Episodic retrieval and forgetting

MIHÁLY RACSMÁNY

AKADÉMIAI DOKTORI ÉRTEKEZÉS

2016

CONTENTS:

ACKNOWLEDGEMENTS

First, I am deeply grateful to Martin A. Conway for his long-term contribution to our joint works and his support as my scientific mentor.

I want to thank to Attila Keresztes, Péter Pajkossy, Gyula Demeter, Ágnes Szőllősi, and Renáta Dombovits, the member of my research group, their contribution to the presented research projects. Their enthusiastic work and scientific ideas shaped significantly our joint research. Special thanks to Ágnes Szőllősi for editing and correcting this dissertation.

I would like to express my deepest appreciation to my following scientific collaborators István Szendi, Csaba Pléh, Gyula Kovács, Zoltán Janka, András Harsányi, Katalin Csigó, Attila Németh, Tamás Kurimay, Attila Krajcsi, Csongor Cimmer, and Daniel Kaiser. Their generous help, expertise and wise advices were fundamental for the presented researches.

I owe a very important debt to Katalin Szentkuti, Luca Frankó, Anna Ágó, Dorottya Pápai and Dóra Bakos for their help in technical assistance of the presented experiments.

I would like to express my gratitude to Ágnes Lukács, Szabolcs Kéri, Márta Zimmer, Anna Babarczy, Péter Simor, Bertalan Polner, Anna Torma, and Csaba Müller, and to the entire community of the Department of Cognitive Science of Budapest University of Technology and Economics for providing me and for my research group a lively and supportive atmosphere.

I also want to thank to two of my former graduate students Edit Anna Garab, Gábor Nagymáté for their assistance in data collections.

I am grateful for the financial support for my research over the years from the Hungarian Scientific Research Fund (Országos Tudományos Kutatási Alapprogramok, K84019, IN77932, K68463), and the Hungarian National Brain Research and Program (Nemzeti AgyKutatási Program, KTIA NAP 13-2-2014-0020), and the János Bolyai Research Scholarship of the Hungarian Academy of Sciences.

Finally, I owe my deepest gratitude to my wife, Patrícia Balázs for the patience and support that allowed me to dedicate time to this effort.

LIST OF PAPERS BY THE AUTHOR CITED IN THE DISSERTATION

Demeter, G., Szendi, I., Juhász, M., Kovács, Z. A., Boncz, I., Keresztes, A., Pajkossy, P., & **Racsmány, M.** (2016). Hypnosis attenuates executive cost of prospective memory. *International Journal of Clinical and Experimental Hypnosis, 64*, 200-212.

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

Szőllősi, Á., Pajkossy, P., & **Racsmány, M.** (2015). Depressive symptoms are associated with the phenomenal characteristics of imagined positive and negative future events. *Applied Cognitive Psychology, 29*, 762-767.

Szőllősi, Á., Keresztes, A., Conway, M., & **Racsmány, M.** (2015). A diary after dinner: How the time of event recording influences later accessibility of diary events. *Quarterly Journal of Experimental Psychology, 68*, 2119-2124.

Pajkossy, P., Simor, P., Szendi, I., & **Racsmány, M.** (2015). Hungarian validation of the Penn State Worry Questionnaire (PSWQ): Comparing latent models with one or two method factors using both paper-pencil and online versions of the PSWQ. *European Journal of Psychological Assessment, 31*, 159-165.

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

Harsányi, A., Csigó, K., Rajkai, C., Demeter, G., Pajkossy, P., Németh, A., & **Racsmány, M.** (2014). The probability of association between response inhibition and complusive symptoms of obsessive-compulsive disorder: Response to Abramovitch and Abramowitz. *Psychiatry Research, 217*, 255-256.

Demeter, G., Keresztes, A., Harsányi, A., Csigó, K., & **Racsmány, M.** (2014). Obsessed not to forget: Lack of retrieval-induced supression effect in obsessive compulsive disorder. *Psychiatry Research, 218*, 153-160.

Harsányi, A., Rajkai, C., Harsányi, K., Demeter, G., Németh, A., & **Racsmány, M.** (2014). Two types of inhibition impairments in OCD: Obsessions, as a problem of thought suppression; compulsions, as behavioral-executive impairment. *Psychiatry Research, 215*, 651-658.

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

Racsmány, M., Conway, M. A., Keresztes, A., & Krajcsi, A. (2012). Inhibition and interference in the think/no-think task. *Memory and Cognition, 40*, 168-176.

Racsmány, M., Keresztes, A., Pajkossy, P., & Demeter, G. (2012). Mirroring intentional forgetting in a shared-goal learning situation. *PLoS ONE, 7*(1), e29992.

Racsmány, M., Demeter, G., Csigó, K., Harsányi, A., & Németh, A. (2011). An experimental study of prospective memory in obsessive-compulsive disorder. *Journal of Clinical and Experimental Neuropsychology, 33*, 85-91.

Csigó, K., Döme, L., Harsányi, A., Demeter, G., & **Racsmány, M.** (2010). Deep brain stimulation for treatment refractory obsessive-compulsive disorder - a case report. *Clinical Neuroscience, 63*, 137-143.

Csigó, K., Harsányi, A., Demeter, G., Harkai, C., Németh, A., & **Racsmány, M.** (2010). Long-term follow-up of patients with Obsessive-Compulsive Disorder treated by anterior capsulotomy: A neuropsychological study. *Journal of Affective Disorders, 126*, 198-205.

Racsmány, M., Conway, M. A., & Demeter, G. (2010). Consolidation of episodic memories during sleep: Long-term effects of retrieval practice. *Psychological Science, 21*, 80-85.

Harsányi, A., Csigó, K., Demeter, G., Rajkai, C., Németh, A., & **Racsmány, M.** (2009). A Dimenzionális Yale-Brown Obszesszív-Kompulzív Teszt (DY-BOCS) magyar forditása és a teszttel szerzett első tapasztalataink. *Psychiatria Hungarica, 24*, 18-60

Racsmány, M. (2009). Modern neuropszichológiai vizsgálómódszerek a klinikumban: Az emlékezet és a kontrollált viselkedés zavarai es diagnosztikája. *Orvosképzés, 84*, 148-150.

Kovács, G., & **Racsmány, M.** (2008). Handling L2 input in phonological STM: The effect of non-L1 phonetic segments and non-L1 phonotactics on nonword repetition. *Language Learning, 58*, 597-624.

Racsmány, M., Conway, M. A., Garab, E. A., & Nagymáté, G. (2008). Memory awareness following episodic inhibition. *Quarterly Journal of Experimental Psychology, 61*, 525-534.

Racsmány, M., Conway, M. A., Garab, E. A., Cimmer, C., Janka, Z., Kurimay, T., Pléh, C., & Szendi, I. (2008). Disrupted memory inhibition in schizophrenia. *Schizophrenia Research, 101*, 218-224.

Csigó, K., Harsányi, A., Demeter, G., Németh, A., & **Racsmány, M.** (2008). Terápiarezisztens kényszerbetegek műtéti kezelése. *Psychiatria Hungarica, 23*, 94-108.

Demeter, G., Csigó, K., Harsányi, A., Németh, A., & **Racsmány, M.** (2008). A végrehajtó funkciók sérülése obszesszív-kompulzív zavarban. *Psychiatria Hungarica, 23*, 85-93.

Demeter, G., & **Racsmány, M.** (2008). Kontrollált emlékezeti előhívás és a frontális lebeny sérülése. *Pedagógusképzés, 1-2*, 55-68.

Kállai, J., Bende, I., Karádi, K., & **Racsmány, M.** (Eds.), *Bevezetés a neuropszichológiába*. Budapest: Medicina Könyvkiadó, 2008.

Racsmány, M. (2008). Amnézia. In J. Kállai, I. Bende, K. Karádi, & M. Racsmány (Eds.), *Bevezetés a neuropszichológiába* (pp. 249-287). Budapest: Medicina Könyvkiadó.

Racsmány, M., Lukács, Á., & Pléh, C. (2008). Szabályos és kivételes morfológia Williams-szindrómában. In C. Pléh (Ed.), *A fejlődési plaszticitás és az idegrendszer* (pp. 87-102). Budapest: Akadémiai Kiadó.

Harsányi, A., Csigó, K., Demeter, G., Németh, A., & **Racsmány, M.** (2007). Dimenzionalitás és neurokognitív eltérések OCD-ben. *Psychiatria Hungarica, 22*, 366-378.

Racsmány, M., Albu, M., Lukács, Á., & Pléh, C. (2007). A téri emlékezet vizsgálati módszerei: Fejlődési és neuropszichológiai adatok. In M. Racsmány (Ed.), *A fejlődés zavarai és vizsgálómódszerei* (pp. 210-239). Budapest: Akadémiai Kiadó.

Demeter, G., Csigó, K., Harsányi, A., Németh, A., & Racsmány, M. (2007). Impaired executive functions in obsessive compulsive disorder (OCD). *Psychiatria Hungarica, 23*, 85-93.

Racsmány, M. (Ed.), *A fejlődés zavarai és vizsgálómódszerei*. Budapest: Akadémiai Kiadó.

Racsmány, M. (2007). Emlékezeti illúziók és kóros tévedések. In I. Czigler & A. Oláh (Eds.), *Találkozás a pszichológiával* (pp. 135-149). Budapest: Osiris Kiadó.

Racsmány, M. (2007). Az "elsődleges emlékezet" - a rövid távú emlékezés és a munkamemória elméletei. In V. Csépe, M. Győri, & A. Ragó (Eds.), *Általános pszichológia 2.: Tanulás, emlékezés, tudás* (pp. 177-208). Budapest: Osiris Kiadó.

Racsmány, M. (2007). Kódolás és előhívás az emberi emlékezetben. In V. Csépe, M. Győri, & A. Ragó (Eds.), *Általános pszichológia 2.: Tanulás, emlékezés, tudás* (pp. 159-176). Budapest: Osiris Kiadó.

Harsányi, A., Csigó, K., Demeter, G., Németh, A., & **Racsmány, M.** (2007). A kényszerbetegség új megközelítési lehetőségei: A dopaminerg teóriák. *Psychiatria Hungarica, 22*, 248-258.

Racsmány, M., & Conway, M. A. (2006). Episodic inhibition. *Journal of Experimental Psychology: Learning, Memory & Cognition, 32*, 44-57.

Kovács, G., & **Racsmány, M.** (2006). Munkamemória es idegennyelv-elsajátitás: Az idegen beszédhangok hatása a verbális rövid távú emlékezetre. *Magyar Pszichológiai Szemle, 61*, 399-431.

Racsmány, M. (2006a). Elveszett emlékek. In I. Kovács & V. Z. Szamarasz (Eds.), *Látás, nyelv, emlékezet* (pp. 161-166). Budapest: Typotex Kiadó.

Racsmány, M. (2006b). Átmeneti emlékezés. In I. Kovács & V. Z. Szamarasz (Eds.), *Látás, nyelv, emlékezet* (pp. 149-160). Budapest: Typotex Kiadó.

Racsmány, M. (2006c). Interferáló emlékek. In I. Kovács & V. Z. Szamarasz (Eds.), *Látás, nyelv, emlékezet* (pp. 137-147). Budapest: Typotex Kiadó.

Racsmány, M. (Ed.), *Neuropszichológiai diagnosztikai módszerek.* Budapest: Akadémiai Könyvkiadó, 2006.

Racsmány, M., Lukács, Á., Németh, D., & Pléh, C. (2005). A verbális munkamemória magyar nyelvű vizsgálóeljárásai. *Magyar Pszichológiai Szemle, 60*, 479-506.

LIST OF ABBREVIATIONS

ACC	Anterior cingulate cortex
DF	Directed forgetting
DLPFC	Dorsolateral prefrontal cortex
DSF	Digit span forward
DSM-IV	Diagnostic and Statistical Manual of Mental Disorders
ER	Episodic retrieval
F group	Participants received forget instruction
fMRI	Functional magnetic resonance imaging
F response	familiarity response in recognition task
F-words (F-items)	Words (items) instructed to forget instruction in the directed forgetting procedure
HAM-D	Hamilton Rating Scale for Depression
LS	Learning Set
NRP items	Nonpracticed items from unpractised categories
NTBRE	Not to be recollectively experienced
OCD	Obsessive-compulsive disorder
PAs	Paired associates
PANSS	Positive and Negative Syndrome Scale
PFC	Prefrontal cortex
PPL	Posterior parietal lobe
PTSD	Posttraumatic stress disorder
R, K, G responses	Remembering, knowing, guessing responses in recognition task
R group	Participants received remember instruction
RIF	Retrieval-induced Forgetting
ROI	Region of interest
RP	Retrieval practice

Rp+ items	Practiced items from the practiced categories
Rp- items	Nonpracticed items from the practiced categories
R response	Remember response in recognition task
R-words (F-items)	Words (items) received remember instruction in the directed forgetting procedure
SAM	Search of Associative memory
SANS	Scale for the Assessment of Negative Syndromes
STAI	Spielberger State and Trait Anxiety Inventory
TBF items	To-be-forgotten items in directed forgetting procedure
TBR items	To-be-remembered items in directed forgetting procedure
TNT	Think/No-Think Task
VPT	Visual Pattern Test
WCST	Wisconsin Card Sorting Task
WM	Working memory
Y-BOCS	Yale Brown Obsessive-Compulsive Scale

I. Background and objectives

I.1. Historical background: the role of interference and inhibition in human forgetting

Although the history of experimental research on human memory embraces almost 150 years, the investigation of retrieval processes was a neglected area until the late 1960s. For most of the scholars of human learning and memory, a fixed testing condition was a gold-standard of memory experiment to adequately assess learning performance. However, since the pioneer work of Endel Tulving (e.g. Tulving and Pearlstone, 1966, Tulving and Thomson, 1973), retrieval was not regarded as a fixed condition in memory experiments anymore, instead as a dynamic, self-limiting process, which can significantly alter remembering. Forgetting has been a central topic of research on human memory since the seminal work of Ebbinghaus (1885). From the point of view of the presented studies in this dissertation, it is a crucial issue whether forgetting is an adverse side effect of inadequate consolidation processes or an adaptive and inherent feature of retrieval processes. The following short review aims to summarize the most important findings and theories concerning the complex relationship between retrieval and forgetting.

I.1.1. Retrieval cue and forgetting

The first significant scientific quarrel in the realm of memory research concerning human forgetting was between trace decay and interference theories. Whereas trace decay theories focused on the representation of target memory, interference theories emphasised the relationship between retrieval cues and target memories (see reviews of Crowder, 1976; Anderson & Bjork, 1994; Anderson & Neely, 1996). Trace decay theories assumed that the representation of target memories deteriorates with time, memory traces simply dissolve and become no more available (Jenkins & Dallenbach, 1924; Hebb, 1949; Minami & Dallenbach, 1946). In contrast, according to theories of interference, the representation of target memories remains intact, but later learning affects the relationship of cues and target memories. From the numerous experimental paradigms that have been developed to check this latter assumption, those two would be discussed here which hold the greatest significance regarding the theories and experiments of this dissertation. One of these experimental procedures is the A-B, A-D learning paradigm, where ‘A’ items designate the retrieval cues in two learning lists whereas ‘B’ and ‘D’ are for the associated response items.

In the A-B, A-D learning paradigm there is a decrease of retrieval performance of the first learning list, which is a result of the retroactive interference caused by the second learning list. One of the most significant explanations given to this phenomenon was the so-called "occlusion hypothesis" of McGeoch (1936, 1942), claiming that forgetting is a result of the competition of items connected to identical retrieval cues. The more identical the two response items are, the greater their competition would be, so in an A-B, A-D learning situation the recall of B responses would be impaired due to the adverse effect of D responses. It is important to note that, however McGeoch (1942) often uses the term retroactive inhibition in his works, it is only a description of the impaired retrieval of B responses and does not refer to any underlying mechanism involving the active inhibition of an item or a cue-target relationship.

Although according to McGeoch, the retroactive inhibition appearing in the A-B, A-D learning situation depends on the strength of A-D relationships, i.e. the intrusion of D responses, he considered the effect of two other factors as well, that of the *altered stimulus* and of the *searching of the inappropriate set*. These two factors of McGeoch's theory did not receive much attention later, and so his concept has widely been regarded as a simple response-competition model ever since (e.g. Anderson & Neely, 1996; but see Crowder, 1976). However, from the point of view of theses of this dissertation these two factors bear with significance. The phenomenon of *altered stimuli* calls attention to the fact that during episodic coding the representation of a nominal stimulus is modified by the contextual, episodic information of the learning situation and the functional stimulus the subject aims to retrieve later is a result of this coding process. Therefore it is possible that the original nominal stimuli will not be appropriate later to activate the functional stimuli (response) which results in forgetting as a behavioural consequence (McGeoch, 1942; see also Guthrie, 1935; Estes, 1955). The third important factor according to McGeoch is the selection of appropriate searching strategies, as forgetting may be the consequence of the subject's searching an irrelevant learning set, in other words forgetting could take place, even with intact memory representation, as a consequence of inappropriate retrieval cue.

The other significant account of retroactive inhibition effect in the A-B, A-D learning paradigm was introduced by Melton and Irwin (1940). This hypothesis suggested that the later learning of A-D association extinguishes the A-B association. This account presumes that the intrusion of newly learned responses explains only some part of the forgetting effect. This account based on findings showing that the amount of intruding errors deriving from A-D responses stops increasing after a while when repeated intervening learning phases are used, whereas the retroactive inhibition appearing in the number of A-B responses keeps increasing. According to Melton and Irwin (1940) the only explanation for this result can be the gradual decreasing of the strength of A-B relationships while A-D relationships are developed. This 'unlearning' theory therefore claims that a permanent change occurs in the representational activity of cue-target relationships. It should be emphasised that later main stream 'inhibitory theories" do not consider this account an "inhibitory" one for two reasons: one is that the change does not occur in the representation of the target memory but in the cue-target relationship, and the other is that this change is not a temporary suppression but has a more or less permanent nature.

Beside the A-B, A-D learning situation another experimental procedure, called part-set cueing became highly influential in experimental research of forgetting. It refers to a situation affecting episodic memory when the retrieval of certain items of a set (a learning list, a category etc.) reduces the probability of retrieving other items from the same set (Slamecka, 1968, 1969). It has been shown that the part-list cueing effect can be released with the same intensity whether the cues are words of the original learning list, or extra-list words from the same category (Roediger, 1973; Watkins, 1975). Some studies have even found that intra-list words caused a bigger_part-list cueing effect when cues were from the same category as target words, and the effect existed with non-categorised lists, as well (Mueller & Watkins, 1977; Roediger et al., 1977; Roediger & Neely, 1982; Nickerson, 1984).

Although several explanations have been given to the part-list cueing effect, here only those models will be discussed in the following short paragraphs which have some relevance to the theories and paradigms of memory that are in the focus of this dissertation.

The main reason why part-set cueing has evoked such great interest in researchers of human forgetting was that this phenomenon questioned the previously dominant theoretical models assuming a direct association between memory items and an activation spreading along this associative chain. According to these theories the presentation of associating cues should facilitate and not inhibit the retrieval of target memory items (e.g. Anderson, 1972; Postman, 1971; Collins & Loftus, 1975).

Watkins (1975) explains the effect so that cues increase the length of the list to be remembered which, in turn, decreases the retrieval probability of particular items. This idea was, however, undermined by results showing that extra-list cues elicit weaker inhibition and that intra-list words not from the categories of tested words do not elicit an inhibitory effect (Roediger et al., 1977; Mueller & Watkins, 1977). Taking these empirical data into consideration, Mueller and Watkins (1977) modified Watkins' original idea: their so-called *cue-overload account* suggests that target words and cue words are integrated in a higher unit or episode and the more items are associated to a cue, the lower the retrieval probability of particular items would be (Murdock, 1962; Tulving & Pearlstone, 1966). This means that in the retrieval phase cues from the same category overload the retrieval cue which results in a decreased retrieval performance. This assumption is supported by the so-called fan-effect: the more facts a subject has to associate to the same topic area, the longer the retrieval time of particular facts would be (J. R. Anderson, 1974).

The two theories to be presented next are based on the same principle as that of McGeoch (1942): inhibition is the consequence of the fact that some cue-item relationships are strengthened at the expense of other cue-item links. According to the influential *sampling-with-replacement* model of Rundus (1973), memory structures are hierarchically organised and memory search is controlled by various set cues or control elements (the phrase of Estes, 1972). Memories acquired during list learning are organised into a three-level hierarchy, in which the highest level represents the list-wide context and the middle level includes category names (i.e. the control elements) through which we can access the lowest level of particular words. The retrieval probability of particular words depends on how strongly they are related to control elements, and the retrieval of a certain item enhances its relationship with the control element. Memory retrieval stops automatically if repeated retrieval attempts do not produce new items. The part-set cueing effect arises because during retrieval certain control elements are presented, thereby related items are repeatedly retrieved, and after a while the searching process stops. The idea of Rundus is supported by Roediger's (1978) findings that the more category names are presented from a learning list, the lower retrieval probability would words from other categories have. The model of Rundus, therefore, explains part-set cueing with the principle of occlusion without assuming the operation of a specific inhibitory mechanism.

A very similar, and up to the present very popular, concept is the SAM (search of associative memory) model of Raaijmakers and Shiffrin (1981). In contrast with the hierarchical theory, the SAM model assumes that the phenomenon of part-set cueing cannot be fully explained by vertical relationships between cues and target memories, but the importance of inter-item relationships should be considered as well (Slamecka, 1972; Rundus, 1973; Roediger, 1974). According to the SAM model, memory retrieval is a cue-dependent process, in which in each step of the search process a particular set of cues is assembled that activate certain target items (or images, or samples). The information of the sampled image is accessed, unpacked and evaluated, which process is called recovery. Thus, in the free-recall phase subjects first retrieve specific contextual cues associated to the list-learning situation, then the first target items that are the most intensely related to these cues. From this point on cues already involve both these retrieved items and the context, and through item-item associations further items get retrieved. This retrieval cycle goes on as long as new items can be accessed, and when there is no more, a stopping rule terminates the searching process. An interesting suggestion of the SAM model is that although in many cases the changing of the cue-target relationship is responsible for forgetting - for example in the A-B, A-D paradigm -, this is not the case with the part-set cueing phenomenon. Using computer simulation, Raaijamakers and Shiffrin (1981) showed that neither the enhanced relationship between context and particular target items, nor the stopping rule are accountable for the part-set cueing effect (although those were the central features of the theory of Rundus, 1973). The only factor responsible for eliciting the effect is that while in a free-recall situation subjects first recall those items that have the strongest connection with the context and with other items, in a part-set cueing situation the cues given are not those with the strongest connections, so they would not help the recall of so many items. It is important to emphasise that although there is a fundamental difference between the explanations of Rundus (1973) and Raaijamakers & Shiffrin (1981) regarding the part-set cueing phenomenon, both claim that retroactive inhibition is due to the increase in the activation of cue-item *relationships* (or item-item *links* in the SAM model) and do not presume a change in the activation of the representation of cues or target memories themselves.

The following group of theories of retroactive inhibition has the common starting point that one cue can have only one related memory response (Martin, 1971). So, for example, although the "A" words are identical in the first and second learning lists the A-B, A-D paradigm, they are influenced by the various responses associated to them and by the contextual signs during encoding process, so that they become functionally different. After the A-D learning phase the functional "AD" variant of the nominal "A" cue would be more accessible, so subjects would more probably produce a "D" response. This presumption, also referred to as meaning bias, explains part-set cueing effect as a result of the change in the representation of cues (Sloman et al., 1991). According to this idea, the greater the overlapping of functional cues is in the learning and in the test phases, the better the recall performance would be, and since the arbitrary cues in the part-set cueing situation are not congruent with those in the learning situation it would result in an impaired recall performance.

The theory to be finally discussed assumes a change in the organisation of responses to be the underlying mechanism of retroactive inhibition. The *response-set inhibition theory* of Postman et al. (1968) claims that the facilitation of new responses and the inhibition of old ones may underlie forgetting. During recall subjects have to respond to a cue and when several responses are associated with the same cue, subjects have to exclude intruding alternative responses in order to select the right one. Postman assumes this to be the reason why retroactive interference decreases in recognition tests where subjects have the right response in front of them and they only have to pick it from among the others, and also in the so-called "modified-modified" retrieval situation where more than one response can be given to a given cue. Although Postman et al. (1968) define retroactive inhibition as a phenomenon where the whole set of inhibited items gets suppressed, they do not make a clear statement as for the activation level of to-be-forgotten items. Their theory allows for a possibility that retroactive inhibition is elicited by subjects' use of some searching criterion (they call it a "general purpose-selector mechanism") which rejects emerging target memories in case of certain contextual signs (e.g. signs referring to the first list in the A-B, A-D paradigm). In this case, however, the activation level of the representation of target memories would be unchanged and the *retrieval strategy* could be responsible for retroactive inhibition (Postman et al., 1968, Postman & Underwood, 1973). The model of Postman and his colleagues gives a clear explanation for retroactive inhibition appearing in the A-B, A-D paradigm, where neither a competition of responses, nor cue-overload is present. However, it fails to interpret results when retroactive inhibition occurred in a mixed-list experimental design where both experimental A-D items and control C-E items were randomly intermixed in the list following the A-B learning phase (see Crowder, 1976; Anderson & Neely, 1996).

I.1.2. The rise of the concept of retrieval inhibition

In the 1960's a new experimental paradigm was invented, which caused difficulties for contemporary theories of human memory. In the experimental procedure called "directed forgetting" subjects have to intentionally forget some previously learned information. The difference between this situation and those annoying everyday examples of spontaneous forgetting is that in this case the subject *wants* to forget something in order to be able to learn something else more successfully.

Two basic experimental procedures of directed forgetting has been developed: one is the "item method" or "word-by-word" which means that during the learning of a word list (item list) each word is accompanied by a cue indicating whether that particular word should later be remembered (R-word) or forgotten (F-word). The other procedure is the "list method" in which the learning of a whole word list is followed by a cue (remember or forget) and then a second list has to be learned.

A plethora of experimental data show that the recall of words is significantly worse following an F instruction than in the control group receiving an R instruction. An important finding is that F-words do not cause the same level of proactive interference than R-word does: the amount of previously learned F-words does not influence the recall performance of R-words (Bjork, 1970). The impaired recall of F-words and improved recall of R-words (this is called the DF effect) suggest that subjects "obey" the instruction and really forget the F-words. Another interesting phenomenon is that F-words do not intrude the free recall of R-words (Reitman et al., 1973), even if there is an explicit instruction to recall F-words (e.g. if the recall of each F-word is rewarded, (Woodward & Bjork, 1971).

Bjork originally suggested that there are two mechanisms underlying these phenomena. One is the segregation or separate grouping of F- and R-items and the other is the selective rehearsal of F- and R-items (Bjork, 1970). According to the hypothesis of selective rehearsal, subjects keep the items in mind with a shallow rehearsal until they receive the cue (F or R cue), and only when particular items turn out to be R-words do they start a deeper, more elaborate rehearsal. The theory of selective rehearsal is supported by such findings like that extending the time before the appearance of the cue has no influence on recall (Woodward et al., 1973) and that the forgetting curve is significantly steeper in the case of F-words (Weiner, 1968).

According to the principle of segregation, a successful DF effect can only be achieved if subjects are able to separate F- and R-items properly. Shebilske et al. (1971) showed that categorised F-items tend to intrude much less than those not categorised, however, the segregation of F- and R-items is possible within one category, as well (Woodward & Bjork, 1971). The importance of segregating the two sets was further shown by Geiselman and Richle (1975) in their experiment on sentences: when both R- and F-sentences were categorised, the recall performance of R-sentences was much better than in the other case when all sentences were mixed together.

Another interesting result is that when F-items are categorised and R-items are not, the recall performance of R-items is much better than in the reverse case (when R-items are grouped and F-items are not). This means that the F-instruction can reduce proactive interference only if F-items are organised into a set (see MacLeod, 1998). For the theoretical assumptions concerning directed forgetting it is an important result that the forget instruction can reduce only proactive interference but not retroactive interference. For example, when, using the list method, subjects were told to forget the second list after having learned it, no DF effect was found (Block, 1971). It is therefore crucial to give the forget instruction during the encoding process, because if it is given during recall, no DF effect would occur.

There is much debate concerning the nature of the forget instruction, and experimental results are rather contradictory. Elmes et al. (1970) found that the F-cue has to be explicit in order to achieve successful directed forgetting, while Epstein (1969) could elicit a DF effect also when subjects had to recall only the end of a word list (i.e. when the forget instruction was implicit). Weiner and Reed (1969) gave their subjects two kinds of F-cues, "forget it" and "don't rehearse it", and found that the instruction calling for non-rehearsal was much less effective than the explicit forget instruction. The forget instruction proves to be most effective when it is given directly after learning the F-words, and it is much less effective when some R-words have to be learned as well before the F-instruction (Timmins, 1974). The recall of R-words is best when there is a short delay before the cue, while in the case of F-words there should be a long delay. This supports the assumption that the DF effect is partly caused by the different rehearsal techniques of R- and F-words.

Although the mechanism of selective rehearsal is useful in understanding directed forgetting, intentional forgetting is a general characteristic of the cognitive system and, therefore, it does not only concern the learning of word lists. For example, Burwitz (1974) demonstrated directed forgetting in the learning of simple motor responses. In his experiment subjects had to learn the operation of a lever without any visual feed-back. After learning four different movements subjects had to repeat a fifth one and they could repeat it significantly better when they were told to forget the previous four movements. Cruse and Jones (1976) used unrehearsable sounds as stimuli: when some of the sounds given were F-sounds, the reaction time of the recognition of R-sounds was much shorter than in the case when there were only R-sounds. These experiments all support the assumption that selective rehearsal itself cannot fully explain the directed forgetting effect.

It would also be important to know how the elaborateness of encoding affects the pattern of directed forgetting. According to the hypothesis of selective rehearsal, subjects keep the F-words in mind with a shallow rehearsal until they receive the cue, and right after the forget instruction they stop processing F-words. It is not yet clear what happens when the elaborateness of the processing of F-words is manipulated. The experimental results are contradictory, for example Wetzel (1975) found that manipulating the level of processing does not influence the DF effect, while Horton and Petruk (1980) showed that the recall rate of R-words is significantly higher in case of semantic encoding than in case of phonemic or structural encoding. Horton and his colleagues conclude that the deeper level of encoding separates F- and R-sets more clearly. However, Geiselman, Rabow, Wachtel and MacKinnon (1985) found no DF effect in case of the deep processing (synonym generation) of F-words.

Another interesting result is that maintenance of rehearsal affects knowing and not remembering, while elaborative rehearsal affects remembering and not knowing (Gardiner, Gawlik & Richardson-Klavehn, 1994).

Bjork (1970), who originally assumed segregation and selective rehearsal to be the two main causes of the DF effect, suggests that the basis of segregation may be the chronological information concerning the position of R- and F-words. This hypothesis is supported by the fact that there is a difference in how well subjects can determine the sequential position of recalled R- and F-words. According to Tzeng et al. (1979), if subjects have to decide which group of five words contained a particular word, they can locate R-words much more precisely than F-words. These results were later replicated by others as well (e.g. Jackson & Michon, 1984). In the 1970s one of the most challenging questions was to explore the exact differences between the encoding of F- and R-words. The findings are quite contradictory as in many experiments both the recall and the recognition of F-words were less successful, while in other cases a DF effect appeared only in recall but not in recognition tasks where, in the latter case, the advantage of R-words disappeared. Basden and Basden (1998), however, pointed out that these contradictions were due to the use of different experimental techniques. When using the item method there is a DF effect both in recall and in recognition tasks, while in the case of the list method it is only in recall that F-words reliably show a significant decrement.

These contradictions suggest that there are different mechanisms underlying these two forms of directed forgetting. In the item method one of the most important mechanisms seems to be that after receiving the cue subjects do not rehearse the F-words any more. This assumption is supported by the fact that both the recall and recognition performances of F-words are poorer in this situation, subjects simply do not learn the F-words so well. The mechanism eliciting a forgetting effect in the list method was first demonstrated by Geiselman, Bjork and Fishman (1983) in an important experiment. Their subjects had either to learn some words or just decide about their attractiveness, while, at the middle of the word list, they were given a forget instruction. This caused an impairment in the later recall of the first half of the word list, surprisingly to the same extent in the case of learned words and words judged for attractiveness. Selective rehearsal could not have played any role in the case of words judged for attractiveness since those words did not have to be learned at all. To explain their results Geiselman and his colleagues assumed that the recall of F-words declined as a consequence of the retrieval inhibition triggered by the F-cue. This idea is supported by the fact that the recognition performance of F-words is just as good as that of R-words, which means that the appearance of F-words in the recognition test releases inhibition.

Geiselman and Bagheri (1985) gained further evidence in support of the retrieval inhibition hypothesis when they found that if subjects are instructed to re-learn both the F- and the R-words after the DF procedure, then there is a greater improvement in the case of F-words than in the case of R-words. Now subjects were much more likely to recall F-words than R-words from all the words that previously could not be recalled. In the second experiment of Geiselman and Bagheri the improvement in the recall performance of F-words was so significant after re-learning that subjects were able to recall more F-words than R-words. Considering these results Bjork (1989) proposed that an inhibitory process operates in case of the list method that prevents the recall of F-words. In Bjork and his colleagues' definition this inhibition is a "possible mechanism that results in loss of retrieval access to inhibited items, without a commensurate loss, if any, in the availability of those items" (Bjork, Bjork & Anderson, 1998, p.105.).

Bjork emphasises that inhibition is meant in a strong sense (i.e. as suppression), indicating that inhibition is not the automatic result of the strengthening of other items but of an explicit suppression of the to-be-forgotten items (Bjork, 1989). However, this does not necessarily imply that the person can intentionally control the inhibitory process itself. This hypothesis is, indeed, very similar to Freud's original concept of repression (see Erdelyi & Goldberg, 1977). We might even say that many researchers find the directed forgetting procedure so interesting exactly because they expect it to be the experimental model of repression (Anderson & Green, 2001; Conway, 2001, but see Kihlstrom, 2002). The psychoanalytic literature considers repression a process used by the subject for referring certain thoughts, images and memories related to instinctual drives to the domain of the unconscious. Repression happens when the satisfaction of a certain impulse would cause unpleasure from another point of view (Freud, 1926).

Although later psychoanalytic writings consider repression an unconscious defence mechanism, Freud originally described it as an intentional suppressive technique. The relationship between directed forgetting and repression appears so evident that there had already been some attempts to connect the two, even before the inhibition theories of DF were developed. For example, Weiner and Reed (1969) thought that the study of the directed forgetting phenomenon would help to explore the mechanism of repression.

The A-B, A-D learning paradigm - where B and D are loosely associated words - has already provided some relevant results: if subjects got a slight electric shock during learning D, then the re-learning of B was found to be impaired, too (Glucksberg & King, 1967). According to Weiner (1968), forgetting happens faster in the case of those items that are associated with an electric shock both during encoding and retrieval. However, despite Weiner's findings, the similarities between repression and directed forgetting were ignored for a long time, as researchers were mainly interested in the characteristics of F-words and the circumstances under which they may appear in recall.

Inhibition resulting from the competition of different representations is the central element of the hypothesis of Conway and his colleagues concerning directed forgetting (Conway et al., 2000). They propose that inhibition occurs in a directed forgetting procedure only if the to-be-forgotten and to-be-remembered items are similar in type. The idea here is that similar but distinct lists compete in memory in terms of their potential memorability. If two lists have a similar potential, then they interfere with each other and one of them must be inhibited in order to maintain the potential memorability of the other. This hypothesis is based on the function of directed forgetting - the reducing of the disturbing effect of present but irrelevant information -, and this occurs when to-be-forgotten items interfere with to-be-remembered items.

It is also important to explore the retrieval circumstances under which the forget instruction can evoke a retrieval inhibition effect on items to-be-forgotten. Bjork and Bjork (1996) believes that to elicit inhibition the original learning episode should be present somehow. In case the retrieval of to-be-forgotten items happens in a situation when the subject cannot access the original learning episode, then no DF effect would appear (Bjork & Bjork, 1996). This hypothesis is supported by a study of Bjork and Bjork (1996), in which they used a standard, list-learning directed forgetting task, where after the second list but before recall they presented subjects a recognition test containing F-items, and they found that there was no DF effect (the F-words have been released from inhibition). On the other hand, when the F-words were presented before recall in a word fragment completion task, the completion of F-words was similarly successful to that of R-words, but a DF effect was still present at recall.

Not all researchers agree with this idea, though. According to the findings of MacLeod (1989), the performance of R-words is much better than that of F-words in such implicit memory tests like the word fragment completion task, although there is some facilitation in the case of F-words, too. For instance, in a lexical decision task the responses given to F-words are significantly slower than those given to R-words, indicating that the inhibition of F-words can occur without the presence of the original learning episode. The results are controversial, but Paller (1990) pointed out that the two implicit tests used by MacLeod (fragment completion and lexical decision tasks) can be solved with explicit memory strategies as well (see Allen & Vokey, 1998; Hauselt, 1998; Squire et al., 1987; Toth, Reingold & Jacoby, 1994). This means that the impaired performance of F-words could be the consequence of the fact that, after all, subjects did recall the original learning episode. Another problem of MacLeod's study is that he used the item method, in which case the DF effect is primarily caused by selective rehearsal and not by retrieval inhibition (see Basden et al., 1993).

Altogether, the results of directed forgetting experiments can hardly be solely explained by the interference theories. This concerns primarily the experiments done with the list method - although most of the results received with the item method of directed forgetting are well explicable by assuming the operation of a selective coding mechanism, in the list method the information acquired prior to the forget instruction seem to be subject to retroactive inhibition and this inhibition is expressed only in the recall phase. Since there is no difference whatsoever between the "forget" and the "remember" situations regarding the new information to-be-remembered (List 2), this phenomenon cannot be interpreted by assuming that previously learned items had been excluded from recall due to the confirmation of newly acquired items. Nor can the unlearning hypothesis acceptably explain this finding, since the effect is temporary and regards only recall. What is more, the re-presentation of items causes a rebound effect, an impossible phenomenon unless cue-item links are eliminated. Perhaps Postman's (1968) response-set suppression theory is the most acceptable explanation, as it is also supported by Bjork's (1996) finding that if some F-items are presented during the distracting task in the delay phase, then the recall of not presented F-items is also significantly improved which eliminates the DF effect. Postman's theory would, however, have difficulty in explaining results received in the retrieval-induced forgetting procedure, the other inhibitory paradigm used by Anderson and his colleagues.

I.1.3. Context-based explanations of intentional forgetting

Although explanations in the 1970s preferred the idea that the experimental effect is due to different encoding of to-be-forgotten and to-be-remembered items (segregation and selective rehearsal of List 1 items, e.g. Bjork, 1970), the two accounts that have become dominant in the directed forgetting literature are the retrieval inhibition theory (Bjork, 1989; Geiselman, Bjork, & Fishman, 1983) and the context-change account (Sahakyan & Kelley, 2002). The retrieval inhibition theory posits that the F-instruction recruits inhibitory processes in order to suppress the accessibility of the to-be-forgotten items at final recall. It is an important aspect of this theory that the inhibition of List 1 items is regarded a goal-directed and adaptive process (Anderson, 2005; Bjork, 1989; Conway, Harris, Noyes, Racsmány, & Frankish, 2000). The goal of the inhibition of List 1 items is to decrease the interference of these items with List 2 items, in other words, to facilitate the learning of relevant items at the expense of irrelevant information. This adaptive nature of the directed forgetting phenomenon is demonstrated by some experimental results showing that the F-instruction of first list items was not successful without a consecutive to-be-remembered learning list, underlying the assumption that suppression of the first list serves the goal of decreasing the memory load during the learning of the List 2 items (Gelfand & Bjork, 1985; Pastötter & Bäuml, 2007).

