsmány & Conway, 2006), and the framework might be adapted to the TNT effect, based on neuroimaging findings that suggest that hippocampal activation is necessary for the suppression effect to occur (Benoit & Anderson, 2012). Exploring the suggested role of cognitive control in reconsolidating memories as a common mechanism behind modifying cue-target associations during retrieval stopping, competitive retrieval, and repeated retrieval needs future research.
Conclusion

A common great idea behind the works of Michael Anderson, Robert Bjork and Henry Roediger (Anderson & Green, 2001; Anderson, 2003; Bjork & Bjork, 1992; Roediger & Karpicke, 2006a), was to think of retrieval as a memory modifier. Tests have been conceived as mere measurements of what people remember, and not just by lay persons but also by experimental psychologist. In a peculiar twist, as Roediger and Karpicke (2006a) noted, research on the effect of testing has been deemed by a rising interest in forgetting in the 1960's, exactly because experimentalists had been aware that any test before a final test would be a confounding factor in their study.

Now with a renewed interest partly stemming from advances in animal research on reconsolidation (Nader, 2009), the idea of memories being reconstructive (Bartlett, 1932; Schacter et al., 1998) has started again to generate novel research on the factors that might rewrite or modify memories by each act of retrieval, including such wide areas as the effect of sleep, arousal or cognitive control (Finn & Roediger III, 2011; Hardt, Einarsson, & Nader, 2010; Hupbach, Gomez, Hardt, & Nadel, 2007b; Nader & Einarsson, 2010; Nader, 2009; Schiller & Phelps, 2011; St Jacques & Schacter, 2013). This dissertation is part of this research and I hope it will itself foster novel research in the field of memory.
References


Prefrontal Cortex: Evidence from Deficits in Individuals with ADHD.