The context-change account suggests that the F-cue elicits a kind of mental context change in participants and this between-list shift in mental context will cause a mismatch of contexts between encoding and later retrieval for List 1 items, so directed forgetting is simply a further example of context-dependent forgetting (Sahakyan & Delaney, 2003; Sahakyan & Kelley, 2002). Sahakyan and her colleagues provided a range of experimental evidences that instructing participants to change their mental context (e.g. imagine a specific environment during encoding of List 1 and another one during encoding of List 2) can simulate the cost and benefit effects of the standard F-instruction (Sahakyan & Kelley, 2002).

Although inhibitory and context-change accounts of directed forgetting posit different factors in their explanation, recent versions of these theories equally assume that the cost (lower memory performance for to-be-forgotten items) and the benefit (higher performance on List 2 items in the F group) of the F-instruction are due to different mechanisms. Pastötter and Bäuml (2010) in their reset-of-encoding hypothesis suggested that the cost of F-instruction is due to retrieval inhibition of List 1 items, whereas the benefit is the indirect consequence of decreased memory load during encoding of the second list items. Sahaykan and Delaney (2003) suggested that the cost of the F-instruction is due to a change in internal context, whereas the benefit is the consequence of more elaborated encoding of List 2 items.

I.2.The adverse effect of retrieval practice: retrieval induced forgetting (RIF)

Retrieval can enhance learning but interestingly it can also induce forgetting of related memories, a phenomenon known as retrieval-induced forgetting. Anderson, Bjork and Bjork (1994) produced compelling evidences that the cued recall of an item can impair later recall of items previously associated to the same cue, and this phenomenon was labelled retrieval-induced forgetting. According to Anderson and his colleagues (Anderson & Spellman, 1995; Shivde & Anderson, 2001) an important property of retrieval-induced forgetting is cue-independence, i.e. the inhibition caused by retrieval generalises to any other cue used to test that item. This means that the forgotten competitive item itself is impaired by an active suppression when a related target is sufficiently retrieved (Anderson & Neely, 1996). Anderson and his colleagues developed a three-phase paradigm to study the mechanism of how memory retrieval impairs interfering memories (Anderson, Bjork & Bjork, 1994; Anderson, & Spellman, 1995; Anderson, & McCulloch, 1999). In the *study phase* of this procedure subjects study category-exemplar pairs, the standard procedure consisting of six exemplars in each of eight different categories. After the study phase subjects participate in a *practice task* where they recall three exemplars from half (i.e. four) of the categories, induced by presenting the category name together with the first two letters of the exemplar. After a steady retention interval a final, *category-cue-directed recall* is administered. The well-replicated result is that the recall performance of unpractised items from partially practised categories (Rp- items) is significantly below the performance of nonpractised items from unpractised categories (Nrp items) (Anderson & Bjork, 1994; Anderson & Neely, 1996).

According to Anderson and his colleagues, the impaired recall performance of competing unpractised items reflects the operation of an active suppression mechanism (Anderson & Spellman, 1995; Anderson & Neely 1996). This account is in agreement with many inhibitory theories in interference literature, which assume that active deactivation of interfering items plays an important role in human forgetting (e.g. Carr & Dagenbach, 1990; Dagenbach & Carr, 1994; Zacks & Hasher, 1994). Anderson and Spellman (1995) emphasise that inhibited items are rendered generally. According to the suppression theory of Anderson and his colleagues, Rp- items get inhibited because they are associated to the same cues (in this case category names) as practised items, therefore when practised (Rp+) items are presented, an interference occurs (Anderson & Neely, 1996). This idea implies that the more intensely Rp+ and Rp- items compete with each other, the greater the inhibition of Rp- items would be. Anderson and his colleagues received exactly this result in an experiment where they manipulated the intensity of the relationship between cues and items (Anderson et al., 1994), so that exemplars were either of very high frequency (e.g. FRUIT-*APPLE*) or very rare (e.g. FRUIT-*PAPAYA*). When the cue is presented the high frequency items would very probably be recalled, as they are strongly associated to the cue. The suppression theory predicts that the recall probability of high frequency items would be greater than that of low frequency items, in case they are Nrp words (exemplars of unpractised categories), but their recall rate would be lower in case they are Rp- words (exemplars of categories from which other items were practised). That was exactly the result Anderson and his colleagues (1994) received: the recall rate of high frequency exemplars was lower than that of low frequency ones if they were Rp- words, but it was significantly higher when they were Nrp words.

The above findings unquestionably support the theory of cue-independent inhibition, however, they leave open the important question of what exactly gets inhibited in this experiment. Anderson and his colleagues assume that in retrieval-induced forgetting the target items *themselves* (Rp- words) are inhibited, and they call this concept "the theory of cue-independent forgetting" (Anderson & Spellman, 1995; Shivde & Anderson, 2001). This concept is underlined by experiments where retrieval-induced forgetting was elicited in categories with partly overlapping exemplars. The theory of cue-independent forgetting predicts that if subjects learn the exemplar of *STRAWBERRY* together with the cue FOOD - but, according to their prior knowledge, it might as well be a member of the category RED -, and in the practice phase other exemplars of the category RED are practised (e.g. *BLOOD*) but not the word *STRAWBERRY* - i.e. *STRAWBERRY* becomes an Nrp word -, then in the recall phase *STRAWBERRY* will be inhibited as if it were an Rp- word, although no other exemplars from its originally associated category (FOOD) were practised. According to Anderson and his colleagues, this means that it is not in the category-exemplar relationship where inhibition occurs, but the item itself (*STRAWBERRY*) gets inhibited, independent of the category. Obviously, the item would not be inhibited if exemplars were not overlapping, in that case the recall of *STRAWBERRY* and other Nrp words would be significantly better than that of the unpractised Rp- words of the category RED (Anderson & Spellman, 1995; Anderson & Green, 2001). So Anderson and his colleagues believe that retrieval-induced inhibition is independent of the cue-target relationship evolving in the process of learning, but it concerns directly the representation of the intruding target item. The function of inhibition appearing in the recall phase is to keep out intruding but irrelevant items from consciousness. Once inhibition has appeared, it would be independent of specific cue-target associations developed during learning, so the target item would be impossible to recall, whatever cue is used, since the representation of the target item itself would be suppressed. Anderson and his colleagues suggest that in the process of accessing a particular item from a previous set the focus of selective attention turns towards an object no more present in reality (Anderson & Neely; Anderson, Green & McCulloch; Shivde & Anderson, 2001), so as a consequence of focused attention ignored items may become inhibited during recall. According to Anderson et al., retrieval-induced inhibition is the result of a similar process as the one underlying the negative priming effect (Anderson & Spellman, 1995; Anderson & Neely, 1996). This phenomenon occurs when attention has to be focused on one particular set of stimuli among several others, and as a consequence, when attention is later focused on previously ignored stimuli, they will be processed much more slowly than they would if they were not intentionally ignored before (Neill, 1977; Tipper, 1985). Therefore, Anderson and his colleagues presume that something similar happens in retrieval-induced inhibition where competing items that should be ignored become inhibited by the inner focus of attention, controlled by retrieval processes. Disregarding the otherwise important problem that there are alternative explanations for the negative priming phenomenon which refute the hypothesis of inhibition (see Park & Kanwisher, 1994), the theory of cue-independent forgetting provides a fine interpretation of the experimental results of Anderson and Spellman (1995) and a new explanation for several other experimental findings, well-known from the literature of interference.

The inhibition of target memories can occur in several different ways, of which Anderson and Spellman (1995) consider two explanations in depth, lateral inhibition and pattern suppression, and they prefer the latter one. In lateral inhibition, with an analogy to neural networks, target memories activated by cues send an inhibitory signal to other target items closely associated to them and to the cues, thereby preventing a spreading activation within the network which would cause an intolerable degree of interference in the system (Estes, 1972; McClelland & Rumelhart, 1981). The concept of lateral inhibition helps to explain both within-category and cross-category impairments occurring in the retrieval-induced forgetting paradigm (Anderson & Spellman, 1995). The theory of Anderson and his colleagues was widely criticised from another aspect, too. According to this criticism, the idea that a cue-independent inhibition underlies the phenomenon of directed forgetting does not seem to be properly supported. There are many properties of retrieval-induced forgetting which support the assumption that inhibitory effects in this paradigm are based on the cue-item relationship. An important feature of Anderson and his colleagues' procedure is that in the practice phase retrieval is necessary for inducing impairment in the recall of related nonrepeated items, the mere re-exposition of items is not enough (Ciranni & Shimamura, 1999; Anderson, Bjork & Bjork, 2000). This finding again supports the idea that retrieval-induced forgetting is the result of the competition of exemplars associated to the same retrieval cue. Above this, Tim Perfect and his colleagues have recently found that repeated retrieval of practised items without their cues does not produce forgetting in the case of related items from the same category (Perfect et al., 2001). The main empirical evidence of the cue-independent inhibitory theory came from the experimental procedure called independent probe technique by Anderson and Spellman (1995). They used a special set of materials that contained related categories with similar items. Each category contained six exemplars, three of them in addition to being members of their own study category, were also associated to another study category. In a series of experiments Anderson and Spellman found that retrieval-induced forgetting impaired not only Rp- items but those Nrp items as well that were similar to practised exemplars ("similar Nrp items"). The similar Nrp items were studied and tested under cues different from those used to practise Rp+ items. According to Anderson and Spellman (1995), this supports the idea that retrieval-induced forgetting is the consequence of the inhibition of the exemplars themselves and this inhibition is independent from their study or test cues. However, this argument is highly questionable on the ground of the existing experimental data - for example it is possible that during the retrieval practice phase similar Nrp words are recoded as members of the practised categories thus becoming actually Rp- words. Another problem is that Williams and Zacks (2001) could not replicate Anderson and Spellman's results of cue-independent inhibition. They used exactly the same experimental procedure and identical material as Anderson and Spellman (1995), and their results showed that retrieval practice did not impair the recall of related Nrp words. Considering that the phenomenon of cross-category impairment is the most powerful evidence for the cue-independent inhibitory theory, the result of Williams and Zacks rather questions the appropriateness of this hypothesis.

There are other problems, however, with the hypothesis of cue-independent inhibition, for example that it pays no sufficient attention to the fact that the inhibition occurring in the selective retrieval practice paradigm is basically of episodic nature. Although Anderson and his colleagues emphasise in most of their studies that inhibition may appear in episodic representations as well, a theory that treats the RIF phenomenon as a suppression of target memories based on semantic features, completely independent of the episodic context, can hardly be called an episodic one. The episodic nature of the RIF phenomenon can be well demonstrated by the following retrieval-induced forgetting experiment. A problematic point of the original paradigm of Anderson et al. (1994) is that inhibition is influenced by the previous semantic knowledge of subjects about category-exemplar relationships, while Macrae and MacLeod (1999, Experiment 1) were able to elicit a RIF effect in completely arbitrary episodic category-exemplar relationships. In the study phase of their experiment subjects had to associate various personality traits (e.g. trustful, sensible, tolerant etc.) to two names (John and Bill), and in the practice phase they practised half of the traits associated to one of the names, in a similar way and to a similar degree as in the paradigm of Anderson et al. (1994). In the final recall phase a significant RIF effect appeared, as subjects recalled a significantly smaller amount of the unpractised traits associated to the practised name than of the traits associated to the unpractised name. Macrae and MacLeod (1999) presume that this mechanism may possibly underlie prejudiced, stereotype thinking, when some (usually unfavourable) personality traits are more readily retrieved concerning certain persons or groups and this impedes the recall of other (favourable) traits. Looking at this finding from the aspect of Anderson and Spellman's (1995) pattern suppression theory we have to assume that unpractised traits related to the same cue (person) will try to intrude during the practice phase, so a control process has to decrease their activation in order to reduce interference. As a consequence, these traits will be difficult to recall later, not only in connection with that particular cue (the particular person, in this case) but in connection with all other cues. Regarding stereotype thinking this would mean that the favourable traits we are unable to recall in connection with certain persons or groups are just as inaccessible for us in connection with everybody else, too, even with people or groups close to us. Obviously, this would result in a rather unevenly operating cognitive system, so inhibition of traits makes sense only together with the inhibition of their episodic cue.

Importantly, research on RIF has shown that only retrieval practice and not restudying of items leads to decreased performance on unpracticed items from the same set (Anderson, Bjork, & Bjork, 2000; Ciranni & Shimamura, 1999; Bäuml, 2002; Staudigl, Hanslmayr, & Bäuml, 2010; Bäuml & Aslan, 2004; but see Verde, 2009), although both types of practicing could strengthen the cue-target associations equally. The most influential theory of RIF – the inhibitory control based accounts – posit that when participants practice retrieval of half of the members from a given category, the other half would compete for retrieval (Anderson & Bell, 2001; Anderson et al., 1994; Anderson et al., 2000; Anderson & McCulloch, 1999; Bäuml & Hartinger, 2002; Storm, Bjork, Bjork, Nestojko, 2006; Storm &

Nestojko, 2010). This competition is then resolved by executive control guided active inhibition, which renders the memories of competitors less accessible for later recall (Anderson, 2005; Anderson & Levy, 2007). Interference based accounts explain RIF without inhibition (Camp, Pecher, & Schmidt, 2007; Camp, Pecher, Schmidt, & Zeelenberg, 2009; Jakab & Raaijmakers, 2009). These models assume that strengthening some category-member associations is enough to lead to interference at any later attempt to retrieve competitors. Here, it is this interference at final recall that leads to RIF. In sum, interference based accounts assume that RIF is the consequence of a sampling failure, i.e., a bias in relative associative strengths, whereas inhibitory models assume that RIF occurs due to recovery failure, i.e., due to a decreased item strength.

Studies on retrieval-induced forgetting typically used categorized lists as learning sets (LS) and applied interim retrieval practice sessions during which some of the items from certain categories received repeated retrieval practice while other category members were not practiced (Storm, 2011). Most of the studies in this literature put only a few minutes delay between practice and final recall (Anderson and Bell, 2001; Anderson et al., 1994; Anderson, Bjork, Bjork, 2000; Anderson and McCulloch, 1999; Bäuml and Hartinger, 2002; Racsmány and Conway, 2006; Storm, Bjork, Bjork, Nestojko, 2006; Storm and Nestojko, 2010). However, two recent studies applied longer practice-recall delays (Racsmány, Conway, Demeter, 2010; Abel et al. 2012). They found that following longer delays between practice and final recall (12 hours or a day), the ratio of recall probabilities of practiced and unpracticed items within the LS was the same than in experimental conditions using shorter delays (Racsmány, et al., 2010; Abel et al. 2012). Note that the ratio of recall probabilities of practiced and practiced items within the practiced LS could be independent from the presence of retrieval-induced forgetting (RIF), as this later phenomenon is measured as the difference between recall of unpracticed items from the practiced LS and the recall of elements from unpracticed LSs (Macrae and Macleod, 1998; MacLeod and Hulbert, 2011). A possible explanation for this is that the recall success of the entire practiced LS could be higher than that of the baseline LS, when there is a longer delay between practice and final test (Racsmány and Keresztes, 2015). In other words, the forgetting rate is higher for the unpracticed LS than for the practiced LS, as a consequence, the nominal recall percent of unpracticed items of the practiced and unpracticed LSs will be different after short delay (RIF present), and will be the same following a longer delay (lack of RIF), although the ratio of recall probabilities for practiced and unpracticed items within the practiced LS is similar or even the same after short and long-term delays (Bauml & Schlichting; 2014). This could be one reason why RIF is usually regarded as a short-term phenomenon (Macrae and Macleod, 1998; MacLeod and Hulbert, 2011). Here we argue that the proper way of measuring the long-term effects of repeated practice is not the comparison of unpracticed items from practiced and unpracticed LSs, rather to measure the ratio of recall probabilities within the practiced LS before and after practice.

I.3. Stopping retrieval: the Think/No-Think Task (TNT)

Anderson and Greene (2001) have recently provided evidence that a phenomenon similar to retrieval-induced forgetting can be elicited in an intentional forgetting task, as well. The so-called Think/No Think paradigm is actually a mixture of the A-B, A-D learning situation and the directed forgetting procedure. Subjects learn loosely associated word-pairs (e.g. ORDEAL-ROACH), then they are trained to provide the second word as a response when they are given the first word as a cue. Next, they participate in the Think/No Think phase, when the first words are given as cues, but together with the cues a signal is provided, too, indicating whether they should remember the second word or try not to think of it. Word-pairs were tested for recall/avoiding between 1 and 16 times. The results of Anderson and Green (2001) show that the more times subjects recalled the second halves of the word-pairs in this phase, the more probably they could recall them in the final recall phase. On the other hand, the more times they were instructed not to think of these words, the worse their final recall performance was compared to the baseline, where there was only a delay phase instead of the Think/No Think phase. Anderson and Green (2001) not only demonstrated that the more times the retrieval of a learned item is intentionally inhibited, the more difficult it would be later to access that item, but also that this retrieval impairment is independent of the cue given during the learning phase. This idea is supported by the finding that words instructed to ignore in the Think/No Think phase, were not only difficult to retrieve with the help of the first word (the episodic cue) but also with such closely associated but unlearned words that would otherwise activate the semantic memory trace of these items (e.g. INSECT-R______ for ROACH). Thus, in Anderson and his colleagues' opinion these experimental results can be generalised to explain other experimental and clinical phenomena because it seems that the temporary, context-independent inhibition of target memories can be elicited in several situations other than the original retrieval-practice paradigm (Anderson et al., 1994). In the following section some clinical data will be presented which reflect the consequences of the impairment of the executive inhibitory control mechanism, described by Anderson and his colleagues.

The most influential family of theories – the inhibitory control based accounts – posit that when participants practice retrieval of half of the members from a given category, the other half would compete for retrieval (Anderson & Bell, 2001; Anderson et al., 1994; Anderson et al., 2000; Anderson & McCulloch, 1999; Bäuml & Hartinger, 2002; Storm, Bjork, Bjork, Nestojko, 2006; Storm & Nestojko, 2010). This competition is then resolved by executive control guided active inhibition, which renders the memories of competitors less accessible for later recall (Anderson, 2005; Anderson & Levy, 2007; Anderson, 2003; Anderson & Bell, 2001; Anderson, Bunce, & Barbas, 2014; Anderson & Hanslmayer, 2014; Anderson & Levy, 2007; 2011).

Interference based accounts – the second family of theories – explain RIF without inhibition (Camp, Pecher, & Schmidt, 2007; Camp, Pecher, Schmidt, & Zeelenberg, 2009; Jakab & Raaijmakers, 2009). These models assume that strengthening some category-member associations is enough to lead to interference at any later attempt to retrieve competitors. Here, it is this interference at final recall that leads to RIF. The most influential of these models, the Search of Associative Memory (SAM) model (Raaijmakers & Shiffrin, 1981) assumes that retrieval occurs in two steps. First – in the *sampling* phase – cues are assembled into a short-term store for activated memory sets, and items are sampled into these sets based on the relative strength of their associations to the given cue. In a second step – the *recovery* phase – sampled items are retrieved based on the absolute strength of their associations to the given cue. It is only a successful recovery that leads to conscious retrieval of a memory item. Using these terms, interference based accounts assume that RIF is the consequence of a sampling failure, i.e., a bias in relative associative strengths, whereas inhibitory models assume that RIF occurs due to recovery failure, i.e., due to a decreased item strength.

The third family of theories pinpoint episodic retrieval as the source of RIF, suggesting that selective retrieval creates and reshapes highly contextualized episodic memory representations (Conway, 2005; 2009; Jonker, Seli, and MacLeod, 2013; Karpicke, Lehman, & Aue, 2014; Racsmány & Conway, 2006; Racsmány, Conway, Keresztes, & Krajcsi, 2012). One line of these theories assume that episodic memory sets contain context, cue, and item features (Conway, 2009; Racsmány & Conway, 2006). In this framework, selective episodic retrieval of a studied memory set transcribes the contextual features and the current ratios of cue-item associations of the learnt memory set into a constrained episodic representation, and RIF occurs whenever these association strengths are reestablished through reinstatement of contextual episodic memory sets of the latest retrieval phase (Racsmány & Conway, 2006; Racsmány et al., 2010). Another line of these theories emphasize the role of context shift between studying a memory set and selective retrieval of parts of this set (Jonker et al., 2013). According to these theories, the mental context of the study phase is changed due to retrieval processes activated during the selective retrieval phase (Sahakyan & Hendricks, 2012), and remains the same throughout the rest of these experiments; therefore RIF is found because the mental context of the final recall is biased to mimic retrieval patterns of the previous selective retrieval.

1.4. Retrieval and long-term facilitation: the testing effect

The classical view on human learning treated memories as formed during studying, and testing as an assessment of the efficiency of studying (e.g. Crowder, 1976). However, a novel research approach has shown that testing is a strong memory enhancer, and could be more beneficial for long-term retention than restudying (Karpicke and Roediger, 2008; 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 one 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. Presently, there is no widely accepted theoretical account of the testing effect. Here, we discuss two popular theories that have been raised in recent discussions. Both theories stressed the role of retrieval cues in its explanation of the testing effect. However, they significantly differ in the specific role they have postulated for cue-related processing.

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 Smith, 2012; Karpicke and Zaromb, 2010; Karpicke, 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 & Blunt, 2011; Karpicke & Zaromb, 2010; 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. According to the set constraining hypothesis, retrieval prompts a process which narrows the cue-related semantic network, 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). Retrieval is accompanied by a cognitive state frequently termed retrieval mode (Tulving, 2002), and produces the opposite process of spreading activation, a network narrowing process involving discrimination between stimuli (Kahana et al., 2008). The act of study and retrieval differ not solely in their background processes, but also in their goal of cue processing. Whereas the task during retrieval is to constrain the search set to a limited size suitable for the reconstruction of the targeted knowledge, the goal of study is to elaborate the cue through activating a large associated semantic network (Karpicke & Blunt, 2011).

II. Theses of the dissertation

The dissertation includes 10 published papers presenting the results of 29 experiments. The following theses mainly based on the results and the conclusion of these publications.

Thesis 1: The concept of 'episodic inhibition' proposes that knowledge in episodic memories preserves a pattern of activation/inhibition derived from the original experience or generated in it by subsequent access of memory details. Thus, an item inhibited in episodic memory may nonetheless be activated in a conceptual knowledge structure.

Thesis 2: Here it is proposed that target items following an intentional forgetting procedure represented in an episodic memory of the study phase are marked as to-be-forgotten, and these episodic representations are specifically tagged not to be recollectively experienced.

Thesis 3: If people observe another person with the same intention to learn, and see that this person is instructed to forget previously studied information, then they will produce the same intentional forgetting effect as the person they observed. This seems to be an important aspect of human learning: if we can understand the goal of an observed person and this is in line with our behavioural goals then our learning performance will mirror the learning performance of the model. Our results support the assumption that suppression of episodic memories is not automatically generated by environmental cues but depends on the goals of the person who encodes and retrieves them.

Thesis 4: Our results indicate that possible disrupted executive functions (e.g. in schizophrenia) considerably weaken the ability of patients to intentionally avoid recent memories. This can occur even when other incidentally initiated inhibitory processes appear to function relatively normally.

Thesis 5: Retrieving memories does not induce forgetting of related memories among participants with obsessive-compulsive disorder (OCD). Lack of forgetting in OCD occurred in spite of the fact that overall memory and the mnemonic effect of practicing memories was almost identical to that among healthy controls. Our results suggest that suppression of irrelevant, interfering memories during competitive recall is impaired in OCD. The lack of retrieval-induced forgetting (RIF) among OCD patients is not related to overall recall performance, rather, we suggest that it is related to differences in resolving interference during competitive retrieval.

Thesis 6: It is proposed that RIF occurs only when interference during competitive retrieval reaches moderate levels, but not when it is too low or too high. This proposal indicates that low levels of interference do not trigger interference resolution, whereas interference resolution can fail when the interference reaches extremely high levels.

Thesis 7: An initial retrieval of the learning set shields against the adverse effect of retrieval practice; RIF is absent either when measured in comparison to baseline performance on the initial retrieval or to members of unpracticed categories. Here it is proposed that retrieval is the key process that enhances long-term accessibility of retrieved memories and it is the process that can hinder retrieval of items through search set restriction or can shield against the adverse effect of later selective retrieval.

Thesis 8: Here a revised form of the episodic inhibition account is proposed: retrieval practice establishes a pattern of activation and inhibition over the contents or features of an episodic memory of the study phase. As the episodic memory is consolidated in long-term memory, the pattern of activation and inhibition, which determines the accessibility of the contents of the memory, stabilizes and becomes resistant to further change. As a memory is repeatedly retrieved and its contents are accessed, its durability in long-term memory increases, and the accessibility levels of its contents become fixed. Our findings suggest that sleep is important to this process of consolidation. It is proposed that consolidation processes occurring during sleep, and possibly featuring some form of offline rehearsal, mediate these long-term effects of retrieval practice.

Thesis 9: Here it is proposed that recalling two associated items can be simultaneously attenuated or primed depending on how the association is accessed. Furthermore, not thinking about a target item, as compared with thinking about an alternative, can produce the same decrements in cued recall or, sometimes, differences. Our findings suggest that the locus of inhibition in the Think/No-think Task (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.

Thesis 10: Based on the results of a functional neuroimaging study it is proposed that the long-term behavioral advantage of repeated retrieval over repeated study is due to the differential activation of a large network. Specifically, when the retention interval is long, participants cannot effectively process the cue and a large percentage of retrieval attempts fail. Thus, the so-called 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.

III. Applied experimental paradigms in the dissertation

III.1. The list method directed forgetting procedure

List-method directed forgetting refers to the experimental procedure when participants typically learn two lists of items for a later memory test, and following the first list (List 1) and right before the learning of the second list (List 2) they are instructed to forget the first list. The typical result demonstrated in a plethora of publications is the decreased recall of List 1 items (called the cost of the F-instruction) and the increased recall rate of List 2 items (called the benefit of the F-instruction), when the performance is compared to a control condition in which participants instructed to remember both study lists (Bjork, 1989; see also Johnson, 1994; MacLeod, 1998 for detailed reviews of the directed forgetting literature).

III.2. The retrieval practice paradigm

In the retrieval practice paradigm (Anderson et al., 1994) participants study category–member pairs (e.g. animal–tiger, furniture–couch, animal–chicken, etc.); then, in a selective retrieval practice phase, they repeatedly retrieve half of the members from half of the categories (e.g. animal - t...?; labelled as Rp+ items). Typically, final recall administered after a delay reveals that repeated selective retrieval leads to forgetting of related material (e.g. 'animal - c...?'; labelled as Rp- items) compared to unpracticed baseline categories (e.g. furniture - c...?; labelled as Nrp items) – this effect is referred to as retrieval-induced forgetting (RIF).

III.3. The Think/No-Think Task

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.

III.4. Retest vs. restudy learning paradigm

This memory task usually consists of three main phases: an initial learning phase, a practice phase, and a final test phase (see Roediger and Butler 2011 for a detailed review of this paradigm). In the initial learning phase, participants are presented with all word pairs (typically Swahili words as a cue, and words from the native language of the participants as a response) in random order on the computer screen, with the Swahili word on the left and its Hungarian equivalent on the right (40 word pairs in our experiments). Participants are instructed to memorize as many word pairs as possible. Immediately after the initial learning phase, participants practice the word pairs in six cycles (practice phase). Word pairs are randomly assigned into a restudy (20 word pairs) or a retest condition (20 word pairs). Each cycle begins with a restudy or a retest block (the order of the restudy and retest blocks varies randomly across the learning cycles), and each restudy-retest block is followed by a feedback block. In the *restudy* blocks, participants see 20 Swahili words together with their Hungarian meanings in random order. In the *retest* blocks, 20 Swahili words is presented in random order on the computer screen. Participants are instructed to press the space button on a standard keyboard of the computer when the right answer came to their mind. Participants are allowed to type the Hungarian meanings of the Swahili words only when they pressed the space button.

Following a 20 minutes or a 7-day retention interval, participants' memory for all 40 word pairs is tested in the final test phase. Swahili words are presented in random order. Similarly as in the practice phase, participants are instructed to press the space button when the right answer came to their mind. Participants were allowed to type the Hungarian meanings of the Swahili words only when they pressed the space button.

IV. Episodic retrieval and memory suppression effects

As it was detailed before, episodic and context-based accounts suggest mechanisms inherent to episodic retrieval processes to explain the positive and adverse effects of retrieval. For instance, the context-based accounts of RIF and retrieval-enhanced learning (Jonker et al., 2013; Karpicke et al., 2014) emphasize the role of context change between initial study of category-member pairs on the one hand, and selective retrieval and final recall on the other. These accounts predict that an initial retrieval of the entire learning set after the study phase will already have participants change their mental context and later selective retrieval practice will cause no further change in this mental context. As a consequence, the context of the initial retrieval will be the active context at final recall.

The starting point for studies, presented in this dissertation, is the episodic inhibition concept of human forgetting (Racsmány and Conway, 2006). Beyond an emphasis on a passive contextual shift, episodic accounts of forgetting phenomena (Conway, 2009; Racsmány & Conway, 2006; Racsmány et al., 2012) highlight the active role of retrieval processes in creating and reshaping episodic memory representations. According to these accounts, episodic retrieval transcribes current contextual information and cue-item association strength ratios of a learning set into an episodic representation. Whenever the same episodic representation is accessed through episodic cues, the encoded cue-item association strength ratios are reinstated. Initial retrieval of the entire learning set can eliminate the adverse effect of later selective retrieval because it transcribes the entire learning set into an episodic representation. When – in the absence of an initial test – the first retrieval is in the practice phase, then final retrieval using the episodic context of the practice phase restricts the search set to practiced items and some arbitrarily activated competitors, whereas other competitors are not involved into the search set at all. That is why RIF is a long-term phenomenon, if final retrieval can reinstate the context of practice phase RIF will be detected following longer delay (Racsmány et al., 2010, see also Abel & Bäuml, 2012). A critical point, and the purpose of this research program, is that if a poststudy task contains items from the study phase but neither requires access of the memory of the study phase nor automatically cues access, then the pattern of activation/inhibition represented in the memory will not influence performance on the poststudy task.

The larger theoretical frame of the episodic inhibition account is the concept of episodic memory described by originally Tulving as a system (1983, 1985) and later by Conway as a representation (2005, 2009). According to Conway, episodic memory is a representation containing the summary representation of sensory-perceptual processing, recollectively experienced, accessed through episodically relevant (spatial and temporal) cues, and associated with short-term goals. Based on this conceptual framework we assumed that intentional and retrieval-induced forgetting phenomena will be characterized by features of episodic representations. Therefore, suppression effects will only be manifested in memories accessed through episodically relevant cues, in recollectively retrieved experiences and will be shaped by short-term goals of the participants. The following three papers investigated these interrelated aspects of human memory suppression effects.

IV.1. The concept of episodic inhibition (Study 1)

Paper: Episodic inhibition. Mihály Racsmány, Martin A Conway, *Journal of Experimental Psychology: Learning Memory and Cognition* 02/2006; 32(1):44-57. DOI:10.1037/0278-7393.32.1.44

Journal of Experimental Psychology:
Learning, Memory, and Cognition
2006, Vol. 32, No. 1, 44–57

0278-7393/06/$12.00 DOI: 10.1037/0278-7393.32.1.44

Episodic Inhibition

Mihály Racsmány
Hungarian Academy of Sciences and University of Szeged

Martin A. Conway
University of Leeds

Six experiments examined the proposal that an item of long-term knowledge can be simultaneously inhibited and activated. In 2 directed forgetting experiments items to-be-forgotten were found to be inhibited in list-cued recall but activated in lexical decision tasks. In 3 retrieval practice experiments, unpracticed items from practiced categories were found to be inhibited in category-cued recall but were primed in lexical decision. If, however, the primes and targets in lexical decision were taken directly from the study list, inhibition was observed. Finally, it was found that when items highly associated with a study list were processed in between study and test, no inhibition in recall was present. These, and a broad range of other findings, can be explained by the concept of "episodic inhibition," which proposes that episodic memories retain copies of semantic knowledge structures that preserve patterns of activation/inhibition originally generated in those structures during encoding.

Keywords: inhibition, retrieval, practice, cued recall, lexical decision, forgetting

An emerging and important finding in the study of inhibitory processes in human memory is that manipulations that apparently induce inhibition in explicit remembering do not have the same effect when memory is assessed implicitly. This was originally observed by Bjork and Bjork (1996, but see also Basden, Basden, & Gargano, 1993), who conducted a list-method directed-forgetting experiment in which participants were instructed to forget the first list learned (TBF items) prior to learning a to-be-remembered (TBR) second list. In a novel manipulation, a word fragment completion test, which included TBF and TBR items, was interposed between study and free recall. Although a standard directed forgetting effect was observed in free recall, there was no directed forgetting effect in word fragment completion. In order to explain this unusual finding, Bjork and Bjork (1996) suggested that the word fragment completion test could be completed by accessing long-term memory conceptual/semantic or lexical representations of the to-be-completed word fragments. Because there was no directed forgetting effect in word fragment completion, it follows that if word fragments were completed by accessing conceptual/lexical representations, then no inhibition would have been present in those representations and, hence, there would be no effect of the directed forgetting instruction in fragment completion. In contrast, the free-recall test explicitly requires access of a memory of the episode in which the word lists were learned. However, the episodic memory of the TBF list is inhibited by the forget instruction and consequently cannot be easily accessed, leading to impaired memory performance. According to Bjork and Bjork (1996), "the inhibition involved in the directed-forgetting situation appears to be a type of retrieval inhibition that impairs conscious access to the original learning episodes" (p. 192). In the experiments below, we systematically explore retrieval inhibition in directed forgetting and retrieval-induced forgetting experiments using both explicit and implicit tests of memory. First, however, we introduce a modification to the notion of retrieval inhibition.

Retrieval inhibition proposes that episodic memories are inhibited. A slightly different version of this is that rather than memories being inhibited it is their contents that are inhibited. It is, after all, the case that at least some items from the TBF list are always recalled, and no one forgets that there were in fact two lists—even patients with quite severe brain damage show this pattern (Conway & Fthenaki, 2003). It is not then as though the TBF list has been rendered wholly inaccessible and, given that it can be accessed apparently completely normally in a recognition rather than a free-recall test (see, e.g., Conway, Harries, Noyes, Racsmány, & Frankish, 2000), its accessibility is clearly not severely compromised. Instead, we suggest that the effect of a forget instruction on an episodic memory of a list of items recently and normally acquired is to impose a pattern of activation/inhibition over the contents or features of the memory. In order to distinguish this view from that of retrieval inhibition we refer to it here as *episodic inhibition*. Episodic inhibition emphasizes the idea that for every episodic memory there is a pattern of activation/inhibition over the contents of the memory, and this strongly influences access to specific features of the content, that is, representations of words in a memory of a recently acquired word list. The pattern of activation/inhibition over the features of an episodic memory initially

Mihály Racsmány, Research Group on Neuropsychology and Psycholinguistics, Hungarian Academy of Sciences, and Department of Psychology, University of Szeged; Martin A. Conway, The Leeds Memory Group, Institute of Psychological Sciences, University of Leeds.

Financial support for the research was provided by Országos Tudományos Kutatási Alapprogramok (Hungarian National Science Foundation) Grant No. F046571 awarded to Mihály Racsmány and Nemzeti Kutatási és Fejlesztési Pályázatok (National Research and Development Programs) Hungarian National Research Grant No. 02/05/0079 for the project "Cognitive and Neural Plasticity." Mihály Racsmány is a grantee of the Bolyai János Research Scholarship of the Hungarian Academy of Science. Martin A. Conway is supported by a professorial fellowship, RES-051-27-0127, from the Economic and Social Research Council of Great Britain.

Correspondence concerning this article should be addressed to Martin A. Conway, The Leeds Memory Group, Institute of Psychological Sciences, University of Leeds, Leeds, LS2 9JT England. E-mail: m.a.conway@leeds.ac.uk

reflects processing that occurred during encoding but can be changed by subsequent access of the memory and processing of its content. Later we show how this concept of episodic inhibition might be used to provide a common basis for understanding attenuation of memory in directed forgetting and retrieval practice.

One implication of the notion of episodic inhibition, which derives from Bjork and Bjork (1996) and which we also emphasize here, is that the effect of inhibition on the contents of episodic memories is long lasting. In contrast, the effects of inhibition on other types of long-term knowledge representations may be less enduring and more transitory (see Neely, 1991, for a review). Thus, for example, patterns of activation/inhibition over conceptual, lexical, and perhaps other types of representations, generated for example while words on a list are read, will dissipate in periods measured in seconds and milliseconds. Because these patterns of activation/inhibition are rapidly changing they are unlikely to influence performance on a memory test given some time (often minutes) later. In contrast, it is suggested that representations of items in episodic memories, which are themselves derived from conceptual, lexical, and other types of processing present during encoding, maintain the patterns of activation/inhibition that characterized the epoch an episodic memory represents, (cf. Conway, 2001). Indeed, one possibility is that the patterns of activation/inhibition present over features in an episodic memory will remain unchanged until the contents of the memory are accessed and subjected to further processing (see MacLeod & MaCrae, 2001, for highly relevant findings, and Tipper, 2001, and Tipper, Grison, & Kessler, 2003, for related findings from the study of attention).

Our account of episodic inhibition makes a strong claim, namely, that the same representation (item) can be processed independently according to whether it is accessed in conceptual, lexical, or other knowledge structures or in an episodic memory (see too Perfect, Moulin, Conway, & Perry, 2002). An episodic memory, however, preserves a pattern of activation/inhibition from a previous processing episode whereas other knowledge structures, in which the original pattern of activation/inhibition was first established, do not. According to this reasoning a particular pattern of activation/inhibition will be detected when an episodic memory of an item is accessed. However, when a conceptual, lexical, or other representation of the *same* item is accessed, a different pattern of activation/inhibition will be observed. A representation may then be both inhibited (in an episodic memory) while being noninhibited or even activated in conceptual, lexical, or other knowledge structures. It is this prediction of episodic inhibition that is the main focus of the series of experiments reported below, which investigate the phenomenon first in directed forgetting (Experiments 1 and 2), next in retrieval practice (Experiments 3 through 5), and finally in a novel study suggested by the earlier experiments (Experiment 6).

Experiment 1

The present experiment and Experiment 2 both used a directed forgetting by lists procedure. In this procedure participants learn a list of words. Halfway through the list they receive a mid-list instruction. For half the participants—the F group—this is an instruction to forget the words they have learned thus far and instead to concentrate on the upcoming words, which will have to be recalled. The other half—the R group—are instructed to keep remembering the words they have just studied and to learn the next set of words that will have to be recalled. In this procedure the directed forgetting effect consists of poorer recall for List 1 by the F group relative to their List 2 performance and to the performance of the R group for List 1 (see Conway et al., 2000, for further discussion of this particular pattern of directed forgetting). One current view is that the directed forgetting effect (at least in the lists method) is due to inhibition of the List 1 TBF items in the F group triggered by the intention to forget and by learning List 2 (Bjork, 1989; Bjork, Bjork, & Anderson, 1998; Conway et al., 2000). Other accounts in terms of, for instance, selective rehearsal have not received empirical support (Geiselman, Bjork, & Fishman, 1983; Geiselman, & Bagheri, 1985) and it is also acknowledged that the inhibitory account may not extend to other forms of directed forgetting, that is, by items rather than by lists (see Basden & Basden, 1998, and MacLeod, 1998, for reviews).

The novel procedure introduced here is to interpose an apparently unrelated lexical decision test between study and test. This is a test that includes all items from the study phase in the context of new nonstudied filler words and a matching set of nonwords. A clear prediction of the episodic inhibition view detailed earlier is that performance decrements should be present for F group List 1 items in free recall, but these may not necessarily be present for the same items in lexical decision times. Indeed, episodic inhibition predicts that performance decrements of inhibited List 1 items will only be present if the lexical-decision task is mediated by an F group episodic memory of List 1. If, however, lexical decisions are mediated by lexical and conceptual representations of List 1 items, which do not themselves preserve the inhibition induced by the directed forgetting procedure, then no slowing of lexical decision times should be observed. There is some evidence both in support of this prediction and against it. Against the prediction are findings by MacLeod (1989; see also Fleck, Berch, Shear, & Strakowski, 2001) showing that lexical decision times were slowed for F items in an item-by-item directed forgetting procedure, that is, when the F and R instructions followed presentation of each individual word. However, as directed forgetting effects in item-by-item procedures are thought to reflect changes in rehearsal strategies rather than inhibitory processes (Basden & Basden, 1996), there is no reason why episodic inhibition should provide an account of these particular effects. In contrast, experiments involving the list-directed forgetting procedure have revealed that on a range of implicit tasks (none of which were lexical decision tasks) interposed between study and test, there are often no effects of directed forgetting despite a reliable effect in free recall (Bjork & Bjork, 1996; Perfect et al., 2002). It is this pattern that is predicted by episodic inhibition and that is assessed in the present experiment.

Method

Participants. The participants were 32 undergraduate Hungarian students from the University of Szeged, who participated in return for partial credit in a lower division psychology course. Their age varied between 18 and 24 years. There were 20 women and 12 men

Procedure. Participants were tested individually and were informed that they were participating in an experiment on memory that would test their ability to recall words. The experiment was conducted in four phases: a list learning phase, a distractor phase, a lexical decision phase, and a free-recall phase. Words were presented visually on a computer screen. Each word was displayed for 2 s with a 2-s inter-item interval. After the words of the first list (12 words) had been presented, participants were instructed to stop. At this point, participants in the F group were given the forget instruction and those in the R group were given the R instruction. For the F instruction, the experimenter gave spoken instruction that the previous presentations had been a practice list to familiarize the participants with the method and stimuli and that they should now forget the words they had just studied, put them out of mind, and concentrate on the upcoming experimental list, which they would have to remember. For the remember instruction, also spoken, R-group participants were informed that they had now completed studying a first list, and this was to be followed by a second list that also had to be remembered later in the experiment. Allocation to groups was random. After all words had been studied, participants were given a 5-min arithmetic distractor task. The distractor task was followed by a lexical-decision task. The experimental lists were randomly selected from four study lists, each of which contained 12 high-frequency words naming common objects (see Racsmány, 2003, for the full lists).

The design of the lexical-decision task was the same as that used by MacLeod, (1989). There were 15 practice trials, made up of seven words and eight nonwords not included in the experimental sets. Each trial began with a 250-ms warning ****, followed by a 250-ms blank period prior to the item. Each item was presented in uppercase letters at the center of the screen either until the participant pressed a key to indicate the chosen response or for a maximum of 2 s. There was a 250-ms blank period before the next warning stimulus. The 96 experimental trials were made up of 24 studied words (List 1 and List 2 words), 24 unstudied words, and 48 nonwords. Participants were encouraged to respond as rapidly as possible and at the same time to avoid errors. After the lexical-decision task was completed, participants took part in a free-recall task. For this, they were given a sheet of paper and were instructed to try to recall any words they could first from the first list, then from the second list. The forced order of recall served to eliminate output interference from the second list.

Results and Discussion

Lexical-decision task. There were fewer than 1% errors and most participants made no errors. There was no systematic distribution of errors to conditions and the few errors were, for the purposes of analysis, replaced by the mean for that participant in that condition. We conducted a 2 (group) $\times$ 3 (words) mixed analysis of variance (ANOVA) on the reaction time data; the reaction times of nonword items were not included in the analyses because they were not pertinent to the predictions. The main effect of group was not significant ($F < 1$), whereas the main effect of words was significant $F(1, 60) = 18.20$, $p < .01$. Studied words were reliably responded to more quickly than unstudied words (see Table 1), and there was no significant Group $\times$ Words interaction, $F(1, 60) = 0.98$, $p > .10$. Planned comparisons between F-group Lists 1 and 2, and between F-group List 1 and R-group List 1, showed no reliable differences.

Free-recall performance. The main effect of group was not significant $F(1, 30) = 1.84$, nor was the main effect of List ($F < 1$). There was, however, a significant Group $\times$ List interaction, $F(1, 30) = 16.10$, $p < .01$. Planned comparisons confirmed that the recall of List 1 words was significantly lower than the recall of List 2 words in the F group, $F(1, 15) = 8.27$, $p < .01$. The recall performance of List 1 words in the F group was significantly poorer than the recall of List 1 words in the R group, $F(1, 30) = 13.57$, $p < .01$ (see the lower section of Table 1). Thus, a powerful directed-forgetting effect was present in free recall, but this effect was absent in lexical decision.

The findings of this first experiment are then highly consistent with the predictions of the episodic inhibition account. By this view, representations of the List 1 items in the F group are in an inhibited state in an episodic memory of learning the list. When this episodic memory and its contents are accessed, during the list-cued-recall test, relatively few items from List 1 for the F group can be accessed because their representations are inhibited. This inhibition occurs after the memory has been constructed and representations of the items copied into it. Just prior to presentation of the F instruction, the items would presumably be highly activated and accessible, but the F instruction and the second list learning trigger inhibition of the contents of the memory. It is perhaps important to note that no participant failed to recall that there had been a first list, indicating that the memory itself was not in a state of lowered accessibility. An alternative to the episodic inhibition account of the present experiment might focus on the fact that lexical decision preceded free recall and, perhaps, it is this

Table 1
Mean Lexical Decision Times and Percentage Recalled in Experiment 1

	Type of target words			
Group	List 1 words	List 2 words	New words	Nonwords
	Mean latencies in the lexical decision task			
F group	581.2 (54.1)	586.9 (67.2)	625.7 (79.5)	630.2 (80.1)
R group	576.4 (70.4)	574.2 (60.3)	611.8 (71.4)	627.2 (70.3)
	Mean percentage recall of List 1 and List 2			
	List 1		List 2	
	M	*SD*	*M*	*SD*
F group	24.2	13.2	36.5	15.1
R group	40.1	20.0	30.9	22.9

Note. F group refers to those participants told to forget the words they have learned; R group refers to the participants told to remember the words they have learned. Latencies are presented in milliseconds.

fixed order that is in some way influencing the findings. In the next experiment, we tested this alternative by giving an additional free-recall test just before the lexical-decision task. This strategy should ensure that to-be-forgotten items would be under active suppression at the beginning of the reaction time task and thus maximize the conditions for a directed-forgetting effect in lexical decision.

Experiment 2

Method

The participants were 42 undergraduate Hungarian students from the University of Szeged, who participated in return for partial credit in a lower division psychology course. Their ages varied between 18 and 22 years, with the mean age being 20.1 years. There were 26 women and 16 men. Experiment 2 followed the same procedure as Experiment 1, with one difference. The experiment was conducted in five phases: a list learning phase, a distractor phase, the first list-cued recall phase, a lexical decision phase, and a (second) list-cued recall phase.

Results and Discussion

First free-recall performance. A 2 × 2 (Lists × Group) mixed-factor ANOVA was conducted on the number of words recalled by each subject in the first free-recall phase. Table 2 (upper section) shows the mean probabilities for groups and lists. To compare the critical differences between means, we performed a series of planned comparisons on pairs of means, as in Experiment 1. The main effect of groups was significant, $F(1, 40) = 5.49$, $p < .05$. The main effect of list was not significant ($F < 1$). The analysis of recall scores yielded a significant Group × List interaction, $F(1, 40) = 27.14$, $p < .01$. Planned comparisons confirmed that the recall of List 1 words was significantly lower than the recall of List 2 words in the F group, $F(1, 20) = 15.42$, $p < .01$. The recall performance of List 1 words in the F group was significantly poorer than the recall performance of List 1 words in the R group, $F(1, 40) = 23.70$, $p < .01$. These results, the List × Group interaction, and the critical contrast of F-group List 1 versus R-group List 1 demonstrate a robust directed forgetting effect.

Lexical-decision task. We conducted a 2 (group) × 3 (words) mixed ANOVA on the reaction time data; the reaction times of nonword items were not included in the analyses. The main effect of group was not significant ($F < 1$), the main effect of words was significant, $F(1, 80) = 16.80$, $p < .01$, and overall studied words had faster lexical decision times than new words (see the middle section of Table 2). There was no significant Group × Words interaction ($F < 1$).

Second free-recall performance. The main effects of group and lists were not significant. However, as in the first recall phase, a highly reliable Group × List interaction, $F(1, 40) = 14.60$, $p < .01$, was observed (see the lower section of Table 2). Planned comparisons confirmed that the recall of List 1 words was significantly lower than the recall of List 2 words in the F group, $F(1, 20) = 7.56$, $p < .01$. Recall of List 1 words in the F group was significantly poorer than the recall of List 1 words in the R group, $F(1, 40) = 10.46$, $p < .01$. These results, the List × Group interaction, and the critical contrast of F group List 1 versus R group List 1 indicate that the directed forgetting effect was not released by the lexical-decision task or by the earlier recall. The moderate increase in recall overall in the second recall indicates a weak hyperamnesia effect.

Whether lexical decision precedes or follows recall does not appear to influence the presence of the directed forgetting effect in recall or the lack of it in lexical decision times. Indeed, even when a second recall is undertaken, the pattern of the directed-forgetting effect remains intact. Thus, neither prior recall nor encountering the inhibited items in another processing context was sufficient to overcome the effect. According to the episodic inhibition account, these effects occur because the extended dynamic pattern of activation/inhibition that evolved in perceptual, conceptual, motivational, and affective systems during encoding becomes represented in an episodic memory, or set of such memories, and this pattern determines recall. If, however, items inhibited in the episodic memories are encountered in contexts in which they can be processed without accessing the episodic memories, then no inhibition will be observed. The present findings not only support this view but also show that the episodically inhibited items are in fact primed in lexical decision times. In both Experiments 1 and 2, lexical decision times to studied items, including the critical List 1 F-group items, were quicker than to previously unstudied new words. This paradoxical effect is predicted by episodic inhibition, which proposes that the pattern of activation/inhibition induced by the study phase is preserved only in episodic memories of the study phase. It seems that for conceptual or lexical representations, prior exposure to the words in all lists gave rise to enduring activation or priming.

Table 2
Mean Percentage Recalled and Mean Lexical Decision Times in Experiment 2

	Type of target words			
	List 1		List 2	
Group	*M*	*SD*	*M*	*SD*
Mean percentage recall in the first free-recall task				
F group	19.8	11.1	32.5	12.1
R group	42.1	17.8	27.3	16.3
Mean latencies in the lexical decision task				
	List 1 words	List 2 words	New words	Nonwords
F group	569.6 (62.1)	568.7 (52.8)	610.7 (78.7)	613.2 (84.6)
R group	588.1 (66.3)	584.3 (54.5)	597.3 (79.8)	616.6 (63.5)
Mean percentage recall in the second free-recall task				
	List 1		List 2	
	M	*SD*	*M*	*SD*
F group	26.6	15.8	37.5	16.6
R group	45.8	22.5	34.2	22.4

Note. F group refers to those participants told to forget the words they have learned; R group refers to the participants told to remember the words they have learned. Latencies are presented in milliseconds.

Experiment 3

The episodic inhibition view proposes that by accessing knowledge in an episodic memory, the pattern of activation/inhibition over representations in the memory can be altered for at least some time after encoding. Indeed, this malleability provides an important mechanism for reevaluating memories in response to later experience. An experimental procedure that has extensively examined this is the retrieval practice procedure (Anderson, Bjork, & Bjork, 1994; Anderson & Bell, 2001; Anderson & Spellman, 1995). In the retrieval practice procedure, inhibition is induced by selectively rehearsing a subset of items from a recently learned list. Typically, the first list to be learned consists of several categories, for example, fruits, birds, vehicles, and so forth, and paired with each category are several exemplars, for example, orange, apple, banana. Having studied the list, the participant enters a second phase in which practice is undertaken in the form of cued recall of some of the categories and their exemplars. Items previously learned are recalled to word-stem cues such as "fruit/or_?" In the third phase that follows retrieval practice, an attempt is made to recall all the originally acquired items to the category cues, for example, "fruit: ____, ____, ____?" The typical pattern of findings is that category-cued recall of unpracticed items from categories that were practiced (rp items) is reliably lower than that of items from categories that were not practiced (nrp items) and which constitute the baseline for gauging retrieval-induced forgetting (RIF) in this procedure. Finally, recall of items that were practiced (rp+ items) is reliably greater than that of items from practiced categories that were themselves not retrieval practiced (rp−) and items that were not from categories that contained an item that received retrieval practice (nrp items). Thus, the beneficial effects of rehearsal are shown in the high levels of rp+ recall, and the inhibition of unpracticed highly related category members caused by the practice is reflected in the low recall of the rp− items relative to the nrp items.

There are several explanations of the memory effects in retrieval practice, and we consider these later in the General Discussion. The episodic inhibition approach we have developed here makes much the same predictions for retrieval practice as it did for directed forgetting, namely, that the effects of practice will be to set up a particular pattern of activation in an episodic memory of processing the study list, and it is this pattern that will determine later recall. Thus, the study list items will be represented with varying degrees of accessibility in the episodic memory of learning the list. It seems reasonable to assume that most items will be activated and accessible; after all, the goal set for participants is to learn the study items for a later memory test. Practice of one item from a category in which all the items are at roughly similar levels of activation will have the effect of increasing the activation level of that item (rp+ items) in the memory and perhaps of decreasing and even inhibiting the closely associated items (rp− items). The activation levels of episodic memory representations of items from unpracticed categories (nrp items) will presumably undergo little change, as they are not directly accessed during the practice phase. As with directed forgetting, however, if items that are inhibited in an episodic memory can be processed in a new processing context that does not require or induce access to the episodic memory, then no inhibition and even activation of these inhibited items may be observed. Thus, a lexical-decision task interposed between study- and test-containing words inhibited by retrieval practice, words that cannot be recalled, may show no slowing of lexical decision times and even a speeding of reaction times relative to new previously unstudied words. Such a pattern of findings would generalize our findings in directed forgetting to retrieval practice and provide convergent evidence for the episodic inhibition account.

The present experiment uses the retrieval practice procedure of Anderson et al. (1994), with two modifications. After the category cued recall phase, a lexical-decision task was undertaken in which word/nonword judgments were made of previously studied words, new words, and previously unstudied nonwords. Following this, a second cued-recall test was taken and this was to examine the effect of implicit reexposure on the pattern of cued recall. This sequence of tests will establish whether an RIF effect is (a) present in the first cued recall phase, (b) absent in the lexical-decision task, and (c) present again in the second recall phase. By the episodic inhibition view, the RIF pattern should be present at least in the first recall test and very possibly in the second test as well; this is because both involve access of an episodic memory of the study list in which the RIF pattern of activation/inhibition has been induced by retrieval practice. However, the episodic inhibition account predicts that RIF effects will not be present in the lexical-decision task because this task can be completed by accessing semantic representations that are not inhibited.

Method

Participants. Twenty-five undergraduate Hungarian students from the University of Szeged, Szeged, Hungary, participated in return for partial credit in a lower division psychology course. The students' ages varied between 18 and 24 years, with the mean age being 20.3 years.

Materials. Following Anderson et al. (1994), we constructed 10 categories, 2 of which were used as fillers. Each category contained six examples, the words being drawn from several published Hungarian norms (Füredi & Kelemen, 1989; Kónya & Pintér, 1985), of moderate to high frequency and highly typical members of their category (see Appendix A).

Procedure. Participants were tested individually and were informed that they would be participating in an experiment on memory. The experiment was conducted in six phases: a learning phase, a retrieval-practice phase, a distractor phase, a category-cued recall phase, a lexical decision phase, and a second surprise category-cued recall phase. The learning phase was controlled by a Pentium III personal computer. The participants saw category–exemplar pairs on the monitor screen, which they were to try to remember as best as they could for a later memory test. Each category exemplar pair was presented in uppercase letters at the center of the screen for 5 s. We presented the category–exemplar pairs in an unsystematic intermixed order. When participants completed the learning phase, the experimenter distributed practice booklets. Each page in the booklet contained one of the category names studied in the previous phase of the experiment and the first two letters of one of the members of that category, which they had to complete. They were encouraged not to guess but to retrieve an item studied in the previous phase. Participants were warned that some of the category–exemplar pairs might be repeated and that when this occurred they should again recall the item from the original list. The participants practiced 3 exemplars from half of the 8 learning categories. The practice booklet contained every critical exemplar three times and thus contained 66 category–exemplar stem pairs. The practiced categories were counterbalanced between experimental groups. After the retrieval practice phase had been completed, the booklets were collected, and participants were given an unrelated arithmetic task for 5 min.

In the first recall phase, participants were given recall booklets with the name of one of the categories studied previously at the top of each page. In each booklet, the order of presentation of the categories was random. The participants worked through the 8-page booklet from first page to last, recalling as many previously studied exemplars as they could in the 40 s allocated for each category. The lexical-decision task followed standard practice, and there were 15 practice trials, consisting of 7 words and 8 nonwords not included in the experimental sets. Each trial began with a 250-ms warning of ****, followed by a 250-ms blank screen prior to presentation of the item. Each item was presented in uppercase letters at the center of the screen until the participant pressed one of two keys to indicate the chosen response or for a maximum of 2 s. The "WORD" key was always operated with the right hand and the "NONWORD" key, with the left. There was a 250-ms blank period before the next warning symbol. The 176 experimental trials were made up of 48 studied words (12 rp+, 12 rp−, 24 nrp words), 48 unstudied words, and 80 nonwords. Order of presentation was random. Participants were required to respond as rapidly as possible while avoiding errors. Response time was recorded from item on screen to keypress in milliseconds. After the lexical-decision task had been completed, participants took a second category-cued-recall test following the procedure of the first test but with the order of cues unsystematic with respect to the first test and original learning trial.

Results and Discussion

Category-cued recall. Table 3 shows the percentages of each type of item that was correctly recalled in the first and in the second category-cued recall phase. Following Anderson et al. (1994), retrieval-induced forgetting was assessed by comparing recall performance on unpracticed items from the practiced categories (rp− items) with recall performance on unpracticed items from the previously unpracticed categories (nrp items). If the latter exceeds the former, then retrieval-induced forgetting has occurred. To determine whether this was the case, we conducted a within-subject ANOVA—with item type as the single variable having the three levels rp+, rp−, and nrp—on the raw scores both for the first- and second-recall performances. In the first-recall phase, a reliable effect of item type was found, $F(2, 48) = 76.10$, $p < .01$. Planned comparisons found that recall of rp+ items was significantly higher than that of nrp items, $F(1, 24) = 637.10$, $p < .01$, confirming the benefit of retrieval practice. Recall of rp− items was significantly lower than that of nrp items, $F(1, 53) = 17.70$, $p < .01$, demonstrating retrieval-induced forgetting. For the second-recall phase, conducted after the lexical-decision task, a reliable effect of item type was again found, $F(2, 48) = 23.40$, $p < .01$. Recall of rp+ items was significantly higher than that of nrp items, $F(1, 24) = 1{,}079.90$, $p < .01$, and recall of rp− items was significantly lower than that of nrp items, $F(1, 24) = 21.50$, $p < .01$. The usual RIF pattern of recall in the recall practice procedure was then present in the standard category cued-recall test and in the same test again after an intervening lexical-decision task that featured exactly the same items.

Lexical decision. The lower section of Table 3 shows the mean RTs by item type condition and for new (unstudied) words. Note that although the means for the nonword trials are also shown in Table 3, these were not included in the analyses, as they were not pertinent to the main questions. Thus, a single factor (item type: rp+, rp−, nrp, and unstudied words) within-subject ANOVA was carried out on the latencies. There was a significant effect of item type, $F(3, 72) = 32.80$, $p < .01$. The latency of rp+ words was significantly shorter than that of the nrp words, $F(1, 24) = 3{,}145.60$, $p < .01$, demonstrating a strong priming effect arising from the original study phase and reinforced by subsequent retrieval practice. Note that these two sources of activation, study and retrieval practice, may sum to produce rp+ items that are more strongly activated than any other items in the set. The latency of nrp words was significantly shorter than that of the previously unstudied words $F(1, 24) = 30.69$, $p < .01$, and this demonstrates that the study phase on its own was sufficient to prime lexical-decision times. Critically, however, there was no significant difference between latencies of the nrp words and rp− words ($F < 1$), and this shows that there was no RIF effect in lexical decision times for items inhibited in category cued recall. Thus, the same items can be both inhibited and primed depending on the type of test used to access the items.

Experiment 4

One way in which the lexical-decision task can be extended is into primed lexical decision. This provides an opportunity to explore, albeit in a different task and different manner, the spread of inhibition originally reported by Anderson and Spellman (1995). Thus, in the present experiment, category exemplars were primed in the lexical decision phase with either a studied or unstudied category. For example, if "orange" was studied in the context of "fruit," then in lexical decision, its studied prime would be "fruit" or, if allocated to the unstudied prime condition, its unstudied prime would be "food." Note that, studied and unstudied

Table 3
Mean Percentage Recall and Mean Lexical Decision Times in Experiment 3

Variable	Rp+ items	Rp− items	Nrp items	New word	Nonword
		Mean percentage of items recalled on first- and second-category cued recall			
First recall	72.6	13.9	33.2		
Second recall	72.5	16.5	43.7		
		Mean (and *SD*) lexical decision times			
Latency	568.5 (42.9)	595.9 (70.2)	597.3 (69.6)	679.4 (52.9)	780.3 (74.8)

Note. Rp+ = Items that received retrieval practice; Rp− = items from practiced categories that were themselves not retrieval practiced; Nrp = items that were not from categories that contained an item that received retrieval practice. Lexical decision times are presented in milliseconds.

primes (all category names) were selected to be equally associated to the target exemplar. Collapsed over prime type, we expect to observe the same pattern of findings for lexical decision and category cued recall as that observed in the previous experiment. For studied primes, it is possible that both episodic and conceptual representations may be accessed and, if so, inhibition may be detected in the form of slower latencies for rp− compared with rp items. For unstudied primes, latency may be mediated mainly, or even solely, by conceptual representations and, if so, no reliable differences in latencies between rp− and nrp items should be observed. On the other hand if there was some spread of inhibition, and this was represented in the episodic memory of the study phase (modified by practice) then it is possible that these unstudied category-plus-studied exemplar pairs might access the memory and thus show some inhibition reflected in slow latencies to rp− items.

Method

Participants. Twenty-six undergraduate Hungarian students from the University of Szeged participated in return for partial credit in a lower division psychology course. Their ages varied between 18 and 25 years, with the mean age being 21.3 years.

Procedure. The procedure was identical to Experiment 3 except that the lexical-decision task was replaced by a primed lexical-decision task. In this task, prior to each visually presented target word, a category name appeared as a prime. There were two different types of prime–target trials: previously studied category–word pairs for example, fruit–orange, or unstudied but related category–word pairs, for example, food–orange (see Appendix B). The 176 trials were made up of 48 studied words (12 rp+, 12rp−, 24 nrp words), 48 unstudied words, and 80 nonwords. There were two category label sets; half of the studied words were randomly assigned to studied category primes and the other half, to unstudied category primes. The unstudied (new) words were primed with associated categories but ones that had not been used for studied words. Nonword trials were primed with studied or unstudied category primes. Each trial began with a 250-ms warning of ****; then the prime was shown for 500 ms, followed by a 500-ms blank period, and finally the target word was presented in uppercase letters at the center of the screen until the participant pressed a key to indicate the chosen response or for a maximum of 2 s. There was a 250-ms blank period before the next warning stimulus. As usual, participants responded as quickly and as accurately as possible.

Results and Discussion

Category-cued recall. As in Experiment 3, the number of items recalled in both category-cued recall phases were each analyzed with a single variable (item type: rp+ or rp− or nrp) within-subject ANOVA. Table 4 shows the percentages of each type of item correctly recalled in the first and in the second category-cued recall phase. In the first-recall phase, there was a significant effect of item type on recall, $F(2, 50) = 206.24$, $p < .01$. Planned comparisons found that recall of rp+ items was significantly higher than that of nrp items, $F(1, 25) = 940.28$, $p < .01$, which confirmed the recollective benefits of the retrieval practice. Recall of rp− items was significantly lower than that of nrp items, $F(1, 25) = 63.61$, $p < .01$, demonstrating retrieval-induced forgetting. In the second-recall phase, which was conducted after the lexical-decision task, we again found a significant effect of item type, $F(2, 50) = 160.43$, $p < .01$. The recall of rp+ items was significantly higher than that of nrp items, $F(1, 25) = 869.12$, $p < .01$, and the recall of rp− items was significantly lower than that of nrp items, $F(1, 25) = 38.29$, $p < .01$ (see Table 4). These results demonstrated the presence of retrieval-induced forgetting after a lexical-decision task involving items of the learning phase. Taken together, the present results and those of Experiment 3 demonstrate powerful and extremely robust effects of RIF recall following the retrieval practice procedure.

Primed lexical decision. Table 4 shows the primed lexical decision data. A 2 × 4 within-subject ANOVA was conducted in which the two variables were prime (studied vs. unstudied) and item type (rp+, rp−, nrp, and unstudied words). Both the prime and the item type main effects were significant, $F(1, 25) = 8.35$, $p < .01$, $F(1, 23) = 16.90$, $p < .01$, respectively. The Prime × Item Type interaction was also significant, $F(1, 23) = 3.60$, $p < .03$. Two separate, single-variable within-subject ANOVAs were then conducted on the raw latencies for studied primes and for unstudied primes. When the prime words were studied categories,

Table 4
Mean Percentage Recall and Mean Lexical Decision Times in Experiment 4

Recall group/ type of prime	Rp+	Rp−	Nrp	Overall	New word	Nonword
Mean percentage of items recalled on first and second category cued recall						
	M(%)	*M*(%)	*M*(%)	*M*(%)		
First recall	80.4	12.9	22.2	43.2		
Second recall	83.6	12.8	32.4	50.8		
Mean (and *SD*) lexical decision times						
	Rp+	Rp−	Nrp		New word	Nonword
Studied category	546.8 (55.3)	594.9 (90.6)	568.1 (82.7)		632.5 (71.9)	714.7 (76.3)
Unstudied category	577.8 (85.8)	570.9 (68.1)	603.8 (75.2)		637.2 (72.5)	719.2 (75.8)

Note. Rp+ = items that received retrieval practice; Rp− = items from practiced categories that were themselves not retrieval practiced; Nrp = items that were not from categories that contained an item that received retrieval practice. Lexical decision times are presented in milliseconds.

there was a significant effect of item type, $F(3, 75) = 26.56$, $p < .01$. The latency of rp+ words was significantly shorter than that of the nrp items, $F(1, 25) = 4.23$, $p < .05$, showing a priming effect of retrieval practice (see Table 4). The latencies of nrp words were reliably faster than that of the previously unstudied words, $F(1, 25) = 59.20$, $p < .01$, demonstrating a priming effect of the learning phase on the lexical decision times. These findings are then highly consistent with the findings of Experiment 3. For the critical contrast, however, a reliable difference was found between latencies of nrp words and rp− words, $F(1, 25) = 8.37$, $p < .01$, and it can be seen from Table 4 that rp− latencies were slower than nrp latencies. This constitutes an RIF effect in the lexical-decision task that was absent in Experiment 3, and it is an RIF effect that (only) occurs when the primes are studied categories. When the primes were unstudied but nonetheless highly associated categories, a significant effect of item type was again observed, $F(3, 75) = 13.17$, $p < .01$. The latencies of rp+ words were significantly shorter than those of the nrp words, $F(1, 25) = 7.50$, $p < .01$. The latencies of rp− words, however, were significantly shorter than those of nrp words, $F(1, 25) = 14.50$, $p < .01$, and did not differ reliably from the corresponding latencies for rp+ items. Why this unexpected "rebound" effect occurred is not known, but what is critical for the present argument is that these findings for unstudied primes, unlike the results for the studied primes, show no evidence of inhibition of rp− items. Finally, in all cases, all studied items were responded to more quickly than the new words ($p < .05$), demonstrating a priming effect of similar magnitude over all conditions.

In the present experiment, the pattern of findings for category-cued recall was highly similar to that found in Experiment 3 at both test phases and constitutes a robust RIF effect. The pattern of lexical decision latencies was also highly similar to that observed earlier when this is collapsed over prime type. The lexical decision data by prime type, however, were more complex: When the prime–target pair contained previously studied items, such as fruit–apple, the RIF pattern of inhibition was present in the lexical decision times, but when the prime–target pair consisted of previously unstudied primes paired with previously studied items, such as food–apple, no RIF inhibition was present, and all items were primed relative to new items. These findings suggest that studied prime–targets accessed a representation of the study list that contained the RIF pattern of activation/inhibition. This representation produced the slowing of latencies to rp− items that would be expected if retrieval practice influenced lexical-decision times. In contrast, unstudied primes paired with studied targets did not show the slowing to rp− items that might be expected to occur when the retrieval practice procedure is used. Thus, whatever representation these cues (unstudied primes plus studied targets) are accessing, it cannot be the same as that accessed by the studied primes and targets.

Experiment 5

In this experiment we repeated the preceding experiment but omitted the category-cued recall administered prior to the lexical-decision task. The reason for this is that it may be the case that the cued-recall test has some (undetected) effect on primed lexical decision, and this might be so even if both tasks access different long-term memory representations (a concern that also applied to Experiment 1).

Method

A new group of 32 undergraduate Hungarian students from the University of Szeged participated in return for partial credit in a lower division psychology course. Their ages varied between 18 and 23 years, with the mean age being 20.1 year. The experiment was the same as Experiment 2 in all other respects.

Results and Discussion

Category-cued recall. The analyses were the same as those conducted previously. Table 5 shows the percentages of each type of item that was correctly recalled in the category-cued recall phase, which was conducted after the lexical-decision task. A significant effect of item type, $F(2, 62) = 53.25$, $p < .01$, was found. The recall of rp+ items was significantly higher than that of nrp items, $F(1, 31) = 26.40$, $p < .01$, and the recall of rp− items was significantly lower than that of nrp items, $F(1, 31) = 50.50$, $p < .01$. These results demonstrate again the presence of a powerful and consistent RIF effect after a lexical-decision task involving items from the learning phase.

Primed lexical decision. Table 5 shows the primed mean lexical decision times. Both the prime and the item type main effects

Table 5
Mean Percentages of Items Recalled and Mean Lexical Decision Times in Experiment 5

Recall group/ type of prime	Rp+	Rp−	Nrp	Overall	New word	Nonword
		Mean percentage of items recalled				
	M(%)	*M*(%)	*M*(%)	*M*(%)		
First recall	69.4	17.4	37.4	53.7		
		Mean (and *SD*) lexical decision times				
Studied category	543.8 (54.6)	596.6 (75.4)	568.3 (66.9)		630.1 (54.7)	720.1 (64.9))
Unstudied category	569.3 (85.2)	600.4 (85.1)	606.7 (79.7)		633.4 (58.8)	718.3 (67.5)

Note. Rp+ = items that received retrieval practice; Rp− = items from practiced categories that were themselves not retrieval practiced; Nrp = items that were not from categories that contained an item that received retrieval practice. Lexical decision times are presented in milliseconds.

were significant, $F(1, 31) = 6.13$, $p < .02$, $F(1, 29) = 30.50$, $p < .01$, respectively. This time the Prime × Item Type interaction was only marginally significant, $F(1, 29) = 2.50$, $p < .08$. As in Experiment 2, when the prime words were studied categories, there was a highly reliable effect of item type, $F(2, 62) = 7.60$, $p < .01$. The latency of rp+ words was significantly faster than that of the nrp items, $F(1, 31) = 4.10$, $p < .05$, confirming the priming effect of retrieval practice. There was again a significant difference between latencies of nrp words and rp− words, $F(1, 31) = 7.50$, $p < .01$, demonstrating an RIF effect in the lexical-decision task when the prime word was a studied category. When the prime word was an unstudied but related category a significant effect of Item Type was found, $F(2, 62) = 4.80$, $p < .01$. The latency of rp+ words was significantly shorter than that of the nrp words, $F(1, 31) = 4.20$, $p < .05$, but there was no significant difference between latencies of nrp words and rp− words ($F < 1.3$). Thus, the curious "rebound" effect detected in Experiment 4 did not occur again and, therefore, this is not a consistent effect. Crucially, however, there was no RIF effect. In addition to this, the latency of nrp words with studied primes was reliably shorter than that of nrp words with unstudied prime, $F(1, 31) = 19.15$, $p < .01$, but there was no significant difference between latencies of rp− words and studied or unstudied primes ($F < 2$). These findings are, for most contrasts, highly similar to those of Experiment 4 and clearly show that an initial category cued-recall test does not influence either lexical decision or a second category cued-recall test. The robust RIF effects induced by retrieval practice were again observed in recall, and again there was an RIF effect in lexical decision latencies when items were primed by studied categories but not when items were primed by unstudied categories.

Experiment 6

In all the experiments reported thus far, we have observed impaired recall of items targeted for inhibition (Tables 1 through 5). For lexical decision, however, the pattern of findings is more complicated. Lexical-decision times have been unaffected by manipulations intended to induce inhibition except when the items featured in the lexical-decision task were exact copies of items from the study lists, for example, studied fruit–apple, later primed with fruit, followed by a lexical decision to apple (Experiments 4 and 5). This pattern can be explained by our proposal that when cues access an episodic memory of the learning event, one that has been affected by retrieval practice, then inhibition is observed. When, however, cues access conceptual representations, no inhibitory pattern is observed and, instead, mainly activation (priming) is observed (Tables 1 through 5). This suggests one further test of the episodic inhibition account. When a task intervenes between study and test, if that task can be completed without accessing an episodic memory of the study phase, that is, by using conceptual knowledge, then no RIF pattern should result regardless of how strong the semantic association is between the study and intervening tasks. This assumes, of course, that the items used in the task do not automatically cue access of the episodic memory. For example, if the study phase containing the usual list of categories and exemplars is followed by a category exemplar generation phase, which although highly related to the study items nevertheless fails to access the memory of the study phase, then no RIF effects should occur. In order to test this, we constructed an experiment in which retrieval practice was replaced by category generation. As usual, category names and exemplars were studied, for example, fruit–apple, fruit–banana, fruit–pear, and this was followed by the generation of exemplars to category exemplar word-stem cues, for example, fruit–or______. The category exemplar word-stem cues were designed so that they featured a highly typical or dominant exemplar from a previously studied category. They were also constructed so that they could not be completed by accessing a memory of the previously studied list.

Method

Participants. Seventy undergraduate Hungarian students from the University of Szeged took part in the experiment. They participated in return for partial credit in a lower division psychology course. Their ages varied between 18 and 25 years, and their mean age was 20.4 years.

Procedure. Participants were tested in groups of 3 to 6 participants. The experiment was conducted in four phases: a learning phase, a generation phase, a distractor phase, and a category-cued recall phase. The learning, distractor, and category-cued recall phases were identical to those used in the previous experiments. The only difference from the standard RIF procedure was that participants took part in a generation task rather than a practice task. In the generation task, immediately after the learning phase, participants received two-letter stems together with a studied category cue and generated a word related to the category cues. The cues were designed to elicit exemplars that, when included in a retrieval practice phase, induced inhibition of words encoded in the study phase (as found in Experiments 3 through 5 in the present series). However, the cues were also designed to elicit exemplars that had not been presented in the study phase (see Appendix C). There was no requirement to recall any items from earlier in the experiment, and the generation task was introduced as a filler task, the main requirement of which was to respond as quickly as possible without error. The generation cues elicited previously unpresented words from half of the eight learning categories, and each cue was repeated three times. Participants generated the same items three times, exactly the same number of practice trials as in the practice phase in the standard RIF procedure. The practiced categories were counterbalanced between experimental groups. After completion of the generation task, participants undertook a numerical filler task for 5 min. Finally, the studied category names were represented, and participants were instructed to recall the items from the study list.

Results and Discussion

In this experiment rp− words were the items from the categories from which participants generated items in the generation phase. If the activation of items from the same categories impairs the representation of related items, then the recall performance of rp− items should be lower than that of nrp items (there was no generation of items from these categories). Note, there were no rp+ items in this version of the RIF procedure. The repeated generation of the items during the semantic generation task was successful, and in over 98% of responses the planned items from the correct categories were generated. As in the previous experiments, retrieval-induced forgetting was assessed by comparing the recall performance of unpracticed items from the practiced categories that is, items from categories later involved in the generations phase (rp− items) with the recall performance of unpracticed items from the previously unpracticed categories, that is, items from those categories that were not included in the generation phase later (nrp items). Mean recall of rp− items was 48% compared with a mean of 46% for nrp items (see Table 6). This difference was not reliable ($t < 1$) and shows that processing of

Table 6
Mean Percentage of Items Recalled in Experiment 6

	Retrieval practice status of item			
	Rp−		Nrp	
Type of recall	*M*(%)	*SD*	*M*(%)	*SD*
Category cued recall	47.79	26.58	45.66	26.25

Note. Rp− = Items from practiced categories that were themselves not retrieval practiced; Nrp = items that were not from categories that contained an item that received retrieval practice.

highly associated semantic items in between study and test does not necessarily induce inhibition (see also Anderson, Bjork, & Bjork, 2000).

General Discussion

In this series of experiments, we have sought to identify the locus of effects in directed forgetting and retrieval practice and, in particular, to isolate the types of mental representations that underlie these effects. Our version of episodic inhibition seeks to account for the findings by distinguishing between episodic and other types of long-term memory representations. We posit that during study, activation/inhibition of several different interacting memory systems mediates comprehension and learning. One result of this goal-oriented processing is an episodic memory or set of such memories which represent the study period. The memory represents the period by containing sensory–perceptual as well as conceptual–affective knowledge that was prominent during the period and that is goal-related (see Conway, 2001). Notably, the content or features of the episodic memory also reflect the processing priorities of the study period and those of any subsequent processing directed at the memory, for example, a directed forgetting instruction or period of selective retrieval practice. Thus, a pattern of activation/inhibition is present over the features of the episodic memory and when the memory is accessed, this pattern determines access of content. Highly active features, representations of rp+ items, for example, are highly accessible whereas other features that are lower in accessibility or, possibly, inhibited, for instance, TBF items and rp− items are more difficult to access. Thus, when an episodic memory of the study phase is accessed, as it has to be in directed forgetting and RIF, the pattern of activation/inhibition over its features powerfully influences what is recalled. The resulting effects, good recall of List 2 and rp+ items (Experiments 1 through 5), reflect high levels of activation of these items in the episodic memory of the study phase. Poorer recall of items from the TBF list and of nrp and rp− items shows either lower levels of activation of these items, as most probably is the case for nrp items, or as others have suggested, their inhibition, which may be the case for List 1 TBF items and rp− items (Anderson & Spellman, 1995; Bjork, 1989).

The critical point, and the purpose of the present research, is that if a poststudy task contains items from the study phase but neither requires access of the memory of the study phase nor automatically cues access, then the pattern of activation/inhibition represented in the memory will not influence performance on the poststudy task. In the present series of experiments, lexical decision was used to test this proposal. Lexical decision times to items from the study list were unaffected by manipulations that nonetheless gave rise to striking and marked decrements in recall performance, and this was the case across Experiments 1 through 5. The exception to this occurred when the lexical decision items were pairs of words (e.g., fruit–apple) from the study phase. These items did show the expected effects of retrieval practice, that is, rp+ items were responded to more quickly than all other items and nrp items were responded to more quickly than rp− items (Experiments 4 and 5). This was the case even though in the same experiments, studied items that were preceded by unstudied primes (e.g., food–apple) showed no RIF effects (Tables 4 and 5). These are problematic findings for those theories of inhibitory effects that argue for semantic processing as the locus of the effects, for example, Anderson and Spellman (1995). It is unclear why the semantic representations of some items should be difficult to access compared with others when all have received the same processing thought to induce inhibition. Even more problematic are the findings that in all the experiments, studied words were primed relative to unstudied words regardless of their status as facilitated or attenuated on other measures.

We believe that there are several processes at work here. First, the RIF effects in lexical decision in response to the copy cues may arise because these cues automatically access the episodic memory containing the pattern of activation/inhibition that mediates the RIF effect (so prominently present in category cued recall). Second, the absence of an RIF effect in lexical decision for those targets that were not exact copies of studied items may have occurred because these items could be responded to using conceptual/lexical representations, which do not preserve the RIF pattern of activation/inhibition. Third, the overall priming seen for studied words, in both directed forgetting and RIF, may occur because when conceptual/lexical networks dysfacilitate a pattern of activation/inhibition, they do not return to a baseline resting level immediately but rather change to some raised level of activation that sets them in readiness to process the same or similar items. Thus, the priming observed over all conditions in lexical decision may reflect speeded processing of items recently encountered. In other words, the inhibition (and activation) that would have been present during study and again during second-list learning or retrieval practice was only preserved in the episodic memory of the studied items, hence, the term *episodic inhibition.*

Extending the Episodic Inhibition View

A key feature of episodic inhibition is that the nature and pattern of activation/inhibition of the semantic or conceptual knowledge contained in an episodic memory will determine later recall. It has long been known that when a list of categories and their exemplars are studied, activation spreads through the representations of the categories to related knowledge (see, for example, Rosch, 1973). Thus, unpresented associates, features, and interitem relations, as well as those items explicitly presented in the study list may become part of an episodic memory of the study phase. It is, perhaps, in this way that independent cues (cues not presented during study and practice) can prove effective in recall (Anderson & Spellman, 1995). Thus, a cue such as “color” might be effective in eliciting “orange,” even though orange was originally encoded in terms of "fruit," and this is because the semantic attribute of orange, "color," is activated during the study phase and incorporated into the episodic memory. One advantage of this account is that it shows how, by encoding specificity (Tulving & Thompson, 1973), related but unpresented cues might be either effective or ineffective in eliciting recall. If unpresented cues do not correspond to knowledge in the episodic memory of the study phase and cannot be elaborated into cues that do correspond, then recall will be unsuccessful. Because the experimenter has relatively little control over what additional unpresented knowledge is encoded into a memory (but see Anderson et al., 2000, for manipulations that attempt to impose stronger control), then the effectiveness of independent cues will always fluctuate.

Thus, a useful aspect of the episodic inhibition account is that it offers a way to consider weaker and less consistent retrieval practice effects. For example, Anderson and Spellman (1995) demonstrated that inhibition can spread from an item directly inhibited by retrieval practice to items associated by semantic features but not themselves directly inhibited by retrieval practice. This effect was weaker than the main inhibitory effect and has not been observed in all studies (Williams & Zacks, 2001). Given that this is a relatively small effect it is perhaps not so surprising that it is not always observed. The present account argues that for this effect to occur at all, the semantic feature linking the items must be represented in the episodic memory of the study phase. Therefore, if Foods, strawberry, crackers, and Red, blood, tomato, were studied and Red–blood retrieval was practiced, Foods–strawberry would show some inhibition if the semantic feature "red" was represented in the episodic memory for both "strawberry" and "blood" (in fact the finding of Anderson & Spellman, 1995, and also of Anderson et al., 2000). If this feature was only represented in the episodic memory of the study phase for one of these items and not for both, then RIF should not occur; this is because the items would not under these circumstances compete in terms of overlap of semantic attributes.

Episodic inhibition also provides an account for a wide range of inhibitory findings and not just those limited to aspects of semantic processing. Consider for example the findings of Anderson, Bjork, and Bjork (2000), who replaced the retrieval practice word fragment cues with copy cues, for example, fruit–orange, and simply required that these be read during practice rather than recalled. No inhibitory effects were observed in the reading condition. According to episodic inhibition, this occurs because the read-only copy cues can be read by accessing conceptual/lexical representations in which no extensive representation of processing from the learning phase persists. As there is no requirement to access the episodic memory of the learning phase during the practice phase, knowledge in the episodic memory remains in a form similar to that at encoding, that is, with representations of list items active. When the episodic memory is accessed during category cued recall the pattern of activation/inhibition that would have been present had the standard retrieval practice manipulation been used is not present and therefore memory for RP items is not reliably impaired. Similarly, a study by MacLeod and MaCrae (2001), in which it was found that the effects of RIF dissipated over a 24-hr period, but when a similar retention interval intervened between study and retrieval practice, the usual RIF pattern was observed. This suggests that the pattern of activation imposed on the episodic memory by retrieval practice immediately following study begins to weaken and dissipate over the 24-hr retention interval. Most probably the episodic memory itself is undergoing some process of forgetting. According to the present account, the RIF-inducing effect of the delayed retrieval practice occurs because an episodic memory of the study phase is accessed during the delayed retrieval practice phase, and the RIF pattern of activation/inhibition is then generated by selective practice. If the episodic memory could not be accessed or if information in it was degraded, as may occur at even longer retention intervals, then according to episodic inhibition, no RIF pattern would be observed. Note, that it would be implausible to suggest that these long-lasting effects are mediated by patterns of activation/inhibition in semantic knowledge structures only.

We also note that episodic inhibition can be extended to procedures that induce RIF but that do not use retrieval practice. Directed forgetting is one such procedure, and it is very difficult to envisage how, for example, theories of RIF that are focused mainly on semantic accounts can be extended to directed forgetting and other procedures (Perfect et al., 2002). The findings of Experiments 1 and 2 could not simply be explained by a semantic features account such as that of Anderson and Spellman (1995). This does not mean that the semantic features account has no role in understanding RIF effects; quite clearly, it does (Anderson, 2003). What needs to be added, however, is that it does so by accounting for the pattern of activation/inhibition of conceptual knowledge contained in episodic memories.

According to episodic inhibition, the common mechanism underlying the relatively poor memory performance in directed forgetting and following retrieval practice is the pattern of activation/inhibition that exists over the contents of episodic memories of the study phase. Basden, Basden, and Morales (2003) reported findings that suggest that directed forgetting and retrieval practice might not share a common mechanism or set of processes. In their experiments, retrieval practice was performed several times on the second (remember) list in a directed forgetting experiment. It was reasoned that when List 2 remember items were studied, some List 1 TBF items might be accessed. If this is the case, then practicing recall of List 2 should provide further opportunities to inhibit List 1 and thus increase the directed forgetting effect. No increases in the magnitude of the directed forgetting effect were observed, suggesting that retrieval induced forgetting did not underlie the directed forgetting effect. These findings are not especially problematic for the episodic inhibition view, which simply argues that when an episodic memory is accessed, the pattern of activation/inhibition over the contents of the memory powerfully influences what will be remembered. Thus, retrieval practice in the form of recalling List 2, once or several times, will only influence an episodic memory of the TBF List 1 if that memory is accessed and processed during the retrieval practice phase. There is no guarantee that such access would spontaneously take place. Even if it did take place, it does not follow that inhibition would be increased. For instance, Conway et al. (2000) found that when 50% of the items in the two lists were close associates of each other, so that List 2 powerfully cued access of the TBF List 1 items, then the directed forgetting effect was abolished and instead a paradoxical increase in the recall of TBF List 1 items was observed. Conway et al. argued that this reflected a "release" from inhibition. Thus, the crucial issue is whether the memory is accessed prior to recall. However, we acknowledge that the possibility that different processes may mediate directed forgetting and the effects of retrieval practice, as Basden et al. (2003) argue, remains open. The present work strongly suggests that one feature these two procedures may share is reliance on an episodic memory of the study phase and the pattern of activation/inhibition that exists over the contents of the memory. There may, nonetheless, be other processes differentially associated with the two procedures which have not been considered here.

Finally, we note that Sahakyan and Delaney (2003) have argued that the directed forgetting effect might be conceptualized in terms of context change rather than in terms of inhibition (others too have expressed reservations about the notion of inhibition as it is used to explain changes in memory performance in directed forgetting, e.g., MacLeod, Dodd, Sheard, Wilson, & Bibi, 2003). Our concept of episodic inhibition is entirely compatible with the context change account, the main differences being that we focus on the nature of the episodic memories that are formed in response to context change. We retain the notion of inhibition, however, partly because the greater body of evidence supports this view (Bjork, Bjork, & Anderson, 1998) but also because it allows us to extend our development of theory beyond the directed forgetting task to the retrieval practice task and to other clearly inhibitory tasks such as memory for the inhibition of return (see Tipper, Grison, & Kessler, 2003). Currently, it is not clear how the context-change account of directed forgetting might be extended in this way.

Conclusions

The present series of experiments demonstrated simultaneous inhibition and activation of the same recently learned items in two different experimental procedures: directed forgetting (Experiments 1 and 2) and retrieval practice (Experiments 3 through 5). This can best be explained by postulating fast-changing conceptual/lexical knowledge structures, copies of which become represented in episodic memories. Knowledge in episodic memories is slow changing and preserves that pattern of activation/inhibition derived from the original experience or generated in it by subsequent access of memory details. Thus, an item inhibited in a memory may nonetheless be activated in a conceptual knowledge structure. If the memory is accessed then evidence of inhibition is found. If, in contrast, an item is accessed in a conceptual network, then evidence for inhibition is not detected. We termed this "episodic inhibition" and showed how it can be applied to a wide range of findings using different procedures.

References

Anderson, M. C. (2003). Rethinking interference theory: Executive control and the mechanisms of forgetting. *Journal of Memory and Language, 49,* 415–445.

Anderson, M. C., & Bell, T. (2001). Forgetting our facts: The role of inhibitory processes in the loss of propositional knowledge. *Journal of Experimental Psychology: General, 130,* 544–570.

Anderson, M. C., Bjork, E. L., & Bjork, R. A. (1994). Remembering can cause forgetting: Retrieval dynamics in long-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 1063–1087.

Anderson, M. C., Bjork, E. L., & Bjork, R. A. (2000). Retrieval-induced forgetting: Evidence for a recall-specific mechanism. *Psychonomic Bulletin and Review, 7,* 522–530.

Anderson, M. C., & Spellman, B. A. (1995). On the status of inhibitory mechanisms in cognition: Memory retrieval as a model case. *Psychological Review, 102,* 68–100.

Basden, B. H., & Basden, D. R. (1996). Directed forgetting: A further comparison of the list and item methods. *Memory, 4,* 633–653.

Basden, B. H., & Basden, D. R. (1998). Directed forgetting: A contrast of methods and interpretations. In J. M. Golding & C. M. MacLeod (Eds.), *Intentional forgetting: Interdisciplinary approaches* (pp. 139–172). Mahwah, NJ: Erlbaum.

Basden, B. H., Basden, D. R., & Gargano, G. J. (1993). Directed forgetting in implicit and explicit memory tests: A comparison of methods. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19,* 603–616.

Basden, B. H., & Basden, D. R., & Morales, E. (2003). The role of retrieval practice in directed forgetting. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29,* 389–397.

Bjork, E. L., & Bjork, R. A. (1996). Continuing influences of to-be-forgotten information. *Consciousness and Cognition, 5,* 176–196.

Bjork, E. L., Bjork, R. A., & Anderson, M. C. (1998). Varieties of goal-directed forgetting. In J. M. Golding & C. M. MacLeod (Eds.), *Intentional forgetting: Interdisciplinary approaches* (pp. 103–137). Mahwah, NJ: Erlbaum.

Bjork, R. A. (1989). Retrieval inhibition as an adaptive mechanism in human memory. In H. L. Roediger & F. I. M. Craik (Eds.), *Varieties of memory and consciousness: Essays in honour of Endel Tulving* (pp. 309–330). Hillsdale, NJ: Erlbaum.

Conway, M. A. (2001). Sensory perceptual episodic memory and its context: Autobiographical memory. *Philosophical Transactions of the Royal Society of London B, 356,* 1297–1306.

Conway, M. A., & Fthenaki, A. (2003). Disruption of inhibitory control of memory following lesions to the frontal and temporal lobes. *Cortex, 39,* 667–686.

Conway, M. A., Harries, K., Noyes, J., Racsmány, M., & Frankish, C. (2000). The disruption and dissolution of directed forgetting: Inhibitory control of memory. *Journal of Memory and Language, 43,* 409–430.

Fleck, D. E., Berch, D. B., Shear, P. K., & Strakowski, S. M. (2001). Directed forgetting in explicit and implicit memory: The role of encoding and retrieval mechanisms. *The Psychological Record, 5,* 207–220.

Füredi, M., & Kelemen, J. (1989). *A mai magyar nyelv szépprózai gyakorisági szótára* [A frequency dictionary of the literary language of Hungarian]. Budapest: Akadémiai Kiadó.

Geiselman, R. E., & Bagheri, B. (1985). Repetition effects in directed forgetting: Evidence for retrieval inhibition. *Memory and Cognition, 13,* 57–62.

Geiselman, R. E., Bjork, R. A., & Fishman, D. L. (1983). Disrupted retrieval in directed forgetting: A link with posthypnotic amnesia. *Journal of Experimental Psychology: General, 112,* 58–72.

Kónya, A., & Pintér, G. (1985). Kategória norma a verbális emlékezet vizsgálatához [Category norms for verbal memory research]. *Hungarian Psychological Review, 2,* 93–111.

MacLeod, C. M. (1989). Directed forgetting affects both direct and indirect tests of memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 13–21.

MacLeod, C. M. (1998). Directed forgetting. In J. M. Golding & C. M. MacLeod (Eds.), *Intentional forgetting: Interdisciplinary approaches* (pp. 1–59). Mahwah, NJ: Erlbaum.

MacLeod, C. M., Dodd, M. D., Sheard, E. D., Wilson, D. E., & Bibi, U. (2003). In opposition to inhibition. *Psychology of Learning and Motivation, 43,* 163–214.

MacLeod, M. D., & MaCrae. C. N. (2001). Gone but not forgotten: The transient nature of retrieval-induced forgetting. *Psychological Science, 12,* 148–152.

Neely, J. H. (1991). Semantic priming effects in visual word recognition:

A selective review of current findings and theories. In D. Besner, & G. Humphreys (Eds.), *Basic processes in reading: Visual word recognition* (264–336). Hillsdale, NJ: Erlbaum.

Perfect, T. J., Moulin, C. J. A., Conway, M. A., & Perry, E. (2002). Assessing the inhibitory account of retrieval-induced forgetting with implicit-memory tests. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* 1111–1119.

Racsmány, M. (2003). *Episodic and semantic inhibition in directed forgetting and retrieval-induced forgetting.* Unpublished doctoral dissertation, University of Bristol, Bristol, England.

Rosch, E. (1973). Natural categories. *Cognitive Psychology, 4,* 328–349.

Sahakyan, L., & Delaney, P. F. (2003). Can encoding differences explain the benefits of directed forgetting in the list-method paradigm? *Journal of Memory and Language, 48,* 195–201.

Tipper, S. P. (2001). Does negative priming reflect inhibitory mechanisms? A review and integration of conflicting views. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 54A,* 321–343.

Tipper, S. P., Grison, S., & Kessler, K. (2003). Long-term inhibition of return of attention. *Psychological Science, 14,* 19–25.

Tulving, E., & Thompson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review, 80,* 352–373.

Williams, C. C., & Zacks, R. T. (2001). Is retrieval-induced forgetting an inhibitory process? *American Journal of Psychology, 3,* 329–354.

Appendix A

Category-Exemplar Pairs Used in Experiment 3

Categories	Items
Instrument (musical)	guitar, cello, piano, violin, flute, harp
Vehicle	train, car, ship, tram, wheeler, bicycle
Clothes	coat, gloves, boots, socks, gown, cap
Colour	red, yellow, black, purple, brown, green
Animal	tiger, deer, cat, horse, dog, cow
Furniture	armchair, carpet, wardrobe, lamp, couch, table
Occupation	lawyer, actor, miner, cook, painter, policeman
Fruit	plum, pear, apricot, grape, raspberry, orange
Filler categories: flower, reading matter	

Appendix B

Category-Exemplar Pairs Used in Experiments 4 and 5

Categories	Items
Instrument (musical) or music	guitar, cello, piano, violin, flute, harp
Vehicle or traffic	train, car, ship, tram, wheeler, bicycle
Flower or fragrant	tulip, narcissus, rose, poppy, carnation, violet
Clothes or fashion	coat, gloves, boots, socks, gown, cap
Color or paint	red, yellow, black, purple, brown, green
Animal or mammal	tiger, deer, cat, horse, dog, cow
Furniture or apartment	armchair, carpet, wardrobe, lamp, couch, table
Fruit or food	plum, pear, apricot, grape, raspberry, orange
Filler categories: occupation, reading matter	

Appendix C

English Translation of Category-Exemplar Pairs and Semantically Generated Items Used in Experiment 6

Categories	Items	Generated items
Instrument (musical)	guitar, cello, piano, violin, flute, harp	trumpet, drum, cymbal
Vehicle	train, car, ship, tram, wheeler, bicycle	underground, lorry, boat
Clothes	coat, gloves, boots, socks, gown, cap	skirt, jacket, sweater
Color	red, yellow, black, purple, brown, green	grey, blue, claret
Animal	tiger, deer, cat, horse, dog, cow	donkey, badger, squirrel
Furniture	armchair, carpet, wardrobe, lamp, couch, table	curtain, coat-rack, bookshelf
Occupation	lawyer, actor, miner, cook, painter, policeman	teacher, joiner, plumber
Fruit	plum, pear, apricot, grape, raspberry, orange	melon, currant, blackberry
Filler categories: flower, reading matter		

Received August 24, 2004
Revision received June 23, 2005
Accepted September 7, 2005 ■

IV.2. Autonoetic consciousness and retrieval inhibition (Study 2)

Paper: Memory awareness following episodic inhibition. Mihály Racsmány, Martin A Conway, Edit A Garab, Gábor Nagymáté, *Quarterly Journal of Experimental Psychology* 05/2008; 61(4):525-34. DOI:10.1080/17470210701728750

THE QUARTERLY JOURNAL OF EXPERIMENTAL PSYCHOLOGY
2008, 61 (4), 525–534

Psychology Press
Taylor & Francis Group

Short article

Memory awareness following episodic inhibition

Mihály Racsmány
Research Group on Cognitive Science, Hungarian Academy of Sciences, Budapest University of Technology and Economics, Budapest, Hungary, and University of Szeged, Szeged, Hungary

Martin A. Conway
The Leeds Memory Group, Institute of Psychological Sciences, University of Leeds, Leeds, UK

Edit A. Garab
Department of Cognitive Science, Budapest University of Technology and Economics, Budapest, Hungary

Gábor Nagymáté
Loránd Eötvös University, Budapest, Hungary

Three experiments used directed forgetting (DF) and retrieval practice (RP) to investigate the relation of inhibited items to states of memory awareness occurring at test. In Experiment 1 using list DF robust inhibitory effects were present in cued recall, but in a recognition test these effects were only present in responses accompanied by recollective experience. In Experiments 2 and 3 using RP reliable effects of inhibition were found but these did not relate systematically to states of memory awareness. It is suggested that in DF the to-be-forgotten items are tagged at study as *not to be recollectively experienced* and so have a specific, inhibitory, relation to states of recollective experience occurring during test. In RP no tagging takes place, and inhibition is automatic or nonintentional and consequently does not have a specific relation to states of memory awareness at test.

Keywords: Directed forgetting; Selective practice; Episodic memory; Recollective experience; Retrieval inhibition.

A remarkable finding in the study of human memory is that many of the major memory effects established in laboratory and field studies over the past 30 years are only present when rememberers consciously recollect recently acquired materials (see Gardiner & Richardson-Klavehn, 2000, for a review). If remembering is not accompanied by the experience of recollection (Tulving, 1985; see too Conway, 2005) but instead by some other state of memory awareness, for instance a feeling of familiarity, then standard effects such as level of processing, generation effects, picture superiority

Correspondence should be addressed to Mihály Racsmány, Research Group on Cognitive Science, Budapest University of Technology and Economics, Stoczek u. 2., Budapest, 1111, Hungary. E-mail: racsmany@cogsci.bme.hu

Financial support for the research was provided by OTKA (Hungarian National Science Foundation) 49 840-T53 and K68463. Martin A. Conway is supported by a Professorial Fellowship, RES–051–27–0127, from the Economic and Social Research Council of Great Britain. Mihály Racsmány is grantee of the Bolyai János Research Scholarship of the Hungarian Academy of Science.

http://www.psypress.com/qjep
DOI:10.1080/17470210701728750

effects, and many others are not observed. Thus, the state of consciousness termed by Tulving (1985) as *autonoetic consciousness* or *recollective experience* would seem to be integral to many variables known to determine memory performance.

In the present research we investigate whether this requirement for recollective experience is also integral to manipulations that impair memory performance by inducing inhibition of recently studied items. Memory inhibition procedures take essentially two forms: those that feature a conscious intention to forget and those that do not. Our conjecture is that inhibitory manipulations that contain a conscious intention to forget may have a specific relation to later states of memory awareness in remembering. In particular, it may be that episodic memories that are formed under intentional conditions to forget may be marked in some way as *not to be recollectively experienced* (NTBRE). In contrast, episodic memories formed under conditions of inhibition that do not feature a conscious intention to forget may not be marked in this way. It follows then that intentional inhibitory procedures that give rise to episodic memories, or episodic memory content, that are tagged NTBRE should lead to attenuated recollective experience in remembering. Memory inhibitory procedures that do not entail a conscious intention to forget should not result in NTBRE-tagged memories and as a consequence should not vary in any systematic way with recollection.

Here we examine two states of memory awareness operationally referred to as *remember* and *familiarity* or R and F responses. Recollective experience (R) is known to involve the recall of highly specific details, usually in the form of visual or other modes of imagery (Gardner, Richardson-Klavehn, & Ramponi, 1998). It features a strong sense of the self in the past, and attention turns inwards to focus on the memory construction (Conway, 2005; Tulving, 2002). Familiarity (F) on the other hand is a distinct state of conscious memory awareness that does not have these features and instead is characterized by a strong feeling that some item in the current environment has been encountered recently. We hypothesize that tasks that intentionally induce memory inhibition will impair or reduce R responses but leave F responses unaffected. We also further conjecture that intentionality of remembering may be important here too and that inhibition will be greater in voluntary than in involuntary remembering. This may in part be because the aim of voluntary remembering is to recollectively experience the past.

EXPERIMENT 1

In this experiment we used *directed forgetting* (DF) to induce inhibition of a recently learned list of words (see E. L. Bjork & Bjork, 1996, and R. A. Bjork, 1989, for a review of the DF procedure and MacLeod, 1998, for a more general review of DF). The specific procedure used was DF by lists. In the DF by lists procedure two lists are studied for later recall, and a surprise midlists instruction designates the first list as to be forgotten (TBF) or to be remembered (TBR). The exact procedure followed here is that of Conway, Harries, Noyes, Racsmány, and Frankish, (2000). However, one change to the procedure of Conway et al., is that we use a within-subjects design and refer to the two memory instruction manipulations as the directed forgetting or DF condition and the directed remembering or DR condition. In order to examine memory awareness during remembering three different memory tests were conducted: list cued recall, word stem completion, and a recognition test.

In the cued-recall test List 1 was always recalled first. This test assesses the effect of the forget instruction, by comparing TBF List 1 with TBR List 1 memory performance, while minimizing any potential output interference from memory for List 2 (Conway et al., 2000; Racsmány & Conway, 2006). The word stem completion test was based on a similar test used by Richardson-Klavehn and Gardiner (1996). In this test word stems were completed with words studied on the earlier lists (voluntary or intentional condition). If a stem could not be completed in this way it was completed with any appropriate word that could be generated. When this word was in fact a word from the earlier lists then this constituted involuntary recall. Finally, a recognition test for the two lists was undertaken, and for every word judged to be "old" the state of memory awareness, R or F, accompanying that judgement was recorded.

Method

Participants

A total of 27 undergraduate Hungarian students from the Technical University of Budapest participated in return for partial credit in an introductory psychology course. Their ages varied between 19 and 25 years, and there were 16 females and 11 males.

Procedure

Participants were tested individually and were informed that they were taking part in a memory experiment and that their memory for the studied items would be tested later. The order of the midlists forget/remember instructions was counterbalanced between individuals. In the study phase the TBR words were presented visually on a computer screen. Each word was displayed for 2 s with a 2-s interitem interval. After the first 12 words had been presented, the participants were instructed to stop. At this point participants in the DF condition were given the following forget instruction: "The list you have just studied was a practice list to familiarize you with the experimental procedure. You should now forget these words, try to put them out of your mind. The experimental list will be presented now." In the DR condition the same procedure was followed but the midlists instruction was: "That is the end of List 1. You must try to keep those words in mind while you learn the second list which will be presented now."

After all words had been studied participants were given a 5-minute arithmetic distractor task. The distractor task was followed by the cued recall test. Participants were given a sheet of paper and were instructed to try to recall as many words as they could from both lists. They were asked to start at the top of the page and write each recalled word under the previous word. In order to reduce the role of output interference effect in recall performance we followed the recall instruction of Conway et al. (2000, Exp. 7): Participants had to recall List 1 words first and then List 2 words. Every participant took part in both the DF and the DR conditions, and the order of conditions was counterbalanced among participants. We used four experimental learning lists consisting of words of moderate to high frequency. The order of presentation of the lists was random for each participant. Each study list contained 12 items.

After free recall participants completed the word stem test (based on David & Brown, 2003). This test consisted of 24 word stems, e.g., "ta__" for *table*, and each word stem could be completed with at least two different Hungarian words. The order of presentation of the stems was randomized for each participant. Participants were instructed to complete the stems with words that they had previously studied. If able to do so, participants were then asked to provide a second word stem completion, thus indicating that their first completion was based on recall—that is, voluntary conscious memory. If they were unable to complete the stem with a studied word, participants were asked to complete the stem with the first word that came to mind. Some participants might complete the stem with the first word item coming to mind and then recognize that completion as a studied word. To indicate such occurrence, participants were instructed to place an asterisk next to the completed item. The recognition test consisted of 48 items with 24 studied (old) and 24 unstudied (new) items (every new word had a first syllable that corresponded to a syllable in one of the studied items). For every item judged to be "old" participants also indicated the basis of their judgement, R or F, following standard instructions (see Gardiner et al., 1998).

Results and discussion

In the list cued recall the critical interaction of List × Cue was present, $F(1, 26) = 26.47$, MSE equals; 1.48, $p < .01$. DF List 1 recall was found to be reliably poorer than DR List 1, $t(1, 26) = 11.2$, $p < .01$ (Table 1). Together this pattern shows a strong and reliable DF effect. There were no reliable effects of voluntary versus involuntary

Table 1. *Mean percentages of memory performances in Experiment 1*

	DF		*DR*	
	List 1	*List 2*	*List 1*	*List 2*
Cued recall	*29.7* (12.3)	40 (19.1)	42.3 (19.6)	28.4 (20.1)
Voluntary recall	36.3 (17.6)	37.7 (18.1)	43.4 (13)	39.9 (18.6)
Involuntary recall	6.1 (4.6)	8.1 (5.1)	7.8 (6.1)	9.7 (7.4)
Recognition hits	83 (16.5)	83.9 (11.4)	86.7 (13.2)	82.4 (14.2)
"Remember" responses	*58.4* (25.8)	68 (22.4)	74.4 (28)	63.1 (29.9)
"Familiarity" responses	39.5 (26.1)	31 (23.3)	25.6 (28)	36.9 (29.9)

Note: Standard deviations are presented in parentheses. DF refers to the directed forgetting condition; DR refers to the directed remembering condition. Figures in italics reflect significant to-be-forgotten memory impairment compared to control and to-be-remembered scores.

recall for List 1 compared to List 2 in either condition, $F(1, 26) = 1.7$, $MSE = 0.32$, $p > 0.1$, and $F(1, 26) = 2.8$, $MSE = 0.2$, $p > 0.1$, respectively (Table 1). Similarly, there was no overall effect of DF in recognition, $F(1, 26) = 1.14$, $MSE = 1.1$, $p > 0.1$. A post hoc analysis of power for the critical contrast of List 1 performance in DF and DR conditions showed that with an alpha at .05 the calculated power is 0.66, indicating that the lack of difference in the recognition task was not due to low power (Erdfelder, Faul, & Buchner, 1996).

More interesting was that a powerful DF effect was found in recognition accompanied by recollective experience, R responses, $F(1, 26) = 21.78$, $MSE = 1.45$, $p < .01$. For R responses, DF List 1 recall was found to be reliably poorer than DR List 1. An inversed DF effect was observed in F responses, $F(1, 26) = 19.65$, $MSE = 0.33$, $p < .01$ (see Table 1). These findings quite clearly demonstrate a strong DF effect when remembering is accompanied by recollective experience, but not when it is accompanied by feelings of familiarity. The voluntary/involuntary dimension appeared to be orthogonal to the DF effect and was dominated by a powerful recency effect in both DF and DR groups.

In the directed forgetting paradigm the standard finding with free recall is that the forget instruction decreases the level of first-list items and increases recall of second-list items compared to the condition in which only the remember instruction was used. Another usual finding in this procedure is that the effect of the forget instruction generally observed in free recall is abolished in recognition. This pattern was observed in Experiment 1 where significant and standard directed forgetting effect was found in free recall but not in recognition. In a further analysis of recognition performance it was found that the specific effect of forget instruction was present in "R" items and reversed for "F" items.

EXPERIMENT 2

This second experiment uses the retrieval practice procedure of Anderson, Bjork, and Bjork (1994). In this procedure participants practise selectively recalling items from a previously studied list. The effects of this selective practice are to induce inhibition of closely associated items in a memory of the original study list (Racsmány & Conway, 2006). Memory for these unpractised items is reliably poorer than memory for baseline items. Our question is: Will this effect only be presented in R responses as was the case with DF?

Method

Participants

A total of 48 undergraduate Hungarian students from the Technical University of Budapest participated in return for partial credit in an introductory psychology course. Their ages varied between 18 and 29 years, and there were 27 females.

Procedure and materials

Following Anderson et al. (1994) we constructed 10 categories, 2 of which were used as fillers. Each category consisted of 12 exemplars from each of 8 target categories forming two subsets (6 items) with moderate-to-high-frequency words drawn from two published Hungarian frequency norms (Füredi & Kelemen, 1989; Kónya & Pintér, 1985). The categories chosen were: animals, fruits, sports, cooking utensils, clothes, musical instruments, professions, and reading materials (fillers: flowers, nations). For each category, two counterbalancing sets of 6 items were each created and used an equal number of times as targets versus lures on recognition task. Each item in the first subset had a first syllable that corresponded to a syllable of an item in the other subset. We created two subsets from the 8 target categories and designated them an equal number of times as practised and nonpractised categories. The practised and nonpractised exemplars were counterbalanced as well. There were four phases to the experiment.

Phase 1: Study. A PC controlled the study phase. The participants saw category–exemplar pairs on the monitor screen, and they were told to try to remember the category examples as best as they could. Each category exemplar pair was presented in uppercase letters at the centre of the screen for 5 seconds.

Phase 2: Retrieval practice. When participants had completed the study phase, the experimenter distributed retrieval practice booklets. Each page in the booklet contained one of the category names studied previously and the first two letters of one member of that category also studied previously. Participants were instructed to complete the exemplar fragment with one of the words they had studied earlier. Participants were told that some of the examples might be tested more than once and that in those cases they should respond with the remembered item.

Phase 3: Filled retention interval. After the retrieval practice phase had been completed, booklets were collected, and participants were given an unrelated mathematical task for 5 minutes.

Phase 4: Recall phase. Participants were given recall booklets with the name of one of the previously studied categories on the top of each page. Participants were instructed to recall as many examples as they could in the 10-minute period allocated for this test. They were constrained to keep the order of categories as they were presented in the recall booklet. Order of presentation of category cues was counterbalanced over participants.

As in Experiment 1, category cued recall was followed by a test list consisting of 48 category label–word stem pairs. Each word stem could be completed with at least two different Hungarian words within the same categories. Order of the presentation of category-plus-stems was random with a different random order for each participant. Using these items voluntary and involuntary memory was assessed in the same way as in Experiment 1.

Phase 5: Recognition phase. Finally, participants took a 96-item recognition test. The target stimuli consisted of 12 exemplars from each of eight target categories with moderate-to-high-frequency words drawn from two published Hungarian frequency norms (Füredi & Kelemen, 1989; Kónya & Pintér, 1985). We created two subsets from the eight target categories and designated each an equal number of times as practised and nonpractised categories. In the recognition list old category exemplar pairs from the study list were mixed with an equal number of new pairs. In this task participants were given category–exemplar pairs for an old–new decision. As in Experiment 1, for every item judged to be "old" participants also indicated the basis of their judgement, R or F, following standard instructions (see Gardiner et al., 1998).

Results and discussion

It can be seen from Table 2 that there was a large effect of retrieval practice on category-cued

Table 2. *Mean percentages of memory performances in Experiment 2*

	Rp+ items	*Rp− items*	*Nrp items*
Category-cued recall	76 (17.3)	*46.7* (18.5)	55.7 (13.5)
Voluntary recall	83.5 (17.1)	*57.2* (19.1)	65.2 (17.5)
Involuntary recall	7.7 (8.4)	10.3 (10.9)	12.1 (9.6)
Recognition hits	93.5 (7.1)	*84.1* (15.1)	89.3 (8.4)
False alarms	21.1 (20)	23.8 (16.8)	25.8 (19.6)
"Remember" response	63.7 (27.8)	61.3. (26.3)	61.8 (20.3)
"Familiarity" response	35.9 (28.2)	38.7 (26.3)	37.8 (28.2)

Note: Standard deviations are presented in parentheses. Rp+ = items that received retrieval practice; Rp− = items from practised categories that were themselves not retrieval practised; Nrp = items that were not from categories that contained an item that received retrieval practice. Figures in italics reflect significant to-be-forgotten memory impairment compared to control and to-be-remembered scores.

recall. This main effect of item type was reliable, $F(2, 94) = 81.4$, $MSE = 0.99$, $p < .01$, and so was the critical contrast of Nrp items (items that were not from categories that contained an item that received retrieval practice) with Rp− items (items from practised categories that were themselves not retrieval practised), $t(1, 47) = 4.18$, $p < .01$. There was then a robust retrieval practice effect similar to that observed in many previous studies (for recent findings showing this effect, see Racsmány & Conway, 2006). In terms of voluntary versus involuntary recall there were no comparable effects of retrieval practice, although the critical contrast of Nrp and Rp− items was significant in voluntary cued recall, $t(47) = 2.04$, $p < .05$, but not in the involuntary recall, $t(47) = 0.1$, $p = .8$. As this manipulation has produced mainly null effects in both experiments we conclude that this task is not sensitive to these manipulations, and, therefore, we do not discuss it further. Although memory performance increased markedly in the recognition test, see Table 2, the effects of retrieval practice observed in recall were maintained in recognition, $F(2, 94) = 8.84$, $MSE = 4.2$, $p < .01$, Finally, and of particular interest to the present study, there were no reliable effects of retrieval practice in R and F responses. Memory performance for items recognized with recollective experience was higher than that of items recognized with familiarity: a difference that did not vary significantly over Rp+ (items that received retrieval practice), Nrp, and Rp items, $F(2, 94) = 0.59$, $MSE = 3.39$, $p > .1$; a post hoc analysis of power for the omnibus ANOVA of recognition data revealed that the lack of significant difference was not due to sample size, partial eta squared = .21, critical $F = 3.64$; power = 0.9 (Erdfelder et al., 1996). Thus, unlike the DF effect, the effects of retrieval practice were only present in overall responses and were not selectively confined to R responses.

In summary it was found that prior retrieval of Rp+ items facilitated later retrieval of those items, but decreased the recall of Rp− items relative to the Nrp baseline items. The same pattern was observed in voluntary cued recall and recognition, but not in involuntary cued recall. Critically there was no interaction between the selective practice effect and the recollective judgement of "R" and "F" items in the recognition test. This pattern of data suggests that previous inhibitory consequence of selective practice influences "F" and "R" items equally.

EXPERIMENT 3

In this third experiment we modified the retrieval practice procedure of Experiment 2 in two ways. In Experiment 3 the recognition task preceded cued recall, and we measured not only hits and false alarm rates but also reaction times of recognition decisions.

Method

Participants

A total of 48 undergraduate Hungarian students from the Technical University of Budapest participated in return for partial credit in an introductory psychology course; 4 participants were later discarded as they misunderstood the instruction and exchanged key presses during the recognition task. Their ages varied between 21 and 28 years, and there were 20 females.

Procedure and materials

As in Experiment 2 we constructed 10 categories, 2 of which were used as fillers. The target stimuli consisted of 12 exemplars from each of 8 target categories with moderate-to-high-frequency words drawn from two published Hungarian frequency norms (Füredi & Kelemen, 1989; Kónya & Pintér, 1985). The categories chosen were: animals, fruits, sports, clothes, musical instruments, professions, flowers, nations, and reading materials (fillers: colours, vehicles). Following the procedure of Hicks and Starns (2004) for each category, two counterbalancing sets of 6 items were created and used an equal number of times as targets versus lures in the recognition test. We created two subsets from the 8 target categories and designated each an equal number of times as practised and nonpractised categories. The practised and nonpractised exemplars were counterbalanced. As in Experiment 2 participants viewed category–exemplar pairs on the monitor screen and were instructed to try to remember the category examples as best as they could. Each category–exemplar pair was presented in uppercase letters at the centre of the screen for 5 seconds. A computer program was used to present the study list, and the same program was used to present the recognition task and to record participant's responses. When participants had completed the study phase, they took part in the practice and unrelated filler tasks used in the previous experiment. Following this they took the recognition test. In this test participants were given individual items for an old–new decision (category cues were not used). Each word remained on the computer screen until the person responded with a maximum response window of 2 seconds. We applied a relatively strict response window to force participants to respond as fast as possible in order to detect minor reaction time differences between conditions. Nevertheless, the 2-s response window is far above the average response time (700–1,200 ms) observed in previous recognition studies (see MacLeod, 1999). If participants responded "old" they also indicated whether the response was based on remembering, knowing, or guessing (R, K, or G) using labelled keys. At the beginning of the test participants were given both written and verbal instructions. They also practised the response keys using the filler exemplars and associated new words. Standard remember–know instructions were used (Gardiner et al., 1998). Finally, after the recognition test participants took part in a category cued recall as in Experiment 2. Order of presentation of category cues was counterbalanced across participants.

Results and discussion

It can be seen from Table 3 that RP had a positive effect on recognition performance of Rp+ items, $F(2, 86) = 38.9$, $MSE = 0.5$, $p < .01$; however, the inhibitory effect of RP on the Rp− hits observed in Experiment 2 was not present in this experiment. The critical contrast of Nrp with Rp− items did not show a reliable difference, $t(1, 43) = 0.22$, $p > .1$. A post hoc analysis of power for the recognition data revealed that the failure to find a significant difference was not due to small sample size, partial eta squared = .64, critical $F = 3.09$; power = 0.9. This finding, which it might be noted runs counter to the findings of Hicks and Starns (2004), is not wholly unexpected as the RP effect is not always observed in recognition memory hits (Koutstaal, Schacter, Johnson, & Galluccio, 1999). The same is also true of DF effects (see MacLeod, 1998). The standard explanation is that the copy cues used in recognition memory tests overcome inhibitory effects (E. L. Bjork & Bjork, 1996). In RP this appears to be case in at least some studies

Table 3. *Mean percentages of recognition accuracy, means of RT medians, and means of cued-recall performance in Experiment 3*

	Rp+ items	*Rp− items*	*Nrp items*
Recognition hits	92.3 (9.1)	73.5 (15.5)	74.1 (13.4)
False alarms	7.9 (7.1)	7.9 (7.1)	8.1 (7.4)
"Remember" responses	70.3 (20.8)	70.7 (21.5)	70.5 (21.2)
"Know" responses	25 (18.8)	21.7 (17.3)	21.6 (18.7)
"Guess" responses	4.7 (7.5)	7.6 (11.9)	7.9 (9.3)
Recognition RTs	1,006.1 (217.1)	*1,131.9* (254.4)	1,012.7 (212.9)
Category-cued recall	76.4 (13.6)	57.5 (16.6)	60.1 (14.2)

Note: Standard deviations are presented in parentheses. RT (reaction time) medians in ms. Figures in italics reflect significant to-be-forgotten memory impairment compared to control and to-be-remembered scores.

although systematic RP effects in recognition have been observed in other series of experiments—see, for example, Racsmány and Conway, (2006). Despite the lack of an RP effect in hits there was a strong effect in reaction times. Table 3 shows the mean RTs of recognition decisions, and it can be seen that these varied over Rp+, Nrp, and Rp− items and that this effect was reliable, $F(2, 86) = 18.4$, $MSE = 2220.5$, $p < .001$. Recognition of Rp− items was a full 100 ms slower than recognition of Nrp items, and this difference was also reliable, Tukey HSD test, $df = 86$, $p < .01$. Thus, it seems that although the copy cues of the recognition test were effective in overcoming the inhibition induced by RP, inhibition was nonetheless still present in access times to inhibited items. This is interesting as it suggests that despite previous inconsistent findings in amount of Rp− items recognized there may be an additional and more consistent effect present in retrieval times. The powerful copy cues present in a recognition test may, then, overcome RP induced inhibition but the inhibitory effect of the RP manipulation remains in access times to memory details (list items).

Crucially, for present concerns, memory performance for items recognized with recollective experience was reliably higher than that of items recognized with familiarity or guess, $F(2, 86) = 295.9$, $MSE = 8.8$, $p < .001$, but this did not vary significantly over Rp+, Nrp, and Rp− conditions, $F(2, 164) = 0.76$, $MSE = 2.2$, $p > .1$. This result replicates the findings of Experiment 2 and further demonstrates that there is no systematic relation between RP and recollective experience at test.

Finally there was no RP effect in category-cued recall: The main effect of item type was reliable, $F(2, 86) = 36.18$, $MSE = 0.46$, $p < .01$, showing elevated recall of Rp+ items relative to the other two conditions, but the critical contrast of Nrp with Rp− items was not significant, $t(1, 47) = 1.18$, $p > .1$. The results in cued recall are highly similar to those in the recognition hits rates and most probably are so because of the effect of the recognition test in releasing RP-induced inhibition.

Changing the sequence of recognition and recall tasks in this experiment yielded significant differences in the pattern of data. This time we found no significant difference between Rp− and Nrp baseline items in recognition hits, but found a strong difference in retrieval times: Recognition of Rp− items was reliably slower then recognition of Nrp baseline items. Another difference relative to the previous results was that we did not find the critical difference of Nrp and Rp− items in cued recall either. However, the most important finding of the present experiment was that retrieval practice had no differential effect on R and K items and in this respect quite clearly differs from directed forgetting.

GENERAL DISCUSSION

In the three experiments we found robust effects of directed forgetting (DF) and retrieval practice (RP). In DF items on a list designated as TBF were poorly recalled compared to baseline. This effect was, however, only present for items associated with recollective experience at recall. In RP items semantically associated with practised items but that were themselves not practised were recalled to reliably lower levels than baseline items (Experiment 2). This was the case in responses overall but, unlike DF, there was no selective effect in items recollectively experienced in the recognition test. As hypothesized earlier this pattern of findings suggests that two tasks may have rather different relations to memory awareness in later recall.

According to *episodic inhibition* (Racsmány & Conway, 2006) recall patterns resulting from both DF and RP are mediated by an episodic memory of the study phase. This memory represents patterns of activation that predominated in conceptual and other knowledge networks during study. In the case of DF the explicit instruction to forget gives rise to a general attempt to inhibit the recently acquired list but this is usually unsuccessful unless followed by second-list learning (R. A. Bjork, 1989). During second-list learning a few items, probably fewer than one or two (Conway et al., 2000), spontaneously come to mind and in so doing trigger further inhibition of List 1 items. A critical point here is that in order for this to occur, TBF List 1 items have to be identifiable, in some way, as TBF. In contrast, in RP there is no explicit instruction to forget, and studied items can, in principle, be brought to mind at any time in the practice phase. Indeed, informal enquires of our participants indicated that this does occur at least occasionally on some trials. We assume, however, that as the practice increases such intrusions become much less frequent, and this is because the repeated recall of the practice items repeatedly induces inhibition of closely associated items represented in the episodic memory of the study phase. The important point then is not that Rp− items may intrude into RP but rather that when they do they are not identified as TBF. They may become TBF in order to reduce interference with Rp + items but this is almost certainly a nonconscious process that does not require an explicit intention to forget.

Our suggestion is, then, that in DF List 1 studied items represented in an episodic memory of the study phase are marked as TBF. In particular, and as argued earlier, we propose that these episodic representations are specifically tagged not to be recollectively experienced—NTBRE. It may be that rather than individual items in the episodic memory being tagged NTBRE the entire memory is tagged, and the contents of the memory inherit the tag. Thus when the episodic memory is accessed in cued recall its contents are more difficult to access than those of a comparable untagged memory (memory for List 1 in the remember group, for example). Interestingly, the contents of a memory tagged NTBRE can still strongly influence performance in tasks that do not require conscious recall—for example, stem completion, lexical decision, and so on (E. L. Bjork & Bjork, 1996; Perfect, Moulin, Conway, & Perry, 2002; Racsmány & Conway, 2006). These findings lend further support to the NTBRE tagging idea because they show that it is the explicit conscious representation of inhibited memories that is affected by intentions to forget.

Original manuscript received 25 May 2006
Accepted revision received 14 May 2007
First published online 28 November 2007

REFERENCES

Anderson, M. C., Bjork, E. L., & Bjork, R. A. (1994). Remembering can cause forgetting: Retrieval dynamics in long-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*, 1063–1087.

Bjork, E. L., & Bjork, R. A. (1996). Continuing influences of to-be-forgotten information. *Consciousness and Cognition*, *5*, 176–196.

Bjork, R. A. (1989). Retrieval inhibition as an adaptive mechanism in human memory. In H. L. Roediger & F. I. M. Craik (Eds.), *Varieties of memory and consciousness: Essays in honour of Endel Tulving* (pp. 309–330). Hillsdale, NJ: Lawrence Erlbaum Associates.

Conway, M. A. (2005). Memory and the self. *Journal of Memory and Language*, *53*, 594–628.

Conway, M. A., Harries, K., Noyes, J., Racsmány, M., & Frankish, C. R. (2000). The disruption and dissolution of directed forgetting: Inhibitory control of memory. *Memory and Language*, *43*, 409–430.

David, D., & Brown, R. (2003). The impact of different directed forgetting instructions on implicit and explicit memory: New evidence from modified process dissociation procedure. *Quarterly Journal of Experimental Psychology*, *56A*, 211–231.

Erdfelder, E., Faul, E., & Buchner, A. (1996). GPOWER: A general power analysis program. *Behaviour Research Methods, Instruments, & Computers*, *28*, 1–11.

Füredi, M., & Kelemen, J. (1989). *A mai magyar nyelv szépprózai gyakorisági szótára* [A frequency dictionary of the literary language of Hungarian]. Budapest, Hungary: Akadémiai Kiadó.

Gardiner, J. M., & Richardson-Klavehn, A. (2000). Remembering and knowing. In E. Tulving & F. I. M. Craik (Eds.), *The Oxford handbook of memory* (pp. 229–245). Oxford, UK: Oxford University Press.

Gardiner, J. M., Richardson-Klavehn, A., & Ramponi, C. (1998). Limitations of the signal detection model of the remember–know paradigm: A reply to Hirshman. *Consciousness and Cognition*, *7*, 285–288.

Hicks, J. L., & Starns, J. J. (2004). Retrieval-induced forgetting occurs in item recognition. *Psychonomic Bulletin and Review*, *11*, 125–130.

Kónya, A., & Pintér, G. (1985). Kategória norma a verbális emlékezet vizsgálatához [Category norms for verbal memory research]. *Hungarian Psychological Review*, *2*, 93–111.

Koutstaal, W., Schacter, D. L., Johnson, M. K., & Galluccio, L. (1999). Facilitation and impairment of event memory produced by photograph review. *Memory & Cognition*, *27*, 478–493.

MacLeod, C. M. (1998). Directed forgetting. In J. M. Golding & C. M. MacLeod (Eds.), *Intentional forgetting: Interdisciplinary approaches* (pp. 1–59). Mahwah, NJ: Lawrence Erlbaum Associates.

MacLeod, C. M. (1999). The item and list methods of directed forgetting: Test differences and the role of demand characteristics. *Psychonomic Bulletin & Review*, *6*, 123–129.

Perfect, T. J., Moulin, C. J. A., Conway, M. A., & Perry, E. (2002). Assessing the inhibitory account of retrieval-induced forgetting with implicit-memory tests. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*, 1111–1119.

Racsmány, M., & Conway, M. A. (2006). Episodic inhibition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*, 44–57.

Richardson-Klavehn, A., & Gardiner, J. M. (1996). Cross-modality priming in stem completion reflects conscious memory, but not voluntary memory. *Psychonomic Bulletin & Review*, *3*, 238–244.

Tulving, E. (1985). Memory and consciousness. *Canadian Psychology*, *26*, 1–12.

Tulving, E. (2002). Chronesthesia: Awareness of subjective time. In D. T. Stuss & R. C. Knight (Eds.), *Principles of frontal lobe functions* (pp. 311–325). New York: Oxford University Press.

IV.3. The role of goal in retrieval inhibition (Study 3)

Paper: Mirroring Intentional Forgetting in a Shared-Goal Learning Situation. Mihály Racsmány, Attila Keresztes, Péter Pajkossy, Gyula Demeter, *PLoS ONE* 01/2012; 7(1):e29992. DOI:10.1371/journal.pone.0029992

OPEN ACCESS Freely available online

PLoS one

Mirroring Intentional Forgetting in a Shared-Goal Learning Situation

Mihály Racsmány[1,2]*, Attila Keresztes[1]❡, Péter Pajkossy[1]❡, Gyula Demeter[1]❡

1 Department of Cognitive Science, Budapest University of Technology and Economics, Budapest, Hungary, **2** Institute of Psychology, University of Szeged, Szeged, Hungary

Abstract

Background: Intentional forgetting refers to the surprising phenomenon that we can forget previously successfully encoded memories if we are instructed to do so. Here, we show that participants cannot only intentionally forget episodic memories but they can also mirror the "forgetting performance" of an observed model.

Methodology/Principal Findings: In four experiments a participant observed a model who took part in a memory experiment. In Experiment 1 and 2 observers saw a movie about the experiment, whereas in Experiment 3 and 4 the observers and the models took part together in a real laboratory experiment. The observed memory experiment was a directed forgetting experiment where the models learned two lists of items and were instructed either to forget or to remember the first list. In Experiment 1 and 3 observers were instructed to simply observe the experiment ("simple observation" instruction). In Experiment 2 and 4, observers received instructions aimed to induce the same learning goal for the observers and the models ("observation with goal-sharing" instruction). A directed forgetting effect (the reliably lower recall of to-be-forgotten items) emerged only when models received the "observation with goal-sharing" instruction ($P < .001$ in Experiment 2, and $P < .05$ in Experiment 4), and it was absent when observers received the "simple observation" instruction ($P > .1$ in Experiment 1 and 3).

Conclusion: If people observe another person with the same intention to learn, and see that this person is instructed to forget previously studied information, then they will produce the same intentional forgetting effect as the person they observed. This seems to be a an important aspect of human learning: if we can understand the goal of an observed person and this is in line with our behavioural goals then our learning performance will mirror the learning performance of the model.

Citation: Racsmány M, Keresztes A, Pajkossy P, Demeter G (2012) Mirroring Intentional Forgetting in a Shared-Goal Learning Situation. PLoS ONE 7(1): e29992. doi:10.1371/journal.pone.0029992

Editor: Jan Lauwereyns, Kyushu University, Japan

Received April 27, 2011; **Accepted** December 9, 2011; **Published** January 5, 2012

Funding: 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 (Budapest University of Technology and Economics, Hungary)" project and is supported by the New Hungary Development Plan (Project ID: TÁMOP-4.2.1/B-09/1/KMR-2010-0002) and by OTKA (Hungarian National Science Foundation) K84019. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: racsmany@cogsci.bme.hu

❡ These authors contributed equally to this work.

Introduction

A flexible memory needs a mechanism by which it can disregard earlier encoded information that is no longer reliable, is irrelevant or even disturbing. The experimental procedure called directed forgetting (DF) demonstrates this relevant aspect of human memory. In a typical directed forgetting experiment participants first learn a set of items, usually a list of words (henceforth: List 1), then receive an instruction either to forget or to remember these items. This paradigm is called the list method of directed forgetting and studies using this procedure demonstrated that following learning of further items (henceforth: List 2), participants can recall significantly fewer of the items designated to be forgotten compared to those that were to be remembered [1–4]. The experimental work of the last thirty years has revealed many attributes of the DF effect and the brain mechanisms involved in this phenomenon have also become clear [5,6,7–13].

The dominant theory of directed forgetting was framed by Bjork [1] who suggested that the forget instruction elicits a process in participants which suppresses the access of List 1 items, although this process is modulated by factors such as list segregation and recall output order. According to Bjork [1] the suppression of List 1 items serves an adaptive goal for participants to escape from proactive interference while studying List 2 items (see Racsmány and Conway [14] for an extension of this concept to episodic retrieval). This idea was supported by experimental results showing that recall performance of List 2 items is significantly higher following a forget instruction than following a remember instruction of List 1 items, although this beneficial effect of forget instruction has not been present constantly in directed forgetting experiments (see[2]). The suppression theory of directed forgetting received strong support both from neuroimaging studies and from investigations of patients suffering from brain damage or psychiatric disorders. For instance, Mecklinger, Para and Waldhauser [15] showed that successful forgetting in a directed forgetting experiment elicited a right frontal activation following the forget instruction. This brain area – and especially the right inferior frontal gyrus - is associated to inhibition of prepotent responses [15,16] (see also [6]). Bäuml, Hanslmayr, Pastötter and Klimesch [17] showed that forget instruction induces a change in alpha oscillations which is assumed to be an active neural inhibitory filter. Furthermore, patients with lesion in the right frontal cortex and patients diagnosed with schizophrenia – known to have frontal dysfunctions [18] - were unable to produce a directed forgetting effect [11,19].

An alternative explanation of directed forgetting was proposed by Sahakyan and Kelley [20] who suggested that the forget instruction produce a change in mental context of participants and this change serves as a key factor for later recall patterns. According to this explanation, directed forgetting is just another example of context dependent memory phenomenon. Participants in the forget group change their internal context as a response to the forget instruction, therefore they are studying List 2 items in a changed mental context and finally they try to recall List 1 and List 2 items in this new mental context. In contrast, participants who receive a remember instruction will learn both lists in the same internal context. Sahakyan and Delaney [21] suggested that only the cost of directed forgetting (the decreased List 1 recall performance in the forget group) is explained by contextual change, while other factors, such as changed learning strategy, contribute to the benefit of forget instruction (the higher recall of List 2 items in the forget group). The results of these experiments gave evidence that instructing participants to intentionally change their mental context produced the same level of forgetting of List 1 items as the 'standard' forget instruction (see [22]).

A fundamental difference between these two concepts of directed forgetting is the role of participant's goal in the causal explanation of the phenomenon. According to the framework of Bjork [1], suppression of the first list is a goal-related response to the forget instruction, where the goal of the participant is to learn valid and disregard invalid information. In contrast, the context change hypothesis [20] proposes that the suppression of the first list is a side effect of the instruction. The forget instruction segregates the two learning lists and creates different contexts for them, however the goal of the participant does not play a causal role in this process. It is difficult to discriminate the predictions of the two explanatory concepts in the standard directed forgetting procedure, because we should manipulate independently the goal of participants and the type of instructions they receive. However, the type of the instruction always determines the goal of the participant, thus these two factors are strongly associated in the standard DF procedure. We can discriminate these two factors, if participants are not directly instructed, but observe another person, a model, who receive a forget instruction. This way it is possible to manipulate independently the goal of the observer (congruent or incongruent with the goal of the model) and the type of instruction (forget or remember) given to the model.

Dissociating goal and instruction is also fruitful from a more general point of view. The directed forgetting procedure is a paradigmatic case of intentional learning, where a learner has to keep relevant information in an active form while has to suppress irrelevant information. From the perspective of an adaptive cognitive system we can assume that participants are able to produce an intentional suppression of successfully studied information by detecting which information is relevant and which is irrelevant for an observed model. By applying the directed forgetting procedure in an observational learning task, where the relevant information must be extracted from the interaction of the experimenter and the observed model, it is possible to get evidence for the adaptiveness of intentional forgetting.

The central question of the present study was whether or not observers were able to mirror the learning performance of an observed model who had received a forget instruction. Considering the learning process as a specific action, we aimed to investigate the role of the observer's goal in activating and suppressing memories. In research on action understanding there are many observations of an action eliciting the same brain activity pattern in motor planning areas as the actual execution of that same action [23–26]. Moreover, studies using various stopping paradigms have demonstrated that the observers mirrored inhibitory attention processes along with the perceived person's action [27,28]. However, so far there has been no demonstration of mirroring explicit goal-related memory access.

According to our hypothesis, observers can mirror the intentional forgetting performance of an observed model, but only if they share the same goal in the learning situation. If the observers' goal is simply to observe the behaviour of the observed model, they will not mirror intentional forgetting; therefore, they will remember the to-be-forgotten information. We assume that a forget instruction elicits suppression of earlier encoded information only if this instruction targets goal relevant information for the observer.

We developed a modified version of the DF procedure aimed at investigating whether or not participants are able to simulate the intentional forgetting performance of a model. In this experimental procedure, called observational directed forgetting (oDF), participants (the observers) observe another person (the model) taking part in a directed forgetting experiment.

Methods

We have obtained ethics approval for our study from the ethics committee of the Budapest University of Technology and Economics, Hungary, all participants gave written consent.

Experiment 1 & 2

In two consecutive experiments, a total of 200 native Hungarian speakers were recruited from the Budapest University of Technology and Economics student population. They received course credits for their participation. One hundred participants (45 males and 55 females) took part in each experiment, their ages varied between 19 and 26 years.

In both experiments, participants (referred to as observers throughout the article) watched a movie of a directed forgetting experiment. In this movie, a model learnt a list of words (List 1), then received a midlist instruction (forget or remember), then learnt another list of words (List 2). In both experiments, observers were randomly assigned to either the forget or the remember group.

The two experiments differed only in the instruction given to the observers prior to watching the movie. In Experiment 1, they were told simply to observe everything they saw in order to remember it later on ("simple observation"), whereas in Experiment 2 observers were told to observe everything they saw in order to remember what the model in the movie had to remember ("observation with goal sharing").

In the movie presented to the observers, a male model sat in front of a computer screen and was told by an experimenter that he would be presented with a list of words and that his task was to learn all of the words for a later memory test. Each word was displayed for 2 s with a 2-s inter-item interval. When filming the movies we used two experimental learning lists (List A and List B) consisting of 12 words of moderate to high frequency. Half of the observers saw a version of the movie in which List A served as List 1, and List B served as List 2, while the other half of the observers saw a version in which List A served as List 2 and List B as List 1. After List 1 had been presented on the screen the experimenter gave either a "forget" or a "remember" instruction. In the "forget" condition, the model in the movie received the instruction that the words presented up until this point were only presented by mistake, and the experimenter asked the model to try to forget these words in order to properly carry out the learning of the following words. Following the forget instruction the model was presented with a second list of 12 words. In the "remember" condition, the experimenter in the movie gave a remember instruction following List 1; that is, he asked the model to remember the words presented up until that point and to try to learn the words in the second list as well. Following the presentation of List 2, the experimenter thanked the model for their contribution.

Following the presentation of the movie, observers took part in a distractor task in which they solved simple arithmetic tasks for 10 min. Then they were asked to recall all the words that had been presented to the model in the movie. All observers were first asked to recall List 1, and then List 2 words, in order to avoid a possible output interference of List 2 words in the forget condition.

Experiment 3 & 4

In two further experiments, a total of 208 native Hungarian speakers were recruited from the Budapest University of Technology and Economics student population. One hundred-twenty participants (43 males and 77 females) took part in Experiment 3, and eighty-eight participants (39 males and 49 females) took part in Experiment 4. Their ages varied between 19 and 28 years. Data of four participants (two models and one observer) was excluded from the analysis of Experiment 3, and data of three participants (two models and one observer) was excluded from the analysis of Experiment 4, because they figured out the goal of the experiment, as it was revealed by the debriefing.

The two experiments followed the same logic as Experiment 1 and 2 with the only exception that this time the observed model was a real participant, not only an actor in a movie.

Two participants (one model and one observer) took part in the experiment at the same time. Each participant pair (observer and model) was randomly assigned to either the remember or the forget group and each member of the pair was randomly assigned to be the observer or the model in the experiment. First, the observers were informed that they would take part in a memory experiment as an observer where a model would learn lists of words for a later recall. The observer was also informed that the aim of their participation is to warm up for a later memory experiment. Similarly to Experiment 1, in Experiment 3 observers received a "simple observation" instruction; that is, their task was to watch carefully and observe everything they saw, because later they would have to remember it. Similar to Experiment 2, in Experiment 4 observers received an "observation with goal sharing" instruction; that is, their task was to watch carefully and observe everything they saw, but crucially they were also informed that at the final recall test there would be a possibility to help the model if she/he asks for it.

The model and the observer sat close to each other in front of a computer screen, in a distance from the screen so that both of them could easily read the presented stimuli. Each word was displayed for 2 s with a 2-s inter-item interval. The experimenter gave instructions only to the model, who were informed that they would be presented with a list of words and were to learn all of the words for a later memory test. After the first list of words had been presented on the screen the experimenter gave either a "forget" or a "remember" instruction to the model. In the "forget" condition the models received the instruction that the words presented up until that point were only presented by mistake, and the experimenter asked them to try to forget those in order to properly carry out the learning of subsequent words. After the forget instruction the models were presented with a second list of words. In the "remember" condition the experimenter gave a remember instruction following List 1, asking the models to remember the words presented up until that point and to try to learn the words on the second list as well.

After the presentation of List 2, both the models and the observers took part in a distractor task, solving simple arithmetical problems for 10 minutes. Then they were asked to recall all the words that had been presented to the model in the movie. All observers were first asked to recall List 1, and then List 2 words, in order to avoid a possible output interference of List 2 words in the forget condition.

Results

In all four experiments the same mixed ANOVA was carried out with instruction (Forget/Remember) as between subject variable and list (List 1/List 2) as within subject variable. In Experiment 3 and 4, recall data of models and observers were analysed separately, and when discussing these results, we report data for models first, and data for observers second.

Experiment 1

We found a significant main effect of list, $F(1,98) = 13.15$, $P < .001$, but no significant interaction between list and instruction, $F(1,98) = 0.02$, ns. Independent t-tests showed that, on average, observers in the forget group and the remember group recalled the same proportion of List 1 words, $t(99) = -0.67$, ns., and the same proportion of List 2 words, $t(99) = -0.66$, ns. This supports our hypothesis that observers with an attitude of merely observing a learning action of a model will not produce the same memory performance as the observed model; therefore, they will not produce an intentional forgetting of List 1 in the forget condition (see Figure 1, upper part of panel B).

Experiment 2

The same ANOVA as in Experiment 1 yielded a significant main effect of list, $F(1, 98) = 20.08$, $P < .001$, and more importantly, a significant interaction between list and instruction, $F(1,98) = 17.4$, $P < .001$. Independent t-tests revealed that observers in the forget group recalled fewer List 1 words, $t(99) = 2.19$, $P < .05$, $r = .22$, but more List 2 words, $t(99) = 2.83$, $P < .01$, $r = .27$, than observers in the remember group. This recall pattern shows that our manipulation was successful in inducing a directed forgetting effect. (see Figure 1, lower part of panel B).

Experiment 3

Models. The list X instruction interaction was significant, $F(1,58) = 10.56$, $P < .005$. Independent t-tests revealed that models in the forget group recalled fewer List 1 words, $t(58) = -2.67$, $P < .01$, $r = .33$, but more List 2 words, $t(58) = 1.29$, ns., than models in the remember group. Although this latter effect, the benefit of directed forgetting instruction, was not significant, our manipulation was successful in inducing a directed forgetting pattern among models (see Figure 2, upper part of panel C).

Observers. Observers showed a different pattern compared to the models they had observed. Their recall data showed no significant list X instruction, $F(1,54) = 2.54$, $P = .117$. Also,

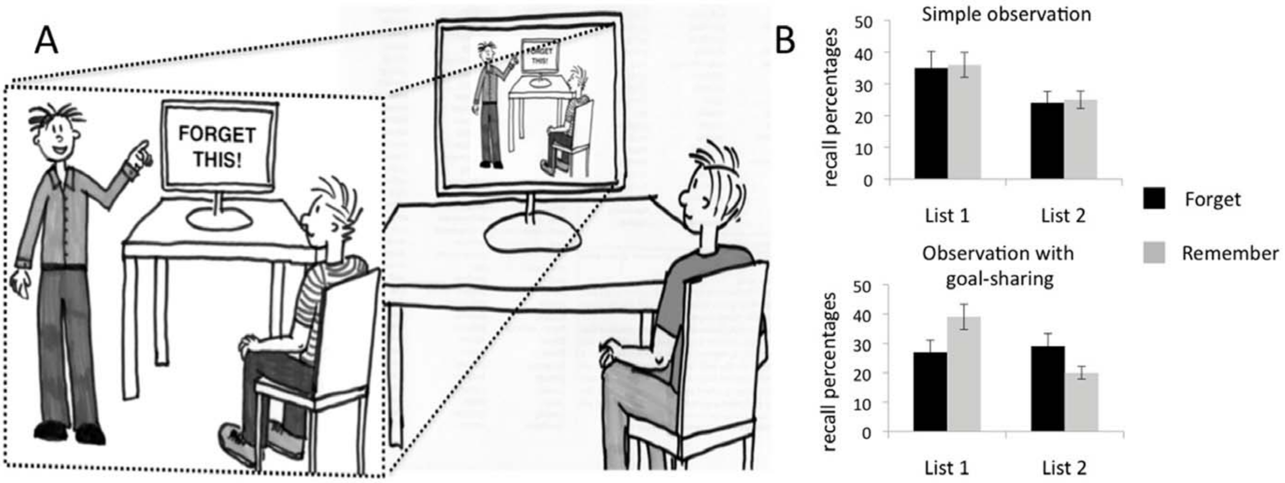


Figure 1. Experimental set-up and results of Experiment 1 and 2. (A) In both experiments the observers sat in front of a computer screen on which they saw a movie of a directed forgetting experiment. In this movie, a model was instructed to learn a list of words shown on a computer screen, and was then shown a second list that was also to be learnt. Immediately before being presented with the second list to learn, the model in this movie received a midlist instruction. Half of the observers saw a movie where the model was instructed to forget the list that they had seen before and to learn the second list (this is the forget condition shown here). The other half of the observers saw a movie where the model was instructed to remember the second list as well (the remember condition). In experiment 1 (upper part of panel B), the observers were simply told to observe the movie in order to remember as many details as possible (simple observation). Here, we found no directed forgetting effect: after watching the movie the observers recalled a similar number of words in the two conditions, $P>.1$. In experiment 2 (lower part of panel B), the observers were told to observe the movie in order to remember everything that the model in the movie had to remember (observation with goal-sharing). Here, we found a significant directed forgetting effect: after watching the movie the observers in the forget condition recalled significantly fewer words from the first list and recalled significantly more words from the second list than the observers in the remember condition, $P<.001$.
doi:10.1371/journal.pone.0029992.g001

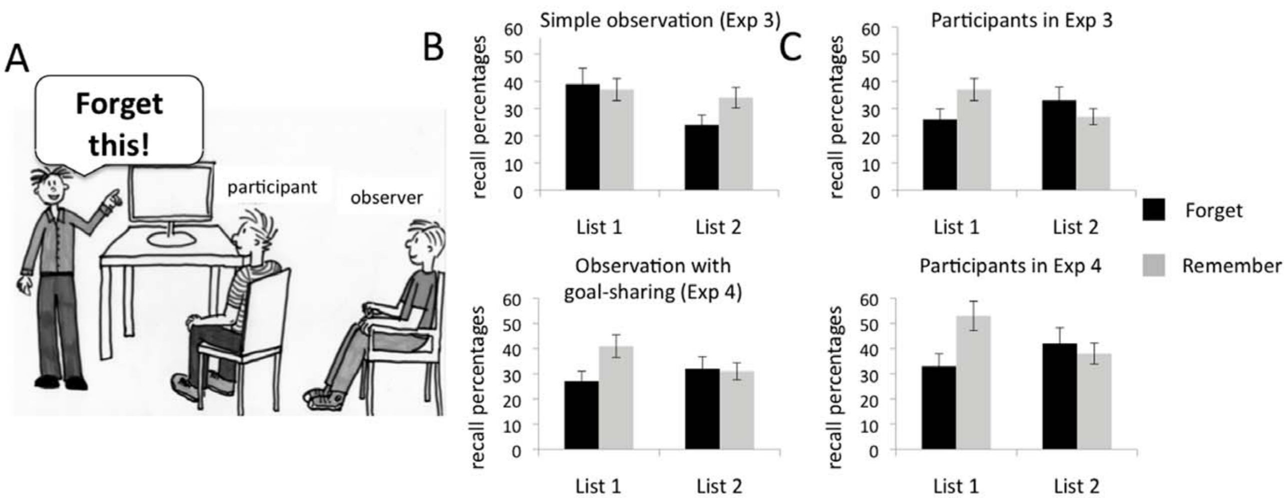


Figure 2. Experimental set-up and results of Experiment 3 and 4. (A) In both experiments participants sat in front of a computer screen and participated in a directed forgetting experiment (we refer to these participants as models and their results are shown in panel C). They were instructed to learn a list of words (List 1) shown on the screen. Immediately after being presented with List 1, the models received a midlist instruction. Half of the models was instructed to forget the list (List 1) they had seen before, and learn the second list (List 2). This is the forget condition shown here. The other half of the models was instructed to remember List 2 as well (remember condition). Models (panel C) in both Experiments showed directed forgetting. Each model was observed by another participant (we refer to these participants as observers and their result are shown in panel B). In experiment 3, observers (upper part of panel B) were told simply to observe the experiment in order to remember as many details as possible (simple observation). Here, we found no directed forgetting effect: after watching the experiment, observers in the forget and remember condition recalled a similar number of List 1 words, and observers in the remember condition recalled more words from List 2 than observers in the forget condition. In experiment 4, observers (lower part of panel B) were told to observe the experiment in order to be able to help the model in the experiment (observation with goal-sharing). Here, we found a significant directed forgetting effect: after watching the experiment, observers in the forget condition recalled significantly less words from List 1 than observers in the remember condition.
doi:10.1371/journal.pone.0029992.g002

independent t-tests revealed that observers in the forget group recalled a similar proportion of List 1 words, $t(54) = 0.43$, ns., and a lower proportion of List 2 words, $t(54) = 1.69$, ns., compared to observers in the remember group. In brief, in this group we found no directed forgetting effect (see Figure 2, upper part of panel B).

Experiment 4

Models. The list X instruction interaction was significant, $F(1,40) = 12.34$, $P<.001$. Independent t-tests revealed that models in the forget group recalled fewer List 1 words, $t(40) = -3.47$, $P<.001$, $r = .48$, but more List 2 words $t(40) = 0.51$, ns., than models in the remember group. Again, as for models in Experiment 3, although the benefit of directed forgetting instruction was not significant, our manipulation was successful in inducing a directed forgetting pattern among participants (see Figure 2, lower part of panel C).

Observers. In contrast to Experiment 3, observers in Experiment 4 showed a similar pattern as the models they had observed. Their recall data showed significant list X instruction, $F(1,41) = 4.24$, $P<.05$. Also, independent t-tests revealed that observers in the forget group recalled fewer List 1 words, $t(41) = -2.36$, $P<.05$, $r = .35$, and a similar proportion of List 2 words, $t(41) = .12$, ns., compared to observers in the remember group. In brief, although we found no benefit of the directed forgetting instruction for the forget group, the observers showed a clear directed forgetting effect (see Figure 2, lower part of panel B).

Discussion

In Experiment 1 and 2 we demonstrated that observers mirror the effect of the forget instruction given to an observed model. This mirroring only occurred when the instruction given to the observers induced shared goal representations.

Although Experiment 1 and 2 gave evidence that suppression of the to-be-forgotten items is modulated by the observer's goal, the applied instruction and the specific way of item presentation raised a series of question with respect to the above interpretation of our results. Did the instruction to remember everything that the model had to remember induce any empathy/goal sharing with the model, or did the observers simply interpret the instruction given to the model as an instruction given to them? Another problem in our interpretation might be that the model did not suppress memories (models were actors in a movie). Therefore we cannot infer that the forgetting effect produced by the observers is truly a mirrored effect.

To clarify these questions we changed both the learning situation and the goal-sharing instruction in two following experiments. In Experiment 3 and 4, instead of watching a movie about an experiment, observers observed a directed forgetting experiment in a real-life setting, with real experimental participants (models). In order to induce empathy/goal sharing of observers with the models, we changed the "observation with goal-sharing" instruction of Experiment 2 in a way to stress the shared goal of the two persons. Therefore, the observers were told that they could help the model at the final recall. We reasoned that this instruction not only induces shared goal-representations, but also rules out the possibility that observers simply interpret the instruction given to the model as an instruction they (the observers) should follow. Besides this, the real-life setting, used in Experiment 3 and 4, allowed us to match the recall pattern of observers to the recall pattern of real participants.

The results of Experiment 3 and 4 replicated the results of Experiment 1 and 2. That is, observers mirrored the effect of the forget instruction given to the observed model, but only when the instruction given to the observers induced shared goal representations.

In sum, we demonstrated that directed forgetting effect in the observer was only present if the goal to encode specific memories was the same or similar for the observer and the model. In four experiments we gave evidence that observers suppressed List 1 items if they observed a model who was instructed to forget these items. However, this effect was modulated by the instruction type given to the observers. Observers only produced the directed forgetting effect if they were instructed to share the goal of the model. This means that if the observer's goal is to acquire the same information as the model, then any environmental manipulation of the model's behaviour will influence the accessibility of the observer's memories. It is important to note that goal sharing was manipulated in two fundamentally different ways in Experiment 2 and Experiment 4. In Experiment 2 observers watched a movie about the experiment, they had no contact with the models, and because of this one could argue that observers may have not felt empathy for the models or shared the model's goal. More importantly, as the observer were instructed to remember everything that the model in the movie had to remember, this may have forced them to instruct themselves the same way as the experimenter instructed the model. However, in Experiment 4 observers took part in the same experiment as the model: they sat next to them and they followed their behaviour from close distance. This experimental design should have induced more empathy in the observers for the model. Moreover, the instruction also differed in Experiment 4. Observers were instructed that they might have the chance to help the model at the final test. This instruction probably led the observer to share the goal of the model. Although there are major differences in the observer's instructions in Experiment 2 and 4, the two experiments produced exactly the same pattern of results. This supports the conclusion that shared goal of observers and models was the critical factor in producing this observational directed forgetting effect.

A further contribution of Experiment 3 and 4 compared to Experiment 1 and 2 is that the memory performance of the model is known. The direct comparison of observers' and models' performance gave further evidence that observers mirrored the memory performance of the model in Experiment 4, while their performance was different from that of the model in Experiment 3.

In a narrower interpretation, our results provide relevant evidence for theoretical accounts of directed forgetting. The concept of retrieval inhibition [1] states that the forget instruction, together with further learning of List 2 triggers an inhibitory process in order to attenuate the interference of to-be-forgotten items with to-be-remembered items. Inhibitory processes serve an adaptive role to enhance the accessibility of reliable items and suppress all unimportant and disturbing information. In contrast, the context change hypothesis [20] proposed an account without inhibition by suggesting that participants in the forget group will create a larger than normal change in internal contextual elements, and will treat the two study lists as separate events because of the forget instruction. As a consequence, participants in the forget group will encode List 1 words in a different context than List 2 words, and there will be a contextual mismatch between List 1 and final recall. According to this concept the forget instruction plays no specific role in the directed forgetting phenomenon, and it is replaceable with any other manipulation causing a similar contextual change between the two study lists.

In our opinion the results of the present study fit better to the concept of retrieval inhibition than to the context change hypothesis. The forget instruction will carry the future importance of studied information only if it targets goal-relevant aspects of the previous event. In other words, the forget instruction will trigger inhibitory processes for to-be-forgotten information, because it informs the participant that these items are no longer relevant from the perspective of the present goal of the learner, that is, the successful recall of the studied items. An observer without the goal to recall all relevant information from the point of view of the model will not use the information of the forget instruction.

It is unclear, how the context change hypothesis could explain the present results in a parsimonious way. To explain the recall performance of the observers with this concept we should assume that observers without shared goal with the models did not create a new internal context for the second list as a response to the forget instruction. Contrary, they have changed their mental context if their goal was in accord with the model. Following the logic of this account we should assume that the "observation with goal sharing" instruction increased the encoding of contextual elements compared to the "simple observation" instruction. One problem with this explanation is that there has been no evidence for such an association between goal-directed learning and internal context encoding. Another, and more evident, problem is that enhanced contextual encoding should have lead to a higher average recall rate among observers instructed with "observation with goal sharing" instruction. This is certainly not the case.

In sum, these results underlie the general assumption that activation and suppression of episodic memory representations is based on goal-related action plans [29]. It is important to note that it has been widely documented that the suppression effect in the directed forgetting procedure lowers the accessibility, but not the availability of to-be-forgotten memories, meaning that these memory items remain intact but become inaccessible by episodic retrieval cues [2,14,30]. Our results support the assumption that suppression of episodic memories is not automatically generated by environmental cues but depends on the goals of the person who encodes and retrieves them [29]. In a broader interpretation, these results gave evidence that observers can mirror the suppression memory effect of the model if they take the model's action goals. The central question of action mirroring is whether the mechanism is a direct match between the perception of the model's action and the observer's motor system [26] or whether it is generated from goal interpretation via top-down processes [31]. Our results suggest that the mirroring of intentional forgetting takes place in the latter form. When the observer shares the model's goal, they will encode items that are relevant to the model and then they will manipulate the accessibility of their own memories according to what seems to be relevant to the model in a learning action. The exact nature of this process – whether it is an action simulation or an end state emulation by different means – is presently unclear, but our results point to a relevant aspect of social learning. Human learners manipulate the activation level of their own memory according to the specific goal of the observation, and if this goal matches the goal of the observed model than the observer will mirror the learning performance of the model.

Acknowledgments

We thank Réka Zeley for the drawings showing the experimental set-up and Ágnes Lukács for her comments on a previous version of the manuscript. We also thank Katalin Szentkuti, Luca Frankó, Dorottya Pápai, and Anna Ágó for organizing the experiments, and Bálint Forgács, Kornél Németh and Eszter Szabó for their help in making the video movies for Experiment 1 and 2.

Author Contributions

Conceived and designed the experiments: MR. Performed the experiments: MR. Analyzed the data: MR. Wrote the paper: MR. Discussed the results and commented on the manuscript: PP AK GD.

References

1. Bjork RA (1989) Retrieval inhibition as an adaptive mechanism in human memory. In: Roediger HL, Craik FIM, eds. Varieties of memory and consciousness: Essays in honour of Endel Tulving. Hilsdale, NJ: Lawrence Erlbaum Associates. pp 309–330.
2. MacLeod C (1998) Directed forgetting. In: Golding JM, MacLeod, CM, eds. Intentional forgetting: Interdisciplinary approaches. Mahwah, NJ: Lawrence Erlbaum Associates. pp 1–58.
3. Conway MA, Harries K, Noyes J, Racsmány M, Frankish CR (2000) The disruption and dissolution of directed forgetting: Inhibitory control of memory. J Mem Lang 43: 409–430.
4. Hasher L, Zacks R (1988) Working memory, comprehension, and aging: A review and a new view. In: Bower GH, ed. The psychology of learning and motivation: Advances in research and theory (vol 22). San Diego, CA: Academic Press. pp 193–225.
5. Ludowig E, Möller J, Bien C, Münte TF, Elger CE, et al. (2010) Active suppression in the mediotemporal lobe during directed forgetting. Neurobiol Learn Mem 93: 352–361.
6. Bäuml KH, Pastötter B, Hanslmayr S (2010) Binding and inhibition in episodic memory-Cognitive, emotional, and neural processes. Neurosci Biobehav R 34: 1047–1054.
7. Van Hooff JC, Whitaker TA, Ford RM (2009) Directed forgetting in direct and indirect tests of memory: Seeking evidence of retrieval inhibition using electrophysiological measures. Brain Cognition 71: 153–164.
8. Banich MT, Mackiewicz KL, Depue BE, Whitmer AJ, Miller GA, et al. (2009) Cognitive control mechanisms, emotion and memory: A neural perspective with implications for psychopathology. Neurosci Biobehav R 33: 613–630.
9. Reber PJ, Siwiec RM, Gitleman DR, Parrish TB, Mesulam MM, et al. (2002) Neural correlates of successful encoding identified using functional magnetic resonance imaging. J Neurosci 22: 9541–9548.
10. Ullsperger M, Mecklinger A, Müller U (2000) An electrophysiological test of directed forgetting: The role of retrieval inhibition. J Cognitive Neurosci 12: 924–940.
11. Conway MA, Fthenaki A (2003) Disruption of inhibitory control of memory following lesions to the frontal and temporal lobes. Cortex 39: 667–686.
12. Anderson MC, Green C (2001) Suppressing unwanted memories by executive control. Nature 410: 366–369.
13. Anderson MC, Ochsner KN, Kuhl B, Cooper J, Robertson E, et al. (2004) Neural systems underlying the suppression of unwanted memories. Science 303: 232–235.
14. Racsmány M, Conway MA (2006) Episodic inhibition. J Exp Psychol Learn 32: 44–57.
15. Mecklinger A, Parra M, Waldhauser GT (2009) ERP correlates of intentional forgetting. Brain Res 1255: 132–147.
16. Aron AR, Robbins TW, Poldrack RA (2004) Inhibition and the right inferior frontal cortex. Trends Cog Sci 8: 170–177.
17. Bäuml KH, S, Pastötter B, Klimesch W (2008) Oscillatory correlates of intentional updating in episodic memory. NeuroImage 41: 596–604.
18. Keefe RSE (2000) Working memory dysfunction and its relevance in schizophrenia. In: Sharma, T, Harvey, PD, eds. Cognitive Functioning in Schizophrenia. Oxford: Oxford University Press. pp 16–49.
19. Racsmány M, Conway MA, Garab EA, Cimmer Cs, Janka Z, et al. (2008) Disrupted memory inhibition in schizophrenia. Schiz Res 101: 218–224.
20. Sahakyan L, Kelley CM (2002) A contextual change account of directed forgetting effect. J Exp Psychol Learn Mem Cogn 28: 1064–1072.
21. Sahakyan L, Delaney PF (2003) Can encoding differences explain the benefits of directed forgetting in the list method paradigm? J Mem Lang 48: 195–206.
22. Sahakyan L, Delaney PF (2005) Directed Forgetting in Incidental Learning and Recognition Testing: Support for a Two-Factor Account. Journal of Experimental Psychology: Learning, Memory, and Cognition 31: 789–801.
23. Rizzolatti G, Fadiga L, Gallese V, Fogassi L (1996) Premotor cortex and the recognition of motor actions. Cognitive Brain Res 3: 131–141.
24. Buccino G, Binkofski F, Riggio L (2004) The mirror neuron system and action recognition. Brain Lang 89: 370–376.
25. Gallese V, Keysers C, Rizzolatti G (2004) A unifying view of the basis of social cognition. Trends Cogn Sci 8: 396–403.
26. Rizzolatti G, Craighero L (2004) The mirror neuron system. Annu Rev Neurosci 27: 169–192.
27. Welsh TN, Elliott D, Anson JG, Dhillon V, Weeks DJ, et al. (2005) Does Joe influence Fred's action? Inhibition of return across different nervous systems. Neurosci Lett 385: 99–104.
28. Frischen A, Loach D, Tipper SP (2009) Seeing the world through another person's eyes: Simulating selective attention via action observation. Cognition 111: 212–218.

29. Conway MA (2005) Memory and the self. J Mem Lang 53: 594–628.
30. Bjork EL, Bjork RA (1996) Continuing influences of to-be-forgotten information. Conscious Cogn 5: 176–196.
31. Csibra G (2007) Action mirroring and action understanding: An alternative account. In: Haggard, P, Rosetti, Y, Kawato, M, eds. Sensorimotor Foundations of Higher Cognition. Attention and Performance. XXII. Oxford: Oxford University Press. pp 435–459.

V. Executive control and retrieval accessibility of episodic memories

V.1. Disorder of executive control and memory inhibition (Study 4)

Paper: Disrupted memory inhibition in schizophrenia. Mihály Racsmány, Martin A Conway, Edit A Garab, Csongor Cimmer, Zoltán Janka, Tamás Kurimay, Csaba Pléh, István Szendi, *Schizophrenia Research* 05/2008; 101(1-3):218-24. DOI:10.1016/j.schres.2008.01.002

Available online at www.sciencedirect.com
ScienceDirect
Schizophrenia Research 101 (2008) 218–224
SCHIZOPHRENIA RESEARCH
www.elsevier.com/locate/schres

Disrupted memory inhibition in schizophrenia

Mihály Racsmány [a,b,*], Martin A. Conway [c], Edit A. Garab [a], Csongor Cimmer [d], Zoltán Janka [d], Tamás Kurimay [e], Csaba Pléh [f], István Szendi [d]

[a] *Research Group on Cognitive Science, Hungarian Academy of Sciences — Budapest University of Technology and Economics, Hungary*
[b] *Department of Psychology, University of Szeged, Hungary*
[c] *Institute of Psychological Sciences, University of Leeds, United Kingdom*
[d] *Department of Psychiatry, University of Szeged, Hungary*
[e] *Department of Psychiatry, St. John's Hospital, Budapest, Hungary*
[f] *Department of Cognitive Science, Budapest University of Technology and Economics, Hungary*

Received 27 September 2007; received in revised form 27 December 2007; accepted 4 January 2008
Available online 6 February 2008

Abstract

A feature of schizophrenia is disrupted executive function leading to learning difficulties and memory problems. In two experiments we measured the ability of patients with schizophrenia to suppress irrelevant parts of acquired information by intentional (executive) and autonomic (non-executive) strategies. In the first experiment using directed forgetting by lists patients were found to be unable to intentionally suppress recently acquired episodic memories. In a second experiment using a procedure that induces inhibition automatically schizophrenic patients showed levels of inhibition comparable to those of normal controls. These findings indicate that in schizophrenia memory is most impaired in tasks that load heavily on control or executive processes.

Keywords: Schizophrenia; Directed forgetting; Selective practice; Prefrontal inhibition; Executive function

1. Introduction

Patients diagnosed with schizophrenia show a range of neurocognitive deficits including memory malfunctions (McKenna et al., 1990; Saykin et al., 1991). Indeed, it has been frequently shown that patients with schizophrenia perform poorly on immediate and delayed verbal learning tasks, such as the Rey Auditory Verbal Learning Test or the Wechsler Memory Scale (Aleman et al., 1999; Heinrichs and Zakzanis, 1998; see Cirillo and Seidman, 2003 for a review). Patients with schizophrenia are sensitive to interference and contextual change between learning and recall and the degree of memory impairment has not been found to be related to medication or duration of illness (Sevan-Schreiber et al., 1996; Torres et al., 2004). It seems that patients within the schizophrenia spectrum are compromised in their ability to disregard irrelevant information. This observation is supported by findings demonstrating that patients with schizophrenia produce less release from proactive interference and show a usually high intrusion error rate of items from earlier sets (Chan et al.,

* Corresponding author. Research Group on Cognitive Science, Hungarian Academy of Sciences, Department of Cognitive Science, Budapest University of Technology and Economics, Stoczek u. 2., Budapest, Hungary-1111. Tel.: +36 1 463 12 69; fax: +36 1 463 10 72.
E-mail address: racsmany@cogsci.bme.hu (M. Racsmány).

0920-9964/$ - see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.schres.2008.01.002

2004; Sitskoorn et al., 2002). Moreover, studies of executive system functions have found schizophrenic patients to be disinhibited on executive tasks such as negative priming, Stroop, and the Go/No-Go motor inhibitory task (Perlstein et al., 1998; see Palmer and Heaton, 2000). Taken together these findings all point to a dysexecutive profile in which malfunctioning control processes leave processing sequences open to interference by related but task-irrelevant information. In short, the pattern shows attenuation in the effectiveness of centrally controlled inhibitory processes.

Inhibitory processes can be triggered intentionally in an explicit, conscious, attempt to avoid or suppress unwanted or irrelevant information, or they may be triggered implicitly, nonconsciously, as part of a processing sequence. Note that what is important here is the way in which inhibitory processes are triggered or initiated, there is no suggestion that the processes themselves could come under direct intentional control. Also note that there is no suggestion either that executive or control processes are conscious, we assume along with all other theorists in the area that executive processes are nonconscious but that some of their outputs may occasionally enter consciousness and that they can be initiated by conscious intentions. Thus, a person might, for example, consciously and intentionally attempt to avoid thinking about a memory or a memory may be inhibited from entering consciousness by nonconscious, incidental, process (Barnier et al., 2007; Barnier et al., 2004; Racsmány and Conway, 2006).

Two experimental paradigms that have been widely used to explore intentionally and incidentally triggered inhibition and these are, respectively, direct forgetting (DF) and retrieval practice (RP). The DF procedure has been most extensively investigated by Bjork and colleagues (see Bjork, 1989; Bjork et al., 1998). In the list-method of DF participants are explicitly instructed to intentionally forget a previously learned list. The forget instruction is then followed by a second to-be-learned list (see MacLeod, 1998 for a detailed review of list- and item DF). The critical contrast that defines the DF effect and which has been observed in many studies is lower List 1 memory performance following a forget instruction compared to List 1 memory performance following a remember instruction, (Bjork and Bjork, 1996; Conway et al., 2000). The single study of directed forgetting in patients with schizophrenia found that produced a significant DF effect, although their forgetting performance was somewhat weaker than that of control group (Müller et al., 2005). This study, however, used an item-by-item DF task, in which items are followed by an R or an F instruction. Item DF effects are considered to be mediated by rehearsal strategies rather than by inhibitory processes (Basden and Basden, 1998) and, therefore, not directly relevant to the present goal of investigating intentional and incidental inhibition. Nonetheless, we note that these patients were able to alter their rehearsal strategies, in order to produce an item DF effect, and that indicates at least partly preserved executive function.

The RP procedure has been extensively studied by Anderson and colleagues (Anderson et al., 1994; Anderson and Spellman, 1995; see for a recent review Norman et al., 2007). In the RP procedure participants study a list containing words grouped into categories. The study phase is followed by the RP phase in which participants practice recalling selected items from the study list. This yields three types of items: Rp+ items are items which have been rehearsed, Rp− items which were studied items but which not themselves been rehearsed but which originate from categories that contain an Rp+ item, and finally Nrp items which are from studied categories from which no items have been rehearsed. The standard finding is that Rp+ items are remembered to a high level, Nrp items to a reliably lower level, with Rp− items showing poorest recall. The explanation is that this effect arises because of the effect of recalling Rp+ items during the practice phase is to automatically inhibit Rp− items (see Racsmány and Conway, 2006). Importantly, a recent study (Nestor et al., 2005) using the RP procedure with 15 patients diagnosed with schizophrenia and found a normal RP effect.

The aim of the present study is then to compare these two inhibitory procedures, DF and RP, in a patient group diagnosed with schizophrenia. We expect RP performance in this group to be in the normal range and show the standard inhibitory pattern as, indeed, previous studies have found, (Müller et al., 2005; Nestor et al., 2005). If this is the case then it demonstrates that at least some incidental, automatic, inhibitory processes are intact and function normally in this group. In contrast, we predict that in DF where initiating inhibition is intentional and effortful and requires the normal functioning of executive processes then a standard pattern of inhibition will not be observed. This is because executive function in schizophrenic spectrum disorder is comprised, as it is in brain damaged patients with frontal lobe lesions who also do not show a normal pattern of directed forgetting (Conway and Fthenaki, 2003).

2. Experiment 1

2.1. Method

2.1.1. Participants

A total of thirty patients with a diagnosis of schizophrenia defined by DSM-IV (American Psychiatric

Association, 1994) and ICD-10 criteria for research (World Health Organization, 1993) took part in the experiments. Patients were selected from the outpatient clinic of the Department of Psychiatry, University of Szeged. All patients were in an early stage of the illness, currently in a stable inter-episodic state, and under antipsychotic medication. The thirty control subjects were recruited from hospital staff and community volunteers. They were evaluated with a modified structured interview (Mini International Neuropsychiatric Interview). Control participants with a personal history of psychiatric disorder or a family history of psychotic and affective spectrum disorders, history of neurological illness, any medical illness known to affect brain structure, head injury with loss of consciousness for more than 30 min, clinically significant substance abuse within the last 6 months, or any medical illness that could significantly constrain neurocognitive functions were excluded. All participants were 18 to 50 years of age, minimum 8 years in education (primary school), and able to give informed consent. The patients were excluded if they had previously undergone electroconvulsive therapy or were subject of clinically significant substance abuse.

2.1.2. Clinical and neuropsychological measures

Clinical symptoms were assessed by psychiatrists using the Positive and Negative Syndrome Scale (PANSS) (Kay et al., 1991), and the Scale for the Assessment of Negative Symptoms (SANS) (Andreasen, 1982). The demographic and clinical characteristics are shown in Tables 1 and 2.

Widely used neuropsychological tasks were employed to measure working memory and executive functions. We measured verbal working memory capacity with the Digit Span Task (Racsmány et al., 2005), we used the Visual Patterns Test (VPT, Della Sala et al., 1997) for measuring visuo-spatial working memory capacity. We assessed executive functions with the Wisconsin Card Sorting Test (WCST, Heaton et al., 1993).

2.1.3. Procedure

Participants were tested individually. They were told that they were participating in an experiment on memory with the aim to test their ability to recall words. The experiment was conducted in four phases: a list learning phase, a distracter phase, a free-recall phase, and a cued-recall phase. Words were presented visually on separate sheets of papers. After the words of the first list (7 words) had been presented, participants were instructed to stop. At this point participants in the forget-instruction-condition (F-condition) were given the following forget instruction: The list you have just learned was a practice list to familiarize with the experimental procedure. You should now forget these words, try to put them out of your mind. The real experimental list will be presented now. In the remember-instruction condition (R-condition) the same procedure was followed, but instead of the forget instruction, participants received a remember instruction: That is the end of List 1. You must try to keep those words in mind while you learn the second list which will be presented now.

Following the forget or remember instructions, the second list was presented. After all words had been studied participants were given a 5-minute simple arithmetic filler task. This was followed by the free-recall test. Participants were provided with paper and pen and asked to try to recall as many words as they could from both lists. They were asked to start at the top of the page and write each recalled word under the previous word. In order to reduce the role of output interference we followed the recall instruction of Conway et al. (2000 Experiment 7): participants were required to recall List 1 words first and then List 2 words. Following the free-

Table 1
Demographic characteristics of the subjects

	Patients		Control	
	Mean	SD (+/−)	Mean	SD (+/−)
Age	31.1	9.8	32.6	9.4
Education (years)	11.1	1.8	12.2	1.9
Full scale IQ WAIS-H	103.4	14.8	109.1	11.2

Standard deviations are presented in parentheses.

Table 2
Clinical and neuropsychological characteristics of the subjects

	Mean	SD (+/−)
Age at onset (years)	23.5	6.8
Duration of illness (years)	7.5	7.1
Relapses	4.5	4.7
PANSS		
Positive	11.4	5.5
Negative	16.4	6.2
Global	29.8	11.1
Total	57.6	19.9
SANS		
Affective	1.6	1.1
Alogia	1.5	0.9
Avolition	1.5	1.2
Anhedonia	2.1	1.2
Attention	1.4	1.1
WCST		
Categories	2.8	2.4
PE%	28.8	18.7
CLR%	37.8	26.6
Digit span	5.7	1.1
VPT	6.4	1.6

Standard deviations are presented in parentheses.

recall participants took part in a stem-cued-recall test. They were given printed list containing the first two letters of the words of first and second lists in a randomised order, and completed the stems for the previously studied words. Each subject took part both in F and R conditions and the order of conditions was counterbalanced among participants. Four lists were constructed from a pool of twenty eight (Hungarian) words of moderate to high frequency (Füredi and Kelemen, 1989). The order of presentation of the lists was counterbalanced for each participant in both conditions.

2.1.4. Results and discussion

A 2 x 2 x 2 (group x instruction x list) mixed analysis of variance (ANOVA) was conducted on individual free-recall rates. The main effect of group was significant $F(1,58)=482.9$, $p<0.001$, the main effect of instruction was not significant $F(1,58)=2.7$, $p>0.1$, and the main effect of list was also not significant $F(1,58)=2.8$, $p>0.1$. More interesting was that we found highly significant interaction between instruction and list $F(1,58)=19.42$, $p<0.001$, and also a powerful interaction between group, instruction, and list $F(1,58)=10.7$, $p<0.001$. Together this pattern shows a strong and reliable directed forgetting effect for the control group but no effect for the patient group. As can be seen from Table 3 the forget instruction had no effect on patients' performance, the only detectable change is a small, non-significant reverse directed forgetting effect in which patients recalled more List 1 words in the forget than in the remember conditions. We calculated an inhibitory index for each participant by subtracting F1 (List 1 in F-condition) performance from R1 (List 1 in R-condition) performances. As can be seen from Table 3 the inhibition score is negative for the patient group showing a reverse effect of the forget instruction.

Table 3
Mean percentages of memory performances in Experiment 1

	Patients		Control	
	Mean	SD (+/−)	Mean	SD (+/−)
Recall of F1 words	35.7	17.1	42.8	24.3
Recall of F2 words	31.4	21,4	62.9	19.9
Recall of R1 words	34.3	18.6	61.4	20
Recall of R2 words	24.3	17.4	40	22.1
Inhibitory index for list-cued recall (R1 − F1)	−9.6	18.2	18.4	13.3
Stem-cued recall of F1	55.7	19.2	62.9	21.1
Stem-cued recall of F2	54.3	19.8	72.9	19.7
Stem-cued recall of R1	51.4	20.1	72.7	19.6
Stem-cued recall of R2	45.7	21.1	62.6	19.4
Inhibitory index for stem-cued recall (R1 − F1)	−6.1	14.1	10.4	11.2

Standard deviations are presented in parentheses.
F1 = List 1 words in the forget condition; F2 = List 2 words in the forget condition; R1 = List 1 words in the remember condition; R2 = List 2 words in the remember condition.

An identical 2 x 2 x 2 (group x instruction x list) analysis of variance (ANOVA) was carried out with the stem-cued-recall rates. The pattern of results was the same as for the previous the list-cued-recall task, the main effect of group was significant $F(1,58)=14.75$, $p<0.001$, the main effect of instruction was not significant $F(1,58)=1.7$, $p>0.1$, and the main effect of list was also not significant $F(1,58)=0.57$, $p>0.1$. However, we found again a significant interaction between instruction and list $F(1,58)=8.4.42$, $p<0.01$, and also a powerful interaction between group, instruction, and list $F(1,58)=6.2$, $p<0.01$.

An inhibitory index was again calculated by subtracting List 1 performances in the forget condition from List 1 performance in the remember condition. As can be seen from Table 3 patients with schizophrenia did not produce inhibition at all in either recall test and their inhibitory index was negative reflecting a rebound effect. A one-way ANOVA was carried out on individual's inhibitory indexes which yielded significant differences between groups both for the free-recall $F(1,58)=8.51$, $p<0.01$ and for the stem-cued-recall tasks $F(1,58)=7.34$, $p<0.01$. Overall this pattern of data demonstrates that patients with schizophrenia are not able to intentionally inhibit previously acquired information.

3. Experiment 2

The aim of this second experiment was to establish whether those patients who produced no inhibition in DF task were nonetheless able to produce inhibition in the RP task. Such a result would replicate the main findings of Nestor et al. (2005), and confirm our hypothesis that patients with schizophrenic spectrum disorder cannot initiate inhibition intentionally but are able to initiate inhibition when it is an incidental part of a task.

3.1. Method

3.1.1. Procedure and materials

The same patients and controls who took part in Experiment 1 took part in the present experiment and the data was collected from participants individually. The order of the DF and RP experiments was counterbalanced among participants; there was minimum one day delay between the two experiments. The RP procedure was conducted in four phases, following the procedure of Racsmány and Conway (2006): Study Phase, RP Phase, Distracter Phase, and a surprise cued-recall phase. Ten categories were used two of which were fillers. Each category consisted of twelve exemplars from each of eight target categories forming two subsets (six items) with moderate to high frequency words drawn from two published Hungarian frequency norms (Kónya and Pintér, 1985; Füredi and Kelemen, 1989). We created two subsets from the eight target categories and designated an equal number of items as practiced and nonpracticed categories. The practiced and nonpracticed exemplars were counterbalanced as well. In the study phase participants saw category-exemplar pairs on a screen and were told to try to remember the category examples as best as they could. Each category-exemplar pair was presented in uppercase letters at the centre of the screen for 5 s. When participants had completed the learning phase, the experimenter distributed practice booklets. Each page in the booklet contained one of the category names they had seen previously and the first two letters of one of the members of that category which they had to complete. Their task was to complete the exemplar fragment with one of the words they had studied earlier. Participants were told that some of the examples might be tested more than once but in every case they should complete the word stem with a word studied previously (note that only a single response was possible for each word stem). After the RP phase booklets were collected and participants were given an unrelated mathematical task for 12 min. Finally, participants were given recall booklets with the name of one of the previously studied categories on the top of each page. Participants had 10 min to recall as many examples as they could, and they had to keep the order of categories as they were arranged in the booklet. Order of presentation of category cues was counterbalanced over participants.

3.1.2. Result and discussion

A 2x3 (group x item type) mixed analysis of variance was performed on individual recall percentages. The main effect of group was significant $F(1,58)=44.143$, $p<0.001$. The main effect of list was also significant $F(1,58)=241.2$, $p<0.001$, however and more importantly there was no significant group by item type interaction $F(1,58)=2.1$, $p>0.1$. To detect specific effects of RP in both groups separate one-way analyses of variance on item types (Rp+, Rp−, Nrp items) were conducted for both the patient and the control groups. The findings replicated the results of Anderson et al. (1994), and a reliable effect of item type was observed, $F(1, 29)=165.1$, $p<0.001$ for the control group. Planned comparisons showed that the recall of Rp+ items was significantly higher than that of Nrp items, $F(1,29)=12.7$, $p<.001$, confirming the benefits of practice on subsequent recall. The recall of Rp− items was found to be significantly lower than that of Nrp items, $F(1,29)=5.8$, $p<.001$, indicating inhibition of these items.

Table 4
Mean percentages of memory performances in Experiment 2

	Patients		Control	
	Mean	SD (+/−)	Mean	SD (+/−)
Recall of Rp+ words	57.5	19.1	80	11.9
Recall of Rp− words	15.8	15,8	30	14.25
Recall of Nrp words	20.4	16.8	44.2	13
Inhibitory index (Nrp−Rp−)	4.6	10.3	12.2	11.2

Standard deviations are presented in parentheses.

Importantly the same pattern of results was present in the patient group where the main effect of item type was significant $F(1, 29)=91.6$, $p<0.001$. Planned comparisons showed that the recall of Rp+ items was significantly higher than that of Nrp items, $F(1,29)=11.9$, $p<.001$, showing that the patients too benefited from retrieval practice. The critical finding was, however, that recall of Rp− items was significantly lower than that of Nrp items, $F(1,29)=5.3$, $p<.01$, indicating the standard inhibitory effect. As in Experiment 1 inhibition scores were calculated and the patient group scores although showing the standard inhibitory effect (see Table 4) were nonetheless reliably lower than those of the control group, $F(1,58)=6.69$, $p<0.01$. Thus, normal but weaker inhibition was found in the patient's with schizophrenic spectrum disorder.

4. General discussion

The present study demonstrates that patients with schizophrenia are not able to intentionally forget items from a previously acquired list, and compared to controls patients produced no DF effect. Patients with schizophrenia are not then able to intentionally initiate inhibition. In marked contrast our patients produced a strong and reliable inhibitory effect in the RP procedure (a finding highly consistent with Nestor et al., 2005). Despite this normal pattern of recall in the RP procedure the general level of their inhibitory index was somewhat lower than that of control subjects, perhaps indicating a more widespread memory problem, in addition to problems with intentional forgetting.

Over the two experiments the pattern of performance is very similar to that reported by Conway and Fthenaki (2003) in a group of patients with frontal and temporal lobe lesions. Frontal patients produced an inverted DF effect and normal RP effect (as did the patients in the present study), while temporal lobe patients produced the reverse pattern. Conway and Fthenaki (2003) argue that the actual process of inhibition is the same in both kinds of task and it is the way this process is triggered that differs. In DF inhibition is intentionally elicited by active thought avoidance, a process carried out mainly by networks of the lateral prefrontal cortex and its connections (Anderson et al., 2004; Bunge et al., 2001; Aaron et al., 2004; Wylie et al., in press). Although prefrontal cortex may also have an important role in successful inhibition in RP paradigm, hippocampal and temporal networks can apply inhibitory processes without top-down executive control (see Norman et al., in press, for an interesting neural network model of this).

We suggest then that the pattern of performance by the patients with schizophrenia in DF may be a sign of disrupted frontal function possibly associated with attenuation of fronto-temporal pathways that, under normal circumstances, would mediate inhibition of recently acquired knowledge. In contrast, the intact, albeit somewhat weaker, inhibitory pattern in RP may reflect functioning medial temporal lobe inhibitory processes. In this case the practice phase induces inhibition by establishing retrieval competition between practiced items and unpracticed items from the same category that compete for recall during the category cued practice phase. In this way a pattern of activation and inhibition is created over the contents of a memory of the study list, with some items highly active (Rp+), some active but at a lower level (Nrp), and some inhibited (Rp−). It is this pattern that mediates recall (see Racsmány and Conway, 2006) and leads to the normal pattern of cued recall in the patients and controls.

In conclusion, the present experiments indicate that possible disrupted executive functions may considerably weaken the ability of patients with schizophrenic spectrum disorder to intentionally avoid recent memories and, perhaps, other cognitions too. This can occur even when other incidentally initiated inhibitory processes appear to function relatively normally. The wider consequences for schizophrenic cognition more generally are negative and one implication is that of weakened intentional control of a wide range of recently acquired material.

Role of funding source

Funding for this study was provided by the grants of OTKA (Hungarian Scientific Research Fund), the ETT IV/93/2003 (Hungarian Ministry of Health), and the RES-051-27-0127 (Economic and Social Research Council of Great Britain); these organizations had no further role in study design; in the collection; analysis and interpretation of data; in the writing of the report; and in the decision to submit the paper for publication.

Contributors

Authors Racsmány and Conway designed the study, undertook the statistical analysis and wrote the manuscript. Author Garab managed the experimental work. Authors Janka, Kurimay, Cimmer and Szendi managed the psychiatric and neuropsychological diagnosis. Author Pléh managed the literature searching. All authors contributed to and have approved the final manuscript.

Conflict of interest

Hereby all authors declare that they have no conflicts of interest.

Acknowledgements

This study was supported by the grants of OTKA (Hungarian Scientific Research Fund) T 046152/2004, T 034814, 49 840-T53 and K68463, and ETT IV/93/2003 (Hungarian Ministry of Health). Martin A. Conway is supported by a Professorial Fellowship, RES-051-27-0127, from the Economic and Social Research Council of Great Britain. Mihály Racsmány is a grantee of the Bolyai János Research Scholarship of the Hungarian Academy of Science.

References

Aaron, A.R., Robbins, T.W., Poldrack, R.A., 2004. Inhibition and the right inferior frontal cortex. Trends. Cog. Sci. 8 (4), 170–177.

Aleman, A., Hijman, R., de Haan, E.H.F., Kahn, R.S., 1999. Memory impairment in schizophrenia: a meta-analysis. Am. J. Psychiatry 156, 1358–1366.

American Psychiatric Association, 1994. Diagnostic and Statistical Manual of Mental Disorders (4th edn) (DSM-IV). APA, Washington, DC.

Anderson, M.C., Spellman, B.A., 1995. On the status of inhibitory mechanisms in cognition: memory retrieval as a model case. Psychol. Rev. 102, 68–100.

Anderson, M.C., Bjork, E.L., Bjork, R.A., 1994. Remembering can cause forgetting: retrieval dynamics in long-term memory. J. Exp. Psychol. [Learn. Mem.Cog.] 20, 1063–1087.

Anderson, M.C., Ochsner, K.N., Kuhl, B., Cooper, J., Robertson, E., Gabrieli, S.W., Glover, G.H., Gabrieli, J.D.E., 2004. Neural systems underlying the suppression of unwanted memories. Science, 303, 232–235.

Andreasen, N.C., 1982. Negative symptoms in schizophrenia. Arch. Gen. Psychiatry 39, 784–788.

Barnier, A.J., Hung, L., Conway, M.A., 2004. Retrieval-induced forgetting of emotional and unemotional autobiographical memories. Cogn. Emot. 18, 457–477.

Barnier, A.J., Conway, M.A., Mayoh, L., Speyer, J., Avizmil, O., Harris, C.B., 2007. Directed forgetting of recently recalled autobiographical memories. J.Exp.Psychol. [Gen]. 136, 301–322.

Basden, B.H., Basden, D.R., 1998. Directed forgetting: a contrast of methods and interpretations. In: Golding, J.M., MacLeod, C.M. (Eds.), Intentional Forgetting: Interdisciplinary Approaches. Lawrence Erlbaum Associates, Mahwah, NJ, pp. 139–173.

Bjork, R.A., 1989. Retrieval inhibition as an adaptive mechanism in human memory. In: Roediger, H.L., Craik, F.I.M. (Eds.), Varieties of Memory and Consciousness: Essays in Honour of Endel Tulving. Lawrence Erlbaum, Hillsdale, NJ, pp. 309–330.

Bjork, E.L., Bjork, R.A., 1996. Continuing influences of to-be-forgotten information. Conscious. Cog. 5, 176–196.

Bjork, E.L., Bjork, R.A., Anderson, M.C., 1998. Varieties of goal-directed forgetting. In: Golding, J.M., MacLeod, C.M. (Eds.), Intentional Forgetting: Interdisciplinary Approaches. Lawrence Erlbaum, Mahwah, NJ, pp. 103–139.

Bunge, S.A., Ochsner, K.N., Desmond, J.E., Glover, G.H., Gabrieli, J.D.E., 2001. Prefrontal regions involved keeping information in and out of mind. Brain 124, 2074–2086.

Chan, R.C.K., Chen, E.Y.H., Cheung, E.F.C., Cheung, H.K., 2004. Eur. Arch. Psychiatry Clin. Neurosci. 254, 256–262.

Cirillo, M.A., Seidman, L.J., 2003. Verbal declarative memory dysfunction in schizophrenia: from clinical assessment to genetics and brain mechanisms. Neuropsychol. Rev. 20, 43–77.

Conway, M.A., Fthenaki, A., 2003. Disruption of inhibitory control of memory following lesions to the frontal and temporal lobes. Cortex 39, 667–686.

Conway, M.A., Harries, K., Noyes, J., Racsmány, M., Frankish, C.R., 2000. The disruption and dissolution of directed forgetting: inhibitory control of memory. Mem. Lang. 43, 409–430.

Della Sala, S., Gray, C., Baddeley, A.D., Wilson, L., 1997. Visual Pattern Matrix Test. Thames Valley, Bury St. Edmunds.

Füredi, M., Kelemen, J., 1989. A mai magyar nyelv szépprózai gyakorisági szótára [A Frequency Dictionary of the Literary Language of Hungarian]. Akadémiai, Budapest.

Heaton, R.K., Chelune, G.J., Talley, J.L., Kay, G.G., Curtiss, G. (Eds.), 1993. Wisconsin Card Sorting Test Manual: Revised and Expanded. Psychological Assessment Resources, Odessa, FL.

Heinrichs, R.W., Zakzanis, K.K., 1998. Neurocognitive deficit in schizophrenia: a quantitative review of the evidence. Neuropsychol. 12, 426–445.

Kay, S.R., Opler, L.A., Spitzer, R.L., Williams, J.B., Fiszbein, A., Gorelick, A., 1991. SCID-PANSS: two-tier diagnostic system for psychotic disorders. Compr. Psychiatry 32, 355–361.

Kónya, A., Pintér, G., 1985. Kategória norma a verbális emlékezet vizsgálatához [Category norms for verbal memory research]. Hung. Psychol. Rev. 2, 93–111.

MacLeod, C.M., 1998. Directed forgetting. In: Golding, J.M., MacLeod, C.M. (Eds.), Intentional Forgetting: Interdisciplinary Approaches. Lawrence Erlbaum, Mahwah, NJ, pp. 1–59.

McKenna, P.J., Tamlyn, D., Lund, C.E., Mortimer, A.M., Hammond, S., Baddeley, A.D., 1990. Amnesic syndrome in schizophrenia. Psychol. Med. 20, 967–972.

Müller, U., Ullsperger, M., Hammerstein, E., Sachweh, S., Becker, T., 2005. Directed forgetting in schizophrenia: prefrontal memory and inhibition deficits. Eur. Arch. Psychiatry. Clin. Neurosci. 255, 251–257.

Nestor, P.G., Piech, R., Allen, C., Niznikiewicz, M., Shenton, M., McCarley, R.W., 2005. retrieval-induced forgetting in schizophrenia. Schizophr. Res. 75, 199–209.

Norman, K., Newman, E.L., Detre, G.J., 2007. A neural network model of retrieval-induced forgetting. Psychol. Rev. 114, 887–953.

Palmer, B.W., Heaton, R.K., 2000. Executive dysfunction in schizophrenia. In: Sharma, T., Harvey, P. (Eds.), Cognition in Schizophrenia: Impairments, Importance and Treatment Strategies. Oxford University Press, Oxford, pp. 51–73.

Perlstein, W.M., Carter, C.S., Barch, D.M., Baird, J.W., 1998. The Stroop task and attention deficits in schizophrenia: a critical evaluation of card and single-trial stroop methodologies. Neuropsychol. 12, 414–425.

Racsmány, M., Conway, M.A., 2006. Episodic inhibition. J.Exp. Psychol.[Learn. Mem.Cog.]. 32, 44–57.

Racsmány, M., Lukács, A., Németh, D., Pléh, Cs., 2005. A verbális munkamemória magyar nyelvű vizsgálóeljárásai. [Hungarian diagnostic tools of verbal working memory functions] Hung. Psychol. Rev. 4, 479–505.

Saykin, A.J., Gur, R.C., Gur, R.E., Mozley, P.D., Mosley, L.H., Resnick, S.M., Kester, D.B., Stafiniak, P., 1991. Neuropsychological function in schizophrenia: selective impairment in memory and learning. Arch. Gen. Psychiatry 48, 619–624.

Sevan-Schreiber, D., Cohen, J.D., Steingard, S., 1996. Schizophrenic deficits in the processing of context: a test of a theoretical model. Arch. Gen. Psychiatry 53, 1105–1112.

Sitskoorn, M.M., Nuyen, J., Appels, M.C.M., van der Vee, N.J.A., Kahn, R.S., 2002. Release in proactive inhibition in schizophrenia and its potential as a genotypic marker. J. Clin. Exp. Neuropsychol. 24, 67–81.

Torres, I.J., O1Leary, D.S., Andreasen, N.C., 2004. Symptoms and interference from memory in schizophrenia: evaluation of Frith's model of willed action. Scizophr. Res. 58, 101–115.

World Health Organization, 1993. The ICD-10 Classification of Mental and Behavioural Disorders: Diagnostic Criteria for Research. WHO, Geneva.

Wylie, G.R., Foxe, J.J., Taylor, T.L. in press. Forgetting as an active process: an fMRI investigation of item-method-directed-forgetting. Cerebral Cortex.

V.2. Executive disorder and retrieval-induced forgetting (Study 5)

Paper: Obsessed not to forget: Lack of retrieval-induced suppression effect in obsessive-compulsive disorder. Gyula Demeter, Attila Keresztes, András Harsányi, Katalin Csigó, Mihály Racsmány, *Psychiatry Research* 04/2014; 218(1):153-160. DOI:10.1016/j.psychres.2014.04.022

Psychiatry Research 218 (2014) 153–160

ELSEVIER

Contents lists available at ScienceDirect

Psychiatry Research

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

Obsessed not to forget: Lack of retrieval-induced suppression effect in obsessive-compulsive disorder

Gyula Demeter [a,*,1], Attila Keresztes [a,1], András Harsányi [b], Katalin Csigó [b], Mihály Racsmány [a]

[a] Department of Cognitive Science, Budapest University of Technology and Economics, Budapest, Hungary
[b] Department of Psychiatry, Nyírő Gyula Hospital, National Institute of Psychiatry and Addictions, Budapest, Hungary

ARTICLE INFO

Article history:
Received 4 July 2013
Received in revised form
16 December 2013
Accepted 8 April 2014
Available online 16 April 2014

Keywords:
Memory
Retrieval-induced forgetting
Executive function
Cognitive control
Inhibition
Interference
OCD

ABSTRACT

The aim of the present study was to investigate the role of executive functions in resolving memory interference in a clinical sample of patients with obsessive-compulsive disorder (OCD). Retrieval of memories has been shown to involve some form of executive act that diminishes the accessibility of rival memory traces, leading to retrieval-induced forgetting (RIF). These executive control processes might suppress unwanted thoughts and irrelevant memories during competitive retrieval. We assessed RIF with the retrieval practice paradigm among 25 OCD patients and 25 healthy controls matched for age and education. Retrieval of target memories led to enhancement of target memory recall in both groups, but suppression of related memories (RIF) occurred only among controls. Our results suggest that suppression of irrelevant, interfering memories during competitive recall is impaired in OCD.

1. Introduction

Obsessive-compulsive disorder (OCD) is a highly debilitating neuropsychiatric condition characterized by intrusive unwanted thoughts and/or repetitive, compulsive behavior or mental rituals (American Psychiatric Association, 2000).

The cognitive profile of the disorder is marked by the deficit of executive functions (Olley et al., 2007; Rao et al., 2008; Cavedini et al., 2010). However, some studies of OCD patients have found intact performance on traditional executive neuropsychological tasks (for reviews see Greisberg and McKay, 2003; Kuelz et al., 2004; Chamberlain et al., 2005; Abramovitch et al., 2013).

According to Chamberlain et al. (2005) failures of cognitive and behavioral inhibition could also explain many of the relevant clinical symptoms as well as executive deficits observed on tasks requiring inhibition of prepotent responses, set-shifting, and inadequate strategy use in memory tasks. Lesion and functional neuroimaging studies (e.g., De Bruin et al., 1983; Bokura et al., 2001; Chudasama and Robbins, 2003; Aron et al., 2004) suggest that abnormalities in the lateral orbitofrontal loop might lead to inhibitory dysfunctions. In OCD there is evidence for the hyperactivity of the lateral orbitofrontal (lOFC) and the dorsal anterior cingulate cortex (dACC) and the hypoactivity of the medial orbitofrontal cortex (mOFC) (for a review see Milad and Rauch, 2012).

Behavioral experiments have provided only partial support for cognitive inhibitory deficits in OCD. Some studies found impaired performance on Stroop (Martinot et al., 1990), GoNogo (Bannon et al., 2002; Penades et al., 2007), Antisaccade (Tien et al., 1992) and negative priming tasks (Enright and Beech, 1993a, 1993b; Enright et al., 1995). Also, OCD patients manifested poorer performance on memory tasks that require updating of the executive system, such as the Letter Memory Task (e.g., Morris and Jones, 1990), the n-back Task (e.g., Kashyap et al., 2013; Nakao et al., 2009; Van der Wee et al., 2003), and prospective memory tasks (Racsmány et al., 2011; Harris et al., 2010). However, many studies failed to detect such impairments on executive tasks (e.g., Aronowitz et al., 1994; Maruff et al., 1999; Bannon et al., 2002; Aycicegi et al., 2003; Spengler et al., 2006; Moritz et al., 2010).

Importantly, another controversial body of literature assessing verbal, visual, and spatial memory in OCD (for reviews see Kuelz et al., 2004; Abramovitch et al., 2013) could be explained by a less effective organizational strategy use and impaired executive functioning (Christensen et al., 1992; Savage et al., 2000; Deckersbach et al., 2005).

Executive functions are crucial in everyday memory. Importantly, their role is not restricted to organizing during encoding,

* Corresponding author. Tel.: +36 1 4631 273 (Office); fax: +36 1 4631 072.
E-mail address: gdemeter@cogsci.bme.hu (G. Demeter).
[1] These authors contributed equally to the work.

http://dx.doi.org/10.1016/j.psychres.2014.04.022

planning retrieval, and monitoring memory output, but also in adaptive forgetting (Baddeley, 1996; Anderson, 2003; Dobbins et al., 2002). Indeed, the act of retrieval itself has been shown to cause forgetting of material related to the retrieved memory (Anderson et al., 1994; Anderson, 2003). This research line has shown that when one tries to retrieve a memory that is associated to a given cue, other memories associated to the same cue will become less accessible for later recall (Anderson et al., 1994; Camp et al., 2007; Racsmány et al., 2010).

This phenomenon has been widely studied with the *retrieval practice paradigm* (Anderson et al., 1994). In this paradigm participants study a list of category-exemplar pairs (e.g., vegetables – carrot, vegetables – tomato, sports – cycling, etc.), then practice retrieval of half of the exemplars from half of the categories (e.g., vegetables – ca___?). After a short delay, all exemplars from all categories are tested by a cued recall test. Typically, this final test shows that exemplars (e.g., tomato) associated to practiced exemplars (e.g., carrot) are less accessible than exemplars unrelated to any practiced exemplar (e.g., cycling). This effect has been termed retrieval-induced forgetting (RIF) and was replicated with a wide range of materials and research designs (Anderson and Bell, 2001; Levy and Anderson, 2002; Anderson, 2003; Bajo et al., 2006; Levy et al., 2007; Anderson and Levy, 2011; Storm, 2011).

Several mechanisms have been proposed to explain RIF. These include retrieval inhibition (Anderson et al., 1994), inhibitory executive control (Anderson, 2003), episodic inhibition (Racsmány and Conway, 2006), and, based on the Search of Associative Memory (SAM) theory (Raaijmakers and Shiffrin, 1981), noninhibitory interference processes (e.g., Raaijmakers and Jakab, 2012). Although these explanations contradict each other as to the involvement of inhibitory and/or executive control processes, neuroimaging studies of RIF clearly indicate that competitive retrieval activates cognitive control related areas in the human brain (Kuhl et al., 2007, 2007; Johansson et al., 2007; Kuhl et al., 2008; Wimber et al., 2009). These results show that when one tries to retrieve a target memory associated to a given cue, interference from other competing memories related to the same cue has to be resolved. According to these studies, interference resolution during memory retrieval involves prefrontal areas, as well as the anterior cingulate gyrus.

Problems in interference resolution through cognitive control (e.g., inhibition of intruding memories) have been suggested to be at the core of several psychiatric syndromes (Chamberlain et al., 2005). Therefore, RIF has been a popular tool to assess cognitive control in memory retrieval in schizophrenia (Racsmány et al., 2008) depression (Groome and Sterkaj, 2010), posttraumatic stress disorder (PTSD) (Amir et al., 2009), and OCD (Jelinek et al., 2012).

Jelinek et al. (2012) found intact RIF for neutral words and a tendency for reduced RIF for personally salient OCD relevant words in patients compared to healthy controls. They concluded that OCD is not characterized by a general inhibitory deficit, and that the reduced RIF for OCD-relevant memories is most likely due to cognitive biases. However, at the final test in their experiment, Jelinek et al. (2012) used a category cued recall where participants were given a category (e.g., vegetables) and were instructed to recall all words they had learnt together with that category in the experiment. When using this type of test, the observed RIF can be explained by response competition or output interference at test; practiced items come to mind first, and this blocks access to non-practiced memories (Anderson, 2003). Inhibition is unnecessary for the emergence of RIF in such a procedure, and non-inhibitory models (e.g., Raaijmakers and Shiffrin, 1981; Anderson, 1983) can account for a significant RIF. To eliminate the contribution of output interference to RIF, the final test should use category plus word stem cues which are specific to one given word in the experiment. Such a test could establish whether lower accessibility of a memory is due to interference resolution during an earlier retrieval act (Anderson, 2003). Therefore, in the current study we used this type of final test procedure.

In the clinical studies of RIF reviewed by the authors (Moulin et al., 2002; Nestor et al., 2005; Racsmány et al., 2008; Groome and Sterkaj, 2010; Storm and White, 2010) the final test was a category cued free recall task. As discussed above, in the RIF effects found in such studies output interference and inhibitory mechanisms are confounded. It follows that when output interference is ruled out from mechanisms producing the RIF effect, the effect itself becomes smaller, and less detectable (for a similar argument see Storm, 2011). Therefore, in our study, we focused on differences in recall latencies as a measure of RIF to ensure that any effect that decreases accessibility of memories due to competitive retrieval would be detected. Our choice for measuring RTs was motivated by earlier studies which suggested that RTs may be indeed sensitive to the effect of interference (Anderson, 2003; Keresztes and Racsmány, 2013) and may be more direct measures of the effect of interference resolution (Veling and van Knippenberg, 2004) than retrieval failure per se. Indeed, RTs proved to be a sensitive measure of the magnitude of RIF, even in cases when recall accuracy did not reveal any forgetting effect (Veling and van Knippenberg, 2004; Racsmány and Conway, 2006; Verde and Perfect, 2011).

Our goal was to investigate the role of executive functions in competitive retrieval in OCD. According to the executive deficit hypothesis both adaptive forgetting induced by retrieval (RIF) and suppression of unwanted thoughts are driven by similar executive processes (Levy and Anderson, 2008). In line with this hypothesis, Aslan and Bäuml (2010) found that the RIF effect was modulated by working memory capacity among healthy adults. Therefore we also assessed working memory using an *n*-back task which requires continuous updating of working memory contents. Apart from variables that are known to influence memory, such as symptom severity, depression, we also controlled for stress that has also been suggested to eliminate the RIF effect (Koessler et al., 2009). We hypothesized that OCD patients manifest reduced RIF compared to the matched healthy controls due to impaired executive functions that are supposed to resolve interference during competitive retrieval.

2. Methods

2.1. Participants

Twenty five patients diagnosed with OCD who satisfied the diagnostic criteria in the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV) (American Psychiatric Association, 2000) were examined at the Nyírő Gyula Hospital, Psychiatry II, Budapest, Hungary. A psychiatrist confirmed the diagnosis following the Structural Clinical Interview for DSM-IV Axis I Disorders (SCID-I) (First et al., 1997). The severity of OCD symptoms was assessed using the Yale Brown Obsessive-Compulsive Scale (Y-BOCS) (W.K. Goodman et al., 1989; W.L. Goodman et al., 1989). Severity of depression of the clinical sample was assessed using the Hamilton Rating Scale for Depression (HAM-D, 21-item) (Hamilton, 1960; Warren, 1994). Anxiety was assessed by the Spielberger State and Trait Anxiety Inventory (STAI). We used the State subscale of the STAI to estimate the stress induced by the experiment (Spielberger et al., 1970; Sipos, 1978; Spielberger, 1983). (See Table 1 for a summary of these assessments.)

We excluded participants who met the criteria for severe depression (Hamilton score > 24). Fourteen participants in our OCD sample were mildly depressed (Hamilton score between 7 and 17). We also excluded participants who met criteria for any other comorbid psychiatric diagnosis (Axis I or Axis II) and who had a lifetime history of drug or alcohol dependence or neurological disorder. Regarding medication, two patients had not been medicated for at least three months, eleven were taking selective serotonin reuptake inhibitors (seven paroxetin, two citalopram, one fevarin, and one stimuloton), ten were taking double action noradrenaline and serotonin agents (seven clomipramine and three venlafaxine) and two patients were taking serotonin reuptake inhibitors combined with double action noradrenaline and serotonin agents (paroxetin and clomipramin).

Table 1
Sample demographics and basic assessment results.

Characteristics	OCD (n=25) Mean	S.D.	Healthy control (n=25) Mean	S.D.	Independent t-test t	p
Age (years)	36.4	10.2	36.24	9.69	−0.05	ns
Education (years)	13.36	2.99	14.28	2.76	1.12	ns
Sex (M/F)	17/8		19/6			
DSF	6.04	0.93	6.16	1.31	0.37	ns
2-back hit %	84.36	18.04	93.92	10.71	2.27	<0.05
2-back correct rejection %	95.04	5.66	97.48	2.78	1.93	ns
3-back hit %	72.68	22.98	85.84	17.64	2.27	<0.05
3-back correct rejection %	93.32	5.12	95.88	2.69	2.2	<0.05
STAI-T	55.8	11.57	37.2	4.94	−7.38	<0.001
STAI-S	45.91	10.99	36.28	7.43	−3.57	<0.01
Y-BOCS total	26.6	7.3				
Y-BOCS ORS	13.04	3.65				
Y-BOCS CRS	13.64	4.80				
HAM-D	9.48	4.98				

Note. OCD, obsessive-compulsive disorder; M, male; F, female; DSF, Digit Span Forward; STAI-T, Spielberger Trait Anxiety Score; STAI-S, Spielberger State Anxiety Score; Y-BOCS Total, Yale Brown Obsessive Compulsive Scale Total score, Y-BOCS ORS, Yale Brown Obsessive Compulsive Scale, Obsessions-Severity Score; Y-BOCS CRS, Yale Brown Obsessive Compulsive Scale, Compulsions Severity Score; HAM-D, Hamilton Depressive Rating Scale; and ns, not significant.

Both concerning medication and symptoms, the OCD group was heterogeneous, and due to the sample size, no subgroup data could be analyzed.

The study was approved by the hospital's ethical review board. Patients received a detailed description and explanation of the study before they decided to sign the informed consent forms. All patients were assured that participation in the study would not interfere with their clinical treatment. The healthy control group (n=25) was matched in age and education (see Table 1).

2.2. *Experimental design and materials*

2.2.1. *The retrieval practice paradigm*

2.2.1.1. Design and material. We used 60 category-word pairs, six words belonging to each of ten categories. To induce the competitive retrieval supposed to be necessary to produce RIF, and to avoid moderation of the RIF effect (see Anderson, 2003), we followed strict selection criteria described in detail in Keresztes and Racsmány (2013). Briefly, we used neutral words of moderate frequency, based on the Frequency Dictionary of the Hungarian Webcorpus (Halácsy et al., 2004; Kornai et al., 2006). We used categories that were not associated to each other (either semantically or phonetically), and category members that were not associated to another member of another category. (see for the list of words used.)

Twelve category-word pairs (belonging to two categories) were used as fillers. The other 48 category-word pairs were assigned to one of the four item types. This was done in a fully randomized manner for each experimental session. First, each category was randomly assigned to be practiced or unpracticed (baseline) categories, so that half of the categories were selected for both category type. Words within each category were randomly assigned into two groups. Half of the words (Rp+) in each practiced category was to be practiced during the practice phase, the other half (Rp−) was not. For unpracticed (baseline) categories, half of the words (Nrp+) were assigned to be baseline items for Rp+ words, the other half (Nrp−) were assigned to be baseline items for Rp− words. We used Presentation 14.5 to randomly assign items to conditions, for presentation of stimuli and data recording. All participants performed the experiment on an IBM T40p ThinkPad.

2.2.1.2. Procedure. In the study phase, all 60 words were presented to participants together with their category label, at a rate of 5000 ms with a 500 ms intertrial interval, in the center of the computer 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 study phase started and ended with two filler items to avoid primacy and recency effects.

The practice phase consisted of three blocks, each containing 18 practice trials. Each block consisted of 12 trials with Rp+ items and six trials with fillers. The first and the last two items in each series were filler items. The order of the rest of the items was pseudo-randomized with the constraint that two consecutive trials never involved members of the same category. 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 say aloud the corresponding target. They had 6000 ms in the first block and 4000 ms in the second and third block to answer. Answers were recorded with a voice-key, and the correctness of the answers was manually checked by the experimenter after each session. In each trial, trigger onset time of the voice key was considered the response reaction time.

The three practice blocks followed each other in a repeated spaced retrieval schedule in order to increase the effect of practice (see Karpicke and Bauernschmidt, 2011). We introduced three, and six minutes of delay filled with simple arithmetic (adding or subtracting three-digit numbers randomly generated by the experimental software), before the second, and third practice block and a 5 min delay before the final test, respectively.

During the final test phase we used category plus word stem cues. This way, we ensured that each cue corresponded to only one test item in the experimental set. To further control for the effect of output interference, we tested Rp− items before testing Rp+ items. This procedure ensured that accessibility of Rp− items was not influenced by response competition arising during the final test (i.e., that a practiced word accessed during the test phase blocks access to a related memory that is tested after this practiced word). The final test phase consisted of two blocks. 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 (Camp et al., 2007). The order of items within blocks was randomized. Testing Rp− items first was necessary to avoid output interference created by Rp+ retrieval during test. Without such output control, output interference cannot be ruled out as an alternative explanation for any RIF effect observed. Controlling for output order of Rp− and Rp+ items necessitated the use of two baselines (Nrp− and Nrp+). This was necessary because recall of items tested at the end of a test session is usually lower than recall of items tested at the beginning of the session. This might then lead to a masking of the RIF effect because of a low overall baseline (see Anderson, 2003). Both blocks started and ended with two filler items. Filler items were necessary during the test phase to ensure consistency between experimental phases, and served as warm up trials. Trials in the test phase 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.

2.2.2. *Assessment of short term and working memory*

2.2.2.1. Digit span forward (DSF). We used the Hungarian version of the DSF task (Racsmány et al., 2005) as a measure of verbal short-term memory. In this task, a series of digits are presented orally by the examiner at a rate of one digit per second. The digits are to be repeated by the participant in the same order. Each trial consisted of four series of equal length (three digits in the first trial), and was considered successful if the participant reproduced at least two series correctly. In this case, the examiner advanced to the next trial which included series that were one digit longer. Digit span was determined by the length of the series in the last trial where the participant could recall at least two series correctly (see Table 1).

2.2.2.2. n-Back task. We designed visual 2-back and 3-back tasks with digits to measure the updating function of working memory. Each task consisted of five blocks of 30 trials. The first block served as practice in both the 2-back and the 3-back task. In these practice blocks participants were given feedback about correct hits, false alarms and misses. After each block, participants had a short self-paced break. In each trial, lasting 2000 ms, a digit, randomly sampled from one to nine, appeared in the center of the screen for 700 ms, followed by a blank screen for 1300 ms. Participants had to press the space bar on the keyboard if the digit on the screen was identical to the digit seen two (in the 2-back task), or three (in the 3-back task) trials before (see Table 1).

2.3. *Statistical analysis*

Statistical analyses were performed using mixed analysis of variance (ANOVAs), one-tailed t-tests and bivariate correlation (Pearson correlation coefficient). Partial Eta squared was used as a measure of the effect size for ANOVA and Cohen's d for the t-test analyses (Cohen, 1988; Field, 2005).

3. Results

3.1. *Psychiatric assessment*

Statistics and p values for the differences between scores of OCD patients and controls on the psychiatric scales are shown in Table 1. Patients were at the lower end of the mild depression range as revealed by the HAM-D. Their level of both trait and state anxiety was higher than that of controls, as indexed by the STAI-T and STAI-S respectively.

3.2. Short term and working memory

We performed independent *t*-tests to compare OCD patients' and controls' performance on the digit-span task, and hit and correct rejection rates in the 2-back and 3-back tasks. Statistics and corresponding *p* values are shown in Table 1. In brief, although short term memory span was almost identical in the two groups, working memory performance of OCD patients was lower than that of controls, as qualified by both hit rates and correct rejection rates.

3.3. Performance during retrieval practice

Recall performance during practice cycles can be seen in Fig. 1. To analyze memory improvement during retrieval practice, we conducted a mixed design ANOVA on recall RTs and recall percentages, with practice cycles (1–3) as a repeated measures factor, and group (OCD vs. Control) as a between-subject variable.

As can be seen in Fig. 1, there was a significant decrease in Recall RTs from cycle 1 through cycle 3, $F(2,90)=57.56$, $p<0.0001$, $\eta^2_{partial}=0.64$, indicating that participants' recall performance improved during retrieval practice, although participants' recall accuracy (77%,76%,78% and 77%, 80%, 79% from cycle 1 to cycle 3 in the OCD and the control group respectively) did not improve from cycle 1 through cycle 3, $F(2,94)=0.69$, $p=0.51$, $\eta^2_{partial}=0.01$. These main effects were not qualified by either a main effect of group ($F(1,45)=0.12$, $p=0.73$, $\eta^2_{partial}=0.00$, for recall RTs and $F(1,47)=0.11$, $p=0.74$, $\eta^2_{partial}=0.00$, for recall accuracy) or a group × practice cycle interaction ($F(2,90)=0.34$, $p=0.71$, $\eta^2_{partial}=0.01$, for recall RTs and $F(2,94)=1.12$, $p=0.33$, $\eta^2_{partial}=0.02$, for recall accuracy). In sum, memory improved in both groups during practice cycles, and this improvement was similar among participants with OCD and among controls.

3.4. The effect of retrieval practice on final test performance

Recall performance during the final test can be seen in Fig. 2. In order to see the differential effect of retrieval practice on recall of different item types, we conducted a mixed design ANOVA on recall RTs and recall accuracies (see Table 2) with item type (Rp+, Rp−, Nrp+, Nrp−) as a repeated measures variable, and group (OCD vs. controls) as a between subject variable. Item type had a significant main effect both on recall RTs, $F(3,126)=12.77$, $p<0.001$, $\eta^2_{partial}=0.23$, and recall accuracy, $F(3,144)=57.68$, $p<0.001$, $\eta^2_{partial}=0.55$. Item type did not interact significantly with group, neither for recall RTs, $F(3,126)=1.44$, $p=0.24$, $\eta^2_{partial}=0.03$, nor for recall accuracy, $F(3,144)=0.52$, $p=0.67$, $\eta^2_{partial}=0.01$.

Importantly, the healthy control group recalled more Nrp− items than Nrp+ $t(24)=2.02$, $p=0.027$, $d=0.82$. This was not surprising given that Nrp− items were tested first, and Nrp+ items second, i.e., we observed the effect of output interference (see Anderson, 2003). However this effect was absent among OCD patients, which might indicate that patients were not sensitive to output interference.

3.4.1. Practice effect

Fig. 2 (left panel) shows the positive effect of retrieval practice, the practice effect (Rp+ minus Nrp+) for recall RTs in the two groups separately. To detect a practice effect, we performed one sided paired-samples *t*-tests for the OCD and the control group separately, contrasting Rp+ recall with Nrp+ recall. Retrieval practice enhanced later recall of practiced memories based on recall RTs, ($t(23)=3.75$, $p<0.001$, $d=1.56$, for controls, and $t(19)=3.84$, $p<0.001$, $d=1.76$ for the OCD group) as well as recall accuracy ($t(24)=6.73$, $p<0.001$, $d=2.75$ among controls, and $t(24)=8.66$, $p<0.001$, $d=3.54$ among participants with OCD). In brief, practicing retrieval enhanced recall enhanced later memory for practiced items among both the OCD patients and controls.

3.4.2. Retrieval-induced forgetting

Fig. 2 (right panel) shows the negative effect of retrieval practice, the RIF effect (Nrp− minus Rp− recall) for recall RTs in the two groups separately. To detect a RIF effect, we performed paired-samples *t*-tests (one-sided) for the OCD and the control group separately, contrasting Rp− recall with Nrp− recall. Recall RTs revealed a significant RIF among controls, $t(24)=2.12$, $p=0.022$, $d=0.87$, but not among OCD patients, $t(19)=0.33$, $p=0.75$, $d=0.15$. The same pattern emerged from recall accuracy data, with no RIF observed among OCD patients, $t(24)=1.02$, $p=0.16$, $d=0.42$, but a tendency for a RIF effect among controls, $t(24)=1.67$, $p=0.053$, $d=0.68$. In brief, repeated retrieval of

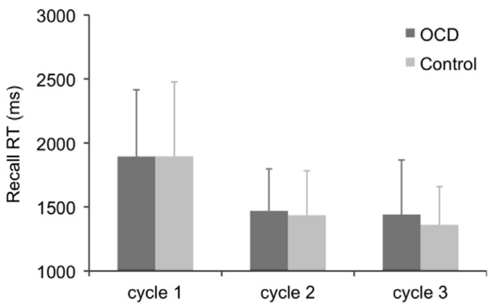


Fig. 1. Average recall reaction times in the three consecutive cycles of retrieval practice in the two groups. *Note.* Error bars represent standard deviations.

Table 2
Recall accuracies for the four item types in the two groups.

	Rp+	Nrp+	Rp−	Nrp−
OCD	0.70 (0.17)	0.38 (0.19)	0.39 (0.17)	0.41 (0.15)
Control	0.64 (0.20)	0.36 (0.16)	0.37 (0.17)	0.43 (0.18)

Note. Values show mean recall percentages (with standard deviations in brackets) during the final test for practiced items (Rp+), their baselines (Nrp+), and items related to practiced items (Rp−) and their baselines (Nrp−).

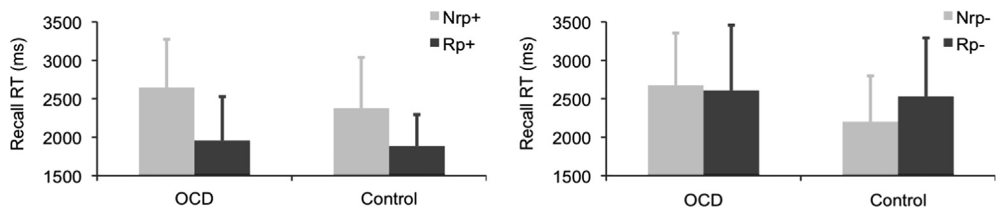


Fig. 2. Average recall reaction times on the final test in the two groups. *Note.* Final recall RTs of practiced items (Rp+), their baselines (Nrp+) (left panel), and items related to practiced items (Rp−) and their baselines (Nrp−) (right panel). The practice effect is evident in both groups (Rp+ items being recalled faster than Nrp+ items). Retrieval induced suppression is evident in the control group (Rp− recall is slower than Nrp− recall), but is absent in the OCD group. Error bars represent standard deviations.

memories caused suppression of related memories in the control group, but not among participants with OCD.

3.5. Potential factors modifying retrieval-induced forgetting

In the analyses below we calculated the Pearson correlation coefficients between the RIF score (calculated as the differences between Rp− recall RT and Nrp− recall RT), and potential factors related to RIF.

3.5.1. Working memory

Working memory did not correlate with RIF in our study: neither hit rates nor correct rejection rates in either the two-back or the three-back condition correlated with the RIF effect. This was true when correlations were calculated for the whole sample (all p's > 0.16, *ns*), as well as when the same correlations were calculated for the two groups separately (all p's > 0.29, *ns* for controls and all p's > 0.46, *ns* for participants with OCD, respectively).

3.5.2. Anxiety

Retrieval-induced forgetting did not correlate significantly with either the state (STAI-S) or the trait (STAI-T) measures of anxiety, $r=0.05$, $p=0.77$, and $r=0.23$, $p=0.13$, respectively. This pattern was the same when we analyzed the OCD group and the control group separately (respective statistics were: $r=0.33$, $p=0.88$; $r=-0.08$, $p=0.71$, for controls, and $r=-0.16$, $p=0.52$; $r=0.15$, $p=0.54$, for participants with OCD). In brief, the RIF effect was not correlated with anxiety.

3.5.3. Symptom severity

As measured by Y-BOCS total scores, symptom severity did not correlate significantly with RIF, $r=-0.19$, $p=0.38$. Also, there was no significant correlation between RIF, and either the obsessive subscale ($r=-0.21$, $p=0.33$), or the compulsive subscale of the Y-BOCS ($r=-0.12$, $p=0.55$).

4. Discussion

In this study, we aimed at assessing the ability to resolve interference during competitive retrieval in a sample of OCD patients, where the core cognitive dysfunction is characterized by executive deficits.

Our results demonstrate that retrieving memories does not induce forgetting of related memories among participants with OCD. Lack of forgetting in OCD occurred in spite of the fact that overall memory and the mnemonic effect of practicing memories was almost identical to that among healthy controls. Importantly, learning curves during the retrieval practice phase were similar in the two groups. The lack of RIF among OCD patients therefore is not related to overall recall performance, rather, we suggest that it is related to differences in resolving interference during competitive retrieval. In brief, despite similar recall performance in the two groups, recall of memories was not accompanied by adaptive suppression of related memories among OCD patients.

In line with previous work (Veling and van Knippenberg, 2004; Racsmány and Conway, 2006; Verde and Perfect, 2011) recall RTs proved to be more sensitive in detecting a RIF effect than simple recall accuracies. Among controls, we found a large and significant RIF effect (Cohen's $d=0.87$) as indexed by the RT data, and a medium size RIF effect (Cohen's $d=0.68$) that was present only at a trend level, when the measure of the effect was recall accuracy. In the OCD sample, according to the same measures, we found no effect for the RT data (Cohen's $d=0.15$) and a small and non significant effect for recall accuracy (Cohen's $d=0.42$).

Earlier, it was suggested by Koessler et al. (2009) that induced stress eliminates RIF among healthy participants by temporarily suspending the inhibitory mechanisms involved. Importantly, we found that both state and trait levels of anxiety were higher among patients than among controls, however these scores did not show any relationship with the amount of RIF. We have to mention that our study assessed stress induced by our experiment indirectly by the subjective evaluation of state anxiety (STAI-S), which could have caused the different results of our study and that of Koessler et al. (2009). In comparison with controls, updating of working memory was impaired among OCD patients, however contrary to the findings of Aslan and Bäuml (2010), WM performance did not correlate with RIF. We have to note that WM in our study was assessed by a different task (*n*-back) than the complex WM-task used by Aslan and Bäuml (2010). Although the *n*-back task and complex WM tasks have been generally thought to measure similar processes of WM, a recent meta-analysis by Redick and Lindsey (2013) implies that they are actually weakly correlated. Another difference between our study and the Aslan and Bäuml study was that our task produced much less variance, and their sample was four times as large as ours, while the effect detected in their study was weak. The correlation analyses showed no linear relationship between the RIF effect and stress, WM capacity, and symptom severity.

Our main findings are in contrast with the results of Jelinek et al. (2012) who found comparable RIF effects among OCD patients and healthy controls. However, in that study, Jelinek and colleagues also found a "tentative evidence for a weakened RIF effect for subjectively salient OCD-relevant material" (Jelinek et al., 2012, pp. 81). One potential confounding factor in their study could be the use of category cued free recall at final test. Such a test fails to control for output interference, whereby accessing memories that had been practiced during the practice phase blocks access to other related memories. In this case, the RIF effect would not be due to the effect of suppression but rather to some output interference process (Anderson, 2003). Here we showed that when item specific cues were used at the final test, retrieval practice did not impair the accessibility of related memories, i.e. no RIF was found.

From previous studies we know that OCD patients manifest problems in the use of organizational strategies during encoding of episodic memories (e.g., Savage et al., 1996; Deckersbach et al., 2005; Muller and Roberts, 2005) and in situations that involve executive functions (see; Kuelz et al., 2004; Abramovitch et al., 2013). These difficulties are particularly pronounced in tasks that are generally thought to tap inhibitory processes, such as the Stroop task (e.g., Martinot et al., 1990), the Go/NoGo task (Bannon et al., 2002; Penades et al., 2007; Watkins et al., 2005), and the antisaccade task (Maruff et al., 1999; Spengler et al., 2006; Tien et al., 1992). A strong hypothesis of Chamberlain et al. (2005) is that deficits of inhibition mechanisms are responsible for the main symptoms and neuropsychological profiles in OCD. In addition to inhibitory mechanisms, deficits in monitoring information also seem to be essential aspects of the cognitive profile of OCD, as suggested by results that indicate an overmonitoring in prospective memory tasks (Racsmány et al., 2011). Both of these processes are thought to be involved in conflict detection and conflict resolution arising during retrieval of competing memory representations (Anderson, 2003; Kuhl et al., 2007; Wimber et al., 2009; Hellerstedt and Johansson, 2013).

For instance, in an fMRI study, Kuhl et al. (2007) found evidence that repeated retrieval of target memories reduced the activity in a control network involving the ACC and dorso and ventrolateral PFC, structures important in detecting and resolving interference (Barch et al., 2000; Botvinick et al., 2004; Carter and Van Veen, 2007). The magnitude of reduction of PFC activity across repeated retrieval attempts of a target memory was associated with increased forgetting of interfering non-target memories at a final test, i.e., increased RIF. Accordingly, another fMRI study demonstrated that when memory competition is successfully resolved, the activity of the left medial and left lateral PFC, as well as activity in the left ACC is reduced (Wimber et al., 2009). Authors of both of the above studies suggest that the frontal structures are important not just in target memory selection but also in inhibition of related memories.

The ACC is of special interest in the context of interpreting our results. This area has been shown to be hyperactive in OCD compared to activity in controls, in tasks requiring cognitive conflict resolution and error detection (Bush et al., 2002; Van Veen and Carter, 2002; Fitzgerald et al., 2005; Maltby et al., 2005; Page et al., 2009; for a detailed review of brain areas affected in OCD see Milad and Rauch, 2012). Milad and Rauch (2012) suggested that the hyperactivity of the dorsal ACC might contribute to the persistence of error signals, producing the obsessive thoughts in OCD.

Given its role in conflict detection, one speculative interpretation of our results would be that the RIF effect is absent in OCD patients due to inappropriate conflict resolution processes during retrieval of competing memories driven by the constant hyperactivity of ACC and prefrontal structures. However, as no neuroimaging was involved in our study, the specific background mechanisms leading to the absence of RIF in OCD need to be addressed by novel experimental and neuroimaging studies.

Our study had a couple of limitations that have to be taken into account when interpreting the results. First, the material used in our study was not selected to be OCD relevant. Although our study was designed to address control processes in memory with emotionally neutral material, an earlier study by Jelinek et al. (2012) found tentative evidence that the use of personally salient material can modify the RIF effect. Second, the majority of the patients was under medication during the study and we included in the sample patients with mild depression. We think that studies with medication-naive patients are critical to obtain a better understanding of the relationship between clinical symptoms and cognitive deficits (e.g., Krishna et al., 2011). Third, a lot of different processes could be involved in the wide range of inhibitory tasks used in studies of inhibition, and cognitive inhibition itself has been defined in many different ways (for a review see Gorfein and MacLeod, 2007). Therefore, our conclusions may have benefited from additional results on another task measuring inhibition.

To conclude, it seems that in OCD interfering memories are not suppressed. Based on the inhibition deficit account of Chamberlain et al. (2005), one interpretation of our data is that the lack of the suppression effect is due to the inefficient suppression of irrelevant, interfering memories during competitive retrieval. However, it is also possible that the suppression effect is not produced by competitive retrieval in OCD because participants with OCD are not sensitive to interference as much as healthy participants. Further experiments are needed to clarify the role of conflict detection processes in the deficit of selective memory suppression in OCD.

Acknowledgments

Financial support for the research was provided by OTKA (Hungarian National Science Foundation) K84019. Gyula Demeter was supported by the Pro Progressio Foundation Scholarship, and the János Bolyai Research Scholarship of the Hungarian Academy of Sciences. We also thank Katalin Szentkuti and Ágnes Szőllősi for their assistance in data collection and Péter Pajkossy for his help in programming the *n*-back task.

Appendix A

See Table A1.

Table A1
English translation of the categories and target words used in the experiment (The original experimental material was Hungarian).

Category cue	Target member	Category cue	Target member
Material	Silk	Bird	Vulture
Material	Marble	Bird	Seagull
Material	Ceramics	Bird	Parrot
Material	Linen	Bird	Owl
Material	Concrete	Bird	Raven
Material	Silver	Bird	Sparrow
Transport	Trolley-bus	Clothes	Pullover
Transport	Truck	Clothes	Stockings
Transport	Airplane	Clothes	Socks
Transport	Helicopter	Clothes	Tyre
Transport	Train	Clothes	Hat
Transport	Metro	Clothes	Boot
Job	Soldier	Sports	Wrestling
Job	Interpret	Sports	Fencing
Job	Model	Sports	Kayaking
Job	Waiter	Sports	Tennis
Job	Policeman	Sports	Volleyball
Job	Driver	Sports	Riding
Fruit	Cranberry	Tools[a]	Pike[a]
Fruit	Ananas	Tools[a]	Saw[a]
Fruit	Strawberry	Tools[a]	Hammer[a]
Fruit	Prune	Tools[a]	Rake[a]
Fruit	Banana	Tools[a]	Shredder[a]
Fruit	Melon	Tools[a]	Welder[a]
Instrument	Saxophone	People[a]	Mexican[a]
Instrument	Clarinet	People[a]	Norvegian[a]
Instrument	Harp	People[a]	Turkish[a]
Instrument	Trumpet	People[a]	Brazilian[a]
Instrument	Viola	People[a]	Afgan[a]
Instrument	Organ	People[a]	Ukrainian[a]

[a] Indicates filler categories and filler items.

References

Abramovitch, A., Abramowitz, J.S., Mittelman, A., 2013. The neuropsychology of adult obsessive-compulsive disorder: a meta-analysis. Clinical Psychology Review 33, 1163–1171.
American Psychiatric Association, 2000. Diagnostic and Statistical Manual of Mental Disorders, Fourth edition American Psychiatric Association, Washington, DC (Text Revision (DSM-IV-TR), 3rd ed.).
Amir, N., Badour, C.L., Freese, B., 2009. The effect of retrieval on recall of information in individuals with posttraumatic stress disorder. Journal of Anxiety Disorders 23, 535–540.
Anderson, J.R., 1983. A spreading activation theory of memory. Journal of verbal learning and verbal behavior 22 (3), 261–295.
Anderson, M.C., 2003. Rethinking interference theory: executive control and the mechanisms of forgetting. Journal of Memory and Language 49, 415–445.
Anderson, M.C., Bell, T., 2001. Forgetting our facts: the role of inhibitory processes in the loss of propositional knowledge. Journal of Experimental Psychology, General 130, 544–570.
Anderson, M.C., Bjork, E.L., Bjork, R.A., 1994. Remembering can cause forgetting: retrieval dynamics in long-term memory. Journal of Experimental Psychology: Learning, Memory, and Cognition 20, 1063–1087.
Anderson, M.C., Levy, B.J., 2011. On the relationship between interference and inhibition in cognition. In: Benjamin, A.S. (Ed.), Successful Remembering and Successful Forgetting: A Festschrift in Honor of Robert A. Bjork. Psychology Press, London, UK, pp. 107–132.
Aron, A.R., Robbins, T.W., Poldrack, R.A., 2004. Inhibition and the right inferior frontal cortex. Trends in Cognitive Science 8, 170–177.
Aronowitz, B.R., Hollander, E., Decaria, C., Cohen, L., Saoud, J.B., Stein, D.J., 1994. Neuropsychology of obsessive-compulsive disorder. Preliminary findings. Neuropsychiatry Neuropsychology and Behavioral Neurology 7, 81–86.
Aslan, A., Bäuml, K.H.T., 2010. Retrieval-induced forgetting in young children. Psychonomic Bulletin & Review 17, 704–709.
Aycicegi, A., Dinn, W.M., Harris, C.L., Erkmen, H., 2003. Neuropsychological function in obsessive-compulsive disorder: effects of comorbid conditions on task performance. European Psychiatry 18, 241–248.

Baddeley, A., 1996. Exploring the central executive. Quarterly Journal of Experimental Psychology: Section A 49, 5–28.
Bajo, M.T., Gómez-Ariza, C.J., Fernández, A., Marful, A., 2006. Retrieval-induced forgetting in perceptually-driven memory tests. Journal of Experimental Psychology: Learning, Memory & Cognition 32, 1185–1194.
Bannon, S., Gonsalvez, C.J., Croft, R.J., Boyce, P.M., 2002. Response inhibition deficits in obsessive-compulsive disorder. Psychiatry Research 110, 165–174.
Barch, D.M., Braver, T.S., Sabb, F.W., Noll, D.C., 2000. Anterior cingulate and the monitoring of response conflict: Evidence from an fMRI study of overt verb generation. Journal of Cognitive Neuroscience 12 (2), 298–309.
Bokura, H., Yamaguchi, S., Kobayashi, S., 2001. Electrophysiological correlates for response inhibition in a Go/NoGo task. Clinical Neurophysiology 112, 2224–2232.
Botvinick, M.M., Cohen, J.D., Carter, C.S., 2004. Conflict monitoring and anterior cingulate cortex: an update. Trends in Cognitive Sciences 8, 539–546.
Bush, G., Vogt, B.A., Holmes, J., Dale, A.M., Greve, D., Jenike, M., Rosen, B.R., 2002. Dorsal anterior cingulate cortex: a role in reward-based decision making. Proceedings of the National Academy of Sciences of the United States of America 99, 523–528.
Camp, G., Pecher, D., Schmidt, H.G., 2007. No retrieval-induced forgetting using item-specific independent cues: evidence against a general inhibitory account. Journal of Experimental Psychology: Learning, Memory, and Cognition 33, 950–958.
Carter, C.S., Van Veen, V., 2007. Anterior cingulate cortex and conflict detection: an update of theory and data. Cognitive, Affective & Behavioral Neuroscience 7, 367–379.
Cavedini, P., Zorzi, C., Piccinni, M., Cavallini, M.C., Bellodi, L., 2010. Executive dysfunctions in obsessive-compulsive patients and unaffected relatives: searching for a new intermediate phenotype. Biological Psychiatry 67, 1178–1184.
Chamberlain, S.R., Blackwell, A.D., Fineberg, N.A., Robbins, T.W., Sahakian, B.J., 2005. The neuropsychology of obsessive-compulsive disorder: the importance of failures in cognitive andbehavioural inhibition as candidate endophenotypic markers. Neuroscience & Biobehavioral Reviews 29, 399–419.
Christensen, K.J., Kim, S.W., Dysken, M.W., Hoover, K.M., 1992. Neuropsychological performance in obsessive-compulsive disorder. Biological Psychiatry 31, 4–18.
Chudasama, Y., Robbins, T.W., 2003. Dissociable contributions of the orbitofrontal and infralimbic cortex to pavlovian autoshaping and discrimination reversal learning: further evidence for the functional heterogeneity of the rodent frontal cortex. Journal of Neuroscience 23, 8771–8780.
Cohen, J., 1988. Statistical Power Analysis for the Behavioral Sciences, 2nd ed. Lawrence Erlbaum Associates Inc, New Jersey.
De Bruin, J.P., Van Oyen, H.G., Van de Poll, N., 1983. Behavioural changes following lesions of the orbital prefrontal cortex in male rats. Behaviour and Brain Research 10, 209–232.
Deckersbach, T., Savage, C.R., Dougherty, D.D., Bohne, A., Loh, R., Nierenberg, A., Rauch, S.L., 2005. Spontaneous and directed application of verbal learning strategies in bipolar disorder and obsessive-compulsive disorder. Bipolar Disorders 7, 166–175.
Dobbins, I.G., Foley, H., Schacter, D.L., Wagner, A.D., 2002. Executive control during episodic retrieval: multiple prefrontal processes subserve source memory. Neuron 35, 989–996.
Enright, S.J., Beech, A.R., 1993a. Further evidence of reduced cognitive inhibition in obsessive-compulsive disorder. Personality and Individual Differences 14, 387–395.
Enright, S.J., Beech, A.R., 1993b. Reduced cognitive inhibition in obsessive-compulsive disorder. British Journal of Clinical Psychology 32, 67–74.
Enright, S.J., Beech, A.R., Claridge, G.S., 1995. A further investigation of cognitive inhibition in obsessive-compulsive disorder and other anxiety disorders. Personality and Individual Differences 19, 535–542.
Field, A., 2005. Discovering Statistics Using SPSS, 2nd ed. Sage Publications Ltd., London.
First, M.B., Spitzer, R.L., Gibbon, M., Williams, J.B.W., 1997. Structured Clinical Interview for DSM-IV Axis I Disorders – Clinician Version (SCID-CV). American Psychiatric Press, Washington, D.C..
Fitzgerald, K.D., Welsh, R.C., Gehring, W.J., Abelson, J.L., Himle, J.A., Liberzon, I., Taylor, S.F., 2005. Error-related hyperactivity of the anterior cingulate cortex in obsessive-compulsive disorder. Biological Psychiatry 57, 287–294.
Goodman, W.K., Price, L.H., Rasmussen, S.A., Mazure, C., Fleischman, R.L., Hill, C.L., Heninger, G.R., Charney, D.S., 1989a. The Yale-Brown Obsessive Compulsive Scale: I. Development, use, and reliability. Archives of General Psychiatry 46, 1006–1011.
Goodman, W.L., Price, L.H., Rasmussen, S.A., Mazure, C., 1989b. The Yale-Brown Obsessive Compulsive Scale II. Validity. Archives of General Psychiatry 46, 1012–1016.
Gorfein, D.S., MacLeod, C.M., 2007. Inhibition in Cognition. American Psychological Association, Washington DC.
Greisberg, S., McKay, D., 2003. Neuropsychology of obsessive-compulsive disorder: a review and treatment implications. Clinical Psychology Review 23, 95–117.
Groome, D., Sterkaj, F., 2010. Retrieval-induced forgetting and clinical depression. Cognition and Emotion 24, 63–70.
Halácsy P., Kornai A., Németh L., Rung A., Szakadát I., Trón V., 2004. Creating open language resources for Hungarian. In: Proceedings of the 4th International Conference on Language Resources and Evaluation, LREC2004. European Language Resources Association.
Hamilton, M., 1960. A rating scale for depression. Journal of Neurology, Neurosurgery & Psychiatry 23, 56–62.
Harris, L.M., Vaccaro, L., Jones, M.K., Boots, M.G., 2010. Evidence of impaired event-based prospective memory in clinical obsessive-compulsive checking. Behaviour Change 27, 84–92.
Hellerstedt, R., Johansson, M., 2013. Electrophysiological correlates of competitor activation predict retrieval-induced forgetting. Cerebral Cortex , http://dx.doi.org/10.1093/cercor/bht019.
Jelinek, L., Rietschel, L., Kellner, M., Muhtz, C., Moritz, S., 2012. The effect of practice on the recall of salient information in obsessive-compulsive disorder. Psychiatry Research 198, 89–93.
Johansson, M., Aslan, A., Bäuml, K.H., Gäbel, A., Mecklinger, A., 2007. When remembering causes forgetting: electrophysiological correlates of retrieval-induced forgetting. Cerebral Cortex 17, 1335–1341.
Karpicke, J.D., Bauernschmidt, A., 2011. Spaced retrieval: absolute spacing enhances learning regardless of relative spacing. Journal of Experimental Psychology: Learning, Memory, and Cognition 37 (5), 1250.
Kashyap, H., Kumar, J.K., Kandavel, T., Reddy, Y.C.J., 2013. Neuropsychological functioning in obsessive-compulsive disorder: are executive functions the key deficit? Comprehensive Psychiatry 54, 533–540.
Keresztes, A., Racsmány, M., 2013. Interference resolution in retrieval-induced forgetting: behavioral evidence for a nonmonotonic relationship between interference and forgetting. Memory and Cognition 41, 511–518.
Koessler, S., Engler, H., Riether, C., Kissler, J., 2009. No retrieval-induced forgetting under stress. Psychological Science 20, 1356–1363.
Krishna, R., Udupa, S., George, C.M., Kumar, K.J., Viswanath, B., Kandavel, T., et al., 2011. Neuropsychological performance in OCD: a study in medication-naive patients. Progress in Neuro-Psychopharmacology and Biological Psychiatry 35, 1969–1976.
Kornai, A., Halácsy, P., Nagy, V., Oravecz, C.s., Trón, V., Varga, D. 2006. Web-based frequency dictionaries for medium density languages. In: Proceedings of the 2nd International Workshop on Web as Corpus, ACL-06. pp. 1–9. Retrieved from: 〈http://people.mokk.bme.hu/%7Ekornai/Papers/webcorp.pdf〉.
Kuelz, A.K., Hohagen, F., Voderholzer, U., 2004. Neuropsychological performance in obsessive-compulsive disorder: a critical review. Biological Psychology 65, 185–236.
Kuhl, B.A., Kahn, I., Dudukovic, N.M., Wagner, A.D., 2008. Overcoming suppression in order to remember: contributions from anterior cingulate and ventrolateral prefrontal cortex. Cognitive, Affective and Behavioral Neuroscience 8, 211–221.
Kuhl, B.A., Dudukovic, N.M., Kahn, I., Wagner, A.D., 2007. Decreased demands on cognitive control reveal the neural processing benefits of forgetting. Nature Neuroscience 10, 908–917.
Levy, B.J., McVeigh, N.D., Marful, A., Anderson, M.C., 2007. Inhibiting your native language: the role of retrieval-induced forgetting during second language acquisition. Psychological Science 18, 29–34.
Levy, B.J., Anderson, M.C., 2002. Inhibitory processes and the control of memory retrieval. Trends in Cognitive Science 6, 299–305.
Levy, B.J., Anderson, M.C., 2008. Individual differences in the suppression of unwanted memories: the executive deficit hypothesis. Acta Psychologica 127, 623–635.
Maltby, D.F., Tolin, P., Worhunsky, A., 2005. Dysfunctional action monitoring hyperactivates frontal–striatal circuits in obsessive-compulsive disorder: an event-related fMRI study. NeuroImage 24, 495–503.
Martinot, J.L., Allilaire, J.F., Mazoyer, B.M., Hantouche, E., Huret, J.D., Legaut-Demare, F., Deslauriers, A.G., Hardy, P., Pappata, S., Baron, J.C., Syrota, A., 1990. Obsessive-compulsive disorder: a clinical, neuropsychological and positron emission tomography study. Acta Psychiatrica Scandinavica 82, 233–242.
Maruff, P., Purcell, R., Tyler, P., Pantelis, C., Currie, J., 1999. Abnormalities of internally generated saccades in obsessive-compulsive disorder. Psychological Medicine 29, 1377–1385.
Milad, M.R., Rauch, S.L., 2012. Obsessive compulsive disorder: beyond segregated cortico-striatal pathways. Trends in Cognitive Science 16, 43–51.
Moulin, C.J., Perfect, T.J., Conway, M.A., North, A.S., Jones, R.W., James, N., 2002. Retrieval-induced forgetting in Alzheimer's disease. Neuropsychologia 40, 862–867.
Moritz, S., Kloss, M., Jelinek, L., 2010. Negative priming (cognitive inhibition) in obsessive-compulsive disorder (OCD). Journal of Behavior Therapy and Experimental Psychiatry 41, 1–5.
Morris, N., Jones, D.M., 1990. Memory updating in working memory: the role of the central executive. British Journal of Psychology 81, 111–121.
Muller, J., Roberts, J.E., 2005. Memory and attention in obsessive-compulsive disorder: a review. Journal of Anxiety Disorders 19, 1–28.
Nakao, T., Nakagawa, A., Nakatani, E., Nabeyama, M., Sanematsu, H., Yoshiura, T., Togao, O., Tomita, M., Masuda, Y., Yoshioka, K., Kuroki, T., Kanba, S., 2009. Working memory dysfunction in obsessive–compulsive disorder: a neuropsychological and functional MRI study. Journal of Psychiatric Research 43, 784–791.
Nestor, P.G., Piech, R., Allen, C., Niznikiewicz, M., Shenton, M., McCarley, R.W., 2005. Retrieval-induced forgetting in schizophrenia. Schizophrenia Research 75, 199.
Olley, A., Malhi, G., Sachdev, P., 2007. Memory and executive functioning in obsessive-compulsive disorder: a selective review. Journal of Affective Disorders 104, 15–23.
Page, L.A., Rubia, K., Deeley, Q., Daly, E., Toal, F., Mataix-Cols, D., Giampietro, V., Schmitz, N., Murphy, D.G.M., 2009. A functional magnetic resonance imaging study of inhibitory control in obsessive-compulsive disorder. Psychiatry Research 174, 202–209.

Penades, R., Catalan, R., Rubia, K., Andres, S., Salamero, M., Gasto, C., 2007. Impaired response inhibition in obsessive compulsive disorder. European Psychiatry 22, 404–410.

Raaijmakers, J.G.W., Jakab, E., 2012. Retrieval induced forgetting without competition: testing the retrieval-specificity assumption of the inhibition theory. Memory & Cognition 40, 19–27.

Raaijmakers, J.G.W., Shiffrin, R.M., 1981. Search of associative memory. Psychological Review 88, 93–134.

Racsmány, M., Demeter, G.y., Csigó, K., Harsányi, A., Németh, A., 2011. An experimental study of prospective memory in obsessive-compulsive disorder. Journal of Clinical and Experimental Neuropsychology 33, 85–91.

Racsmány, M., Conway, M.A., Demeter, G.y., 2010. Consolidation of episodic memories during sleep: long-term effects of retrieval practice. Psychological Science 21, 80–85.

Racsmány, M., Conway, M.A., Garab, E.A., Cimmer, C., Janka, Z., Kurimay, T., Pléh, C., Szendi, I., 2008. Disrupted memory inhibition in schizophrenia. Schizophrenia Research 101, 218–224.

Racsmány, M., Conway, M.A., 2006. Episodic inhibition. Journal of Experimental Psychology: Learning, Memory and Cognition 32, 44–57.

Racsmány, M., Lukács, Á., Németh, D., Pléh, C.s., 2005. A verbális munkamemória magyar nyelvű vizsgálóeljárásai (Hungarian assessment methods of measuring verbal working memory). Magyar Pszichológiai Szemle 60, 479–505.

Rao, N.P., Reddy, Y.C., Kumar, K.J., Kandavel, T., Chandrashekar, C.R., 2008. Are neuropsychological deficits trait markers in OCD? Progress in Neuro-Psychopharmacology and Biological Psychiatry 32, 1574–1579.

Redick, T.S., Lindsey, D.R., 2013. Complex span and *n*-back measures of working memory: a meta-analysis. Psychonomic Bulletin & Review 20, 1102–1113, http://dx.doi.org/10.3758/s13423-013-0453-9 (Epub ahead of print).

Savage, C.R., Deckersbach, T., Wilhelm, S., Rauch, S.L., Baer, L., Reid, T., Jenike, M.A., 2000. Strategic processing and episodic memory impairment in obsessive compulsive disorder. Neuropsychology 14, 141–151.

Savage, C.R., Keuthen, N.J., Jenike, M.A., 1996. Recall and recognition memory in obsessive-compulsive disorder. Journal of Neuropsychiatry and Clinical Neurosciences 8, 99–103.

Sipos, K., 1978. A State-Trait Anxiety Inventory (STAI) magyar nyelvű változatával szerzett első hazai tapasztalatok. (First experiences with the Hungarian version of the State-Trait Anxiety Inventory). In: Dancs, I. (Ed.), 75 éves a Magyar Tudományos Akadémia Pszichológiai Intézete. MTA Pszichológiai Intézete, Budapest, pp. 142–152.

Spengler, D., Trillenberg, P., Sprenger, A., Nagel, M., Kordon, A., Junghanns, K., Heide, W., Arolt, V., Hohagen, F., Lencer, R., 2006. Evidence from increased anticipation of predictive saccades for a dysfunction of fronto-striatal circuits in obsessive-compulsive disorder. Psychiatry Research 143, 77–88.

Spielberger, C.D., 1983. State-Trait Anxiety Inventory. A Comprehensive Bibliography. Consulting Psychologists Press, Palo Alto, CA.

Spielberger, C.D., Gorsuch, R.L., Lushene, R.E., 1970. State-Trait Anxiety Inventory for Adults (Form X). Consulting Psychologists Press, Palo Alto, CA.

Storm, B.C., 2011. Retrieval-induced forgetting and the resolution of competition. In: Benjamin, A.S. (Ed.), Successful Remembering and Successful Forgetting: A Festschrift in Honor of Robert A. Bjork. Psychology Press, London, UK, pp. 89–105.

Storm, B.C., White, H.A., 2010. ADHD and retrieval-induced forgetting: evidence for a deficit in the inhibitory control of memory. Memory 18, 265–271.

Tien, A.Y., Pearlson, G.D., Machlin, S.R., Bylsma, F.W., Hoehn-Saric, R., 1992. Oculomotor performance in obsessive-compulsive disorder. American Journal of Psychiatry 149, 641–646.

Van der Wee, N.J., Ramsey, N.F., Jansma, J.M., Denys, D.A., Van Megen, H.J.G.M., Westenberg, H.M., Kahn, R.S., 2003. Spatial working memory deficits in obsessive compulsive disorder are associated with excessive engagement of the medial frontal cortex. NeuroImage 20, 2271–2280.

Van Veen, V., Carter, C.S., 2002. The anterior cingulate as a conflict monitor: fMRI and ERP studies. Physiology & Behavior 77, 477–482.

Veling, H., van Knippenberg, A., 2004. Remembering can cause inhibition: retrieval-induced inhibition as cue independent process. Journal of Experimental Psychology: Learning, Memory, and Cognition 30 (2), 315.

Verde, M.F., Perfect, T.J., 2011. Retrieval-induced forgetting in recognition is absent under time pressure. Psychonomic bulletin and review 18 (6), 1166–1171.

Warren, W.L., 1994. Revised Hamilton Rating Scale for Depression. Mental Measurements Yearbook, 13th ed. Western Psychological Services, Los Angeles.

Watkins, L.H,A., Sahakian, B.J., Robertson, M., Veale, D.M., Rogers, R.D., Pickard, K.M., Aitken, M., Robbins, T.W., 2005. Executive function in Tourette's syndrome and obsessive-compulsive disorder. Psychological Medicine 35, 571–582.

Wimber, M., Rutschmann, R.M., Greenlee, M.W., Bäuml, K.H., 2009. Retrieval from episodic memory: neural mechanisms of interference resolution. Journal of Cognitive Neuroscience 21, 538–549.

V.3. Interference resolution and retrieval inhibition (Study 6)

Paper: Interference resolution in retrieval-induced forgetting: Behavioral evidence for a nonmonotonic relationship between interference and forgetting. Attila Keresztes, Mihály Racsmány, *Memory & Cognition* 05/2013; 41(4):511-518. DOI:10.3758/s13421-012-0276-3

Mem Cogn (2013) 41:511–518
DOI 10.3758/s13421-012-0276-3

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

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. noninhibitory). 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?

A. Keresztes (✉) · M. Racsmány
Department of Cognitive Science, Budapest University of Technology and Economics, Budapest, Hungary
e-mail: akeresztes@cogsci.bme.hu

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 retrieved, target strength does not influence RIF (e.g., Anderson, 2003). Together, these assumptions imply that RIF is the product of executive processes that resolve interference during competitive retrieval.

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, 1985), repetition priming (Rajaram, Srinivas, & Travers, 2001), and the stop signal RT task (Logan & Cowan, 1984). Blaxton and Neely (1983) showed that RTs to generate the target exemplar were faster if the participant had first read other exemplars from the same category rather than exemplars from a different category. 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., Racsmány & 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–Média 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 Kornai 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 the final 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 $W = .81$; 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 $F(1, 62) = 2.28$, $p = .41$, 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, $F(1, 62) = 3.78$, $p = .057$,

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

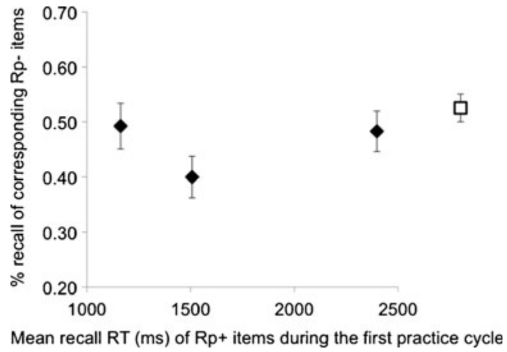


Fig. 1 Recall of Rp– items at test as a function of recall RTs for the associated Rp+ items during the first cycle of the retrieval practice phase. The empty rectangle on the right represents average Nrp– (baseline for the Rp– items) recall. The data are grouped into three tertiles according to the Rp+ 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.

[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.

[2] Binning our data into quartiles did reveal a U-shaped pattern of final recall (standard errors in parentheses): .50 (.04), .41 (.04), .44 (.04), and .48 (.04) for Rp–1.quar, Rp–2.quar, Rp–3.quar, and Rp–4.quar, respectively. Only the second- and third-quartile Rp– items differed significantly from Nrp– recall ($p = .013$ for Rp–2.quar, and $p = .042$ for Rp–3.quar). However, these comparisons do not survive Bonferroni correction. Also, this analysis did not have enough power to detect a quadratic trend in the data [$F(1, 62) = 1.98$, $p = .164$]. We thank one anonymous reviewer for drawing our attention to this problem.

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.

dc_1204_16

Appendix

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

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

The original experimental materials were in Hungarian. [a] Filler categories and filler items.

References

Anderson, M. C. (2003). Rethinking interference theory: Executive control and the mechanisms of forgetting. *Journal of Memory and Language, 49,* 415–445. doi:10.1016/j.jml.2003.08.006

Anderson, M. C., & Bjork, R. A. (1994). Mechanisms of inhibition in long-term memory: A new taxonomy. In D. Dagenbach & T. Carr (Eds.), *Inhibitory processes in attention, memory, and language* (pp. 265–325). San Diego: Academic.

Anderson, M. C., Bjork, R. A., & Bjork, E. L. (1994). Remembering can cause forgetting: Retrieval dynamics in long-term memory. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 20,* 1063–1087. doi:10.1037/0278-7393.20.5.1063

Anderson, M. C., & Levy, B. J. (2010). On the relationhip between interference and inhibition in cognition. In A. S. Benjamin (Ed.), *Successful remembering and successful forgetting: A festschrift in honor of Robert A. Bjork* (pp. 107–132). New York: Psychology Press.

Anderson, M. C., & McCulloch, K. C. (1999). Integration as a general boundary condition on retrieval-induced forgetting. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 25,* 608–629. doi:10.1037/0278-7393.25.3.608

Aslan, A., Bäuml, K.-H., & Pastötter, B. (2007). No inhibitory deficit in older adults' episodic memory. *Psychological Science, 18,* 72–78. doi:10.1111/j.1467-9280.2007.01851.x

Bäuml, K.-H., Pastötter, B., & Hanslmayr, S. (2010). Binding and inhibition in episodic memory—Cognitive, emotional, and neural processes. *Neuroscience and Biobehavioral Reviews, 34,* 1047–1054. doi:10.1016/j.neubiorev.2009.04.005

Blaxton, T. A., & Neely, J. H. (1983). Inhibition from semantically related primes: Evidence of a category-specific inhibition. *Memory & Cognition, 11,* 500–510. doi:10.3758/BF03196987

BME Média és Szociológia Tanszék–Média Oktatási és Kutató Központ (BME-MOKK). (2003). *Frequency dictionary of the Hungarian Webcorpus* [Tab-separated text files]. Retrieved from http://mokk.bme.hu/resources/webcorpus

Camp, G., Pecher, D., & Schmidt, H. G. (2007). No retrieval-induced forgetting using item-specific independent cues: Evidence against a general inhibitory account. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 33,* 950–958. doi:10.1037/0278-7393.33.5.950

Halácsy, P., Kornai, A., Németh, L., Rung, A., Szakadát, I., & Trón, V. (2004). Creating open language resources for Hungarian. In M. T. Lino, M. F. Xavier, F. Ferreira, R. Costa, & R. Silva (Eds.), *Proceedings of the 4th International Conference on Language Resources and Evaluation, LREC2004* (pp. 1201–1204). Paris: European Language Resources Association.

Kornai, A., Halácsy, P., Nagy, V., Oravecz, C., Trón, V., & Varga, D (2006). Web-based frequency dictionaries for medium density languages. In *Proceedings of the 2nd International Workshop on Web as Corpus*, *ACL-06* (pp. 1–9). Retrieved from http://people.mokk.bme.hu/%7Ekornai/Papers/webcorp.pdf

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

Levy, B. J., McVeigh, N. D., Marful, A., & Anderson, M. C. (2007). Inhibiting your native language: The role of retrieval-induced forgetting during second-language acquisition. *Psychological Science, 18,* 29–34. doi:10.1111/j.1467-9280.2007.01844.x

Logan, G. D., & Cowan, W. B. (1984). On the ability to inhibit thought and action: A theory of an act of control. *Psychological Review, 91,* 295–327. doi:10.1037/0033-295X.91.3.295

Norman, K. A., Newman, E. L., & Detre, G. (2007). A neural network model of retrieval-induced forgetting. *Psychological Review, 114,* 887–953.

dc_1204_16

Perfect, T. J., Stark, L.-J., Tree, J. J., Moulin, C. J. A., Ahmed, L., & Hutter, R. (2004). Transfer appropriate forgetting: The cue-dependent nature of retrieval-induced forgetting. *Journal of Memory and Language, 51,* 399–417. doi:10.1016/j.jml.2004.06.003

Racsmány, M., & Conway, M. A. (2006). Episodic inhibition. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 32,* 44–57. doi:10.1037/0278-7393.32.1.44

Rajaram, S., Srinivas, K., & Travers, S. (2001). The effects of attention on perceptual implicit memory. *Memory & Cognition, 29,* 920–930. doi:10.3758/BF03195754

Saunders, J., & MacLeod, M. D. (2006). Can inhibition resolve retrieval competition through the control of spreading activation? *Memory & Cognition, 34,* 307–322. doi:10.3758/BF03193409

Storm, B. C., Bjork, E. L., Bjork, R. A., & Nestojko, J. F. (2006). Is retrieval success a necessary condition for retrieval-induced forgetting? *Psychonomic Bulletin & Review, 13,* 1023–1027. doi:10.3758/BF03194002

Tipper, S. P. (1985). The negative priming effect: Inhibitory priming by ignored objects. *Quarterly Journal of Experimental Psychology, 37A,* 571–590.

Veling, H., & van Knippenberg, A. (2004). Remembering can cause inhibition: Retrieval-induced inhibition as cue independent process. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 30,* 315–318. doi:10.1037/0278-7393.30.2.315

Verde, M. F., & Perfect, T. J. (2011). Retrieval-induced forgetting in recognition is absent under time pressure. *Psychonomic Bulletin & Review, 18,* 1166–1171. doi:10.3758/s13423-011-0143-4

Williams, C. C., & Zacks, R. T. (2001). Is retrieval-induced forgetting an inhibitory process? *American Journal of Psychology, 114*, 329–354.

Wimber, M., Rutschmann, R. M., Greenlee, M. W., & Bäuml, K.-H. (2009). Retrieval from episodic memory: Neural mechanisms of interference resolution. *Journal of Cognitive Neuroscience, 21,* 538–549. doi:10.1162/jocn.2009.21043

VI. Episodic retrieval and long-term memory representations

VI.1. Why retrieval is a protective process of adverse effect of retrieval? (Study 7)

Paper: Initial retrieval shields against retrieval-induced forgetting. Mihály Racsmány, Attila Keresztes, *Frontiers in Psychology* 05/2015; 6:657. DOI:10.3389/fpsyg.2015.00657

dc_1204_16

frontiers in Psychology

ORIGINAL RESEARCH
published: 21 May 2015
doi: 10.3389/fpsyg.2015.00657

Initial retrieval shields against retrieval-induced forgetting

*Mihály Racsmány[1,2]** and *Attila Keresztes[3]*

[1] Department of Cognitive Science, Budapest University of Technology and Economics, Budapest, Hungary, [2] Research Group on Frontostriatal Disorders, Hungarian Academy of Sciences, Budapest, Hungary, [3] Center for Lifespan Psychology, Max Planck Institute for Human Development, Berlin, Germany

Testing, as a form of retrieval, can enhance learning but it can also induce forgetting of related memories, a phenomenon known as retrieval-induced forgetting (RIF). In four experiments we explored whether selective retrieval and selective restudy of target memories induce forgetting of related memories with or without initial retrieval of the entire learning set. In Experiment 1, subjects studied category-exemplar associations, some of which were then either restudied or retrieved. RIF occurred on a delayed final test only when memories were retrieved and not when they were restudied. In Experiment 2, following the study phase of category-exemplar associations, subjects attempted to recall all category-exemplar associations, then they selectively retrieved or restudied some of the exemplars. We found that, despite the huge impact on practiced items, selective retrieval/restudy caused no decrease in final recall of related items. In Experiment 3, we replicated the main result of Experiment 2 by manipulating initial retrieval as a within-subject variable. In Experiment 4 we replicated the main results of the previous experiments with non-practiced (Nrp) baseline items. These findings suggest that initial retrieval of the learning set shields against the forgetting effect of later selective retrieval. Together, our results support the context shift theory of RIF.

Keywords: retrieval-induced forgetting, retrieval-enhanced learning, inhibition, context reinstatement, episodic memory, context effects

OPEN ACCESS

Edited by:
Petko Kusev,
Kingston University London, UK

Reviewed by:
Gregoire Borst,
Université Paris Descartes, France
Eddy J. Davelaar,
Birkbeck, University of London, UK

****Correspondence:***
Mihály Racsmány,
Department of Cognitive Science,
Budapest University of Technology and Economics, Egry József utca 1,
Budapest 1111, Hungary
racsmany@cogsci.bme.hu

Specialty section:
This article was submitted to Cognition,
a section of the journal Frontiers in Psychology

Received: *06 February 2015*
Accepted: *05 May 2015*
Published: *21 May 2015*

Citation:
Racsmány M and Keresztes A (2015) Initial retrieval shields against retrieval-induced forgetting.
Front. Psychol. 6:657.
doi: 10.3389/fpsyg.2015.00657

Introduction

The act of retrieval facilitates later access to retrieved memories. Typically, in comparison with repeated study (restudy), repeated retrieval of memories improves long-term retention, whereas it produces equal or often lower recall performance following a short-term delay (Carrier and Pashler, 1992; Wheeler et al., 2003; Roediger and Karpicke, 2006a,b; Karpicke and Roediger, 2008; Toppino and Cohen, 2009; Keresztes et al., 2013). However, the long-term benefits of retrieval often come with a cost: retrieval-induced forgetting (RIF; Anderson et al., 1994); when retrieval is selective, non-retrieved, but related memories become less accessible.

It has been shown that both selective retrieval and selective restudy of a learning set increase the recall probability of retrieved/restudied memories; however, only selective retrieval induces forgetting of related information from the same set (Ciranni and Shimamura, 1999; Anderson et al., 2000; Bäuml, 2002; Bäuml and Aslan, 2004; Staudigl et al., 2010; but see Verde, 2009). RIF is a robust experimental phenomenon at short delays, and recent findings suggest that it is present also after longer delays (Racsmány et al., 2010; Abel and Bäuml, 2012; Storm et al., 2012; but see MacLeod and Macrae, 2001).

Importantly, this pattern of findings is a potential problem for any educational program using frequent selective retrieval—i.e., testing—of large sets of information as a learning method. In brief, these findings highlight that retrieval has a robust long-term advantage over repeated study of information at the expense of forgetting related, but not retrieved, information. Identifying any factor that could protect these memories from being forgotten, therefore, is key to creating effective learning programs.

In the following sections, we outline the retrieval practice paradigm (Anderson et al., 1994), that is most commonly used to investigate RIF, and then briefly overview three families of theories on associative retrieval processes that can explain RIF. Finally, based on the assumptions of one family of theories, we suggest one critical factor that could shield against the adverse effects of RIF: an initial—non-selective—retrieval of the entire learning set.

In the retrieval practice paradigm (Anderson et al., 1994), participants study category–member pairs (e.g., animal–tiger, furniture–couch, animal–chicken, etc.); then, in a selective retrieval practice phase, they repeatedly retrieve half of the members from half of the categories (e.g., animal–t. . .?). Typically, final recall administered after a delay reveals that repeated selective retrieval leads to forgetting of related material (e.g., "animal–c. . .?") compared to unpracticed baseline categories (e.g., furniture–c. . .?)—this effect is referred to as RIF.

The most influential family of theories—the inhibitory control based accounts—posit that when participants practice retrieval of half of the members from a given category, the other half would compete for retrieval (Anderson et al., 1994, 2000; Anderson and McCulloch, 1999; Anderson and Bell, 2001; Bäuml and Hartinger, 2002; Storm et al., 2006; Storm and Nestojko, 2010). This competition is then resolved by executive control guided active inhibition, which renders the memories of competitors less accessible for later recall (Anderson, 2005; Anderson and Levy, 2007).

Interference based accounts—the second family of theories—explain RIF without inhibition (Camp et al., 2007, 2009; Jakab and Raaijmakers, 2009). These models assume that strengthening some category-member associations is enough to lead to interference at any later attempt to retrieve competitors. Here, it is this interference at final recall that leads to RIF. The most influential of these models, the search of associative memory (SAM) model (Raaijmakers and Shiffrin, 1981) assumes that retrieval occurs in two steps. First—in the sampling phase—cues are assembled into a short-term store for activated memory sets, and items are sampled into these sets based on the relative strength of their associations to the given cue. In a second step—the recovery phase—sampled items are retrieved based on the absolute strength of their associations to the given cue. It is only a successful recovery that leads to conscious retrieval of a memory item. Using these terms, interference based accounts assume that RIF is the consequence of a sampling failure, i.e., a bias in relative associative strengths, whereas inhibitory models assume that RIF occurs due to recovery failure, i.e., due to a decreased item strength.

The third family of theories pinpoint episodic or context-based retrieval as the source of RIF, suggesting that any kind of retrieval creates and reshapes highly contextualized episodic memory representations (Racsmány and Conway, 2006; Conway, 2009; Racsmány et al., 2012; Jonker et al., 2013; Karpicke et al., 2014; see Sahakyan and Hendricks, 2012, for a similar account of directed forgetting). Episodic memory sets contain context, cue, and item features (Racsmány and Conway, 2006; Conway, 2009). The most influential of these theories emphasizes the role of context shift between studying a memory set and retrieval of parts of this set (Jonker et al., 2013; for a similar account of directed forgetting, see Sahakyan and Kelley, 2002). According to the context shift theory, the mental context of the study phase is changed in the following retrieval phase due to processes activated by retrieval of parts of the set. This context then remains the same throughout the rest of these experiments—RIF is found because the mental context of the final recall is biased to mimic retrieval pattern of the previous selective retrieval and not that of the initial study phase.

Importantly for our current research question, the context shift theory leads to the prediction that an initial retrieval attempt of the entire learning set can eliminate the adverse effect of later selective retrieval. This is because an initial retrieval can already establish the episodic context for the rest of the experiment (see Jonker et al., 2013; Karpicke et al., 2014). This way, final recall will bias the retrieval process to mimic the pattern of the initial retrieval and grant access to items not selectively practiced as well.

Retrieval is so central to the wide range of the above discussed theories that retrieval-specificity—the concept that retrieval is necessary to produce RIF—has become a descriptive feature of RIF (Anderson and Spellman, 1995; Anderson, 2003; Storm, 2011). A crucial, and well replicable finding, is that selectively restudying category-member pairs is not enough to produce RIF, category members should be selectively retrieved to induce the effect (Blaxton and Neely, 1983; Bäuml, 1996, 1997, 2002; Ciranni and Shimamura, 1999; Anderson et al., 2000; Anderson and Bell, 2001; Shivde and Anderson, 2001; Levy and Anderson, 2008; Jonker et al., 2013;, but see Raaijmakers and Jakab, 2012). This finding is in line with the inhibitory control based accounts, because these assume that inhibition is only necessary when the retrieval process induces competition between target memories and competitors (Anderson, 2003). It is also in line with theories emphasizing the role of context-based, episodic retrieval in producing RIF, because these theories assume that it is the retrieval process that produces the shift from the study context to the context of retrieval, and creates biased contextualized episodic memory sets (Racsmány and Conway, 2006; Jonker et al., 2013). In contrast, according to the interference accounts, both selective retrieval and restudy should lead to RIF—a prediction incompatible with what is generally found.

However, Verde (2013) suggested that the latest version of the SAM–REM model (Malmberg and Shiffrin, 2005) could explain the same pattern with the additional assumption that retrieval strengthens the context-item associations, whereas restudy strengthens cue-item associations. Because only the former affects the sampling process (by modifying relative strength of associations)—the source of RIF in this model—only retrieval leads to RIF. In support of this suggestion, recent studies (Jonker and MacLeod, 2012; Raaijmakers and Jakab, 2012; Verde, 2013; Experiment 2) showed that selectively strengthening category-member associations and emphasizing context encoding without retrieval might also lead to RIF.

Given the pivotal role of retrieval in shaping episodic memory sets, it is surprising that studies using the retrieval practice paradigm have not investigated the effect of an initial retrieval phase where participants attempt to recall the entire learning set once before selective retrieval. To our knowledge, in the vast amount of experiments investigating the RIF effect, the first retrieval act that occurred in the experiments was selective retrieval, when participants aimed to access only a part of the studied elements[1].

Besides investigating its protective role against RIF, performance in an initial retrieval phase could also provide experimenters with a direct baseline for measuring the extent of forgetting. In the retrieval practice paradigm, baseline is generally measured as the final recall performance of memory items belonging to categories not appearing during the practice phase. Because these categories and corresponding target memories appear in the initial study phase, but neither the category label, nor any member of these categories appear during the selective practice manipulation, these items seem to be a good choice for measuring baseline performance. However, this poses at least three problems in the interpretation of final recall performance. The first is baseline deflation (Anderson, 2003), coined for the phenomenon that during the course of a test session items tested later will suffer interference from items tested earlier, and the probability of successful recall during a test session decreases with the number of previously tested items. The second is cue priming: Cues for selectively retrieved categories appear during the practice phase, and this causes a bias in cue processing at final recall so that practiced items are more probably retrieved and may block access to unpracticed items. Similar cue biases do not occur for cues of categories not selectively retrieved. Third, context biases may add up to cue priming: The context of the retrieval practice phase itself creates uneven recall probabilities for retrieved and non-retrieved memories from categories retrieved during the practice phase. Again, similar context biases do not occur for cues of categories not retrieved during the practice phase of the retrieval practice paradigm (Racsmány and Conway, 2006; Jonker et al., 2013). We suggest that measuring baseline directly with an initial retrieval of the entire learning set can circumvent these issues, and facilitate interpretation of final recall data in the retrieval practice paradigm.

In this paper, we investigated the possible adverse effect of retrieval practice on a part of the studied elements when an initial retrieval accessed the entire memory set studied earlier in the experiment. Additionally, using performance of this initial retrieval, the effect of further selective retrieval on both retrieved and non-retrieved memories could be assessed to a baseline recall level of the same memories. Therefore the following experiments had two aims: first, to measure the interaction between initial testing of the entire learning set and the adverse effect of later selective retrieval practice on related unpracticed items, and second, to introduce a novel baseline measure, the initial retrieval performance, for future RIF experiments.

Based on accounts emphasizing the episodic/contextual nature of retrieval practice (Racsmány and Conway, 2006; Jonker et al., 2013; Karpicke et al., 2014), we predicted that an initial attempt to—non-selectively—retrieve the entire learning set would shield against the adverse effects of later selective retrieval, together with maintaining the positive effects of retrieval practice for retrieved memories. In contrast, interference accounts would predict no effect of an initial retrieval. Because in these accounts, RIF depends on relative cue-item or context-item association strengths, an equally distributed increase in these association strengths would not shift the effect of later selective strengthening of these associations. It is harder to derive predictions based on inhibitory control based accounts. Although strengthening all items via an initial retrieval can lead to larger competition during later selective retrieval—hence to larger RIF, the effect could also be the opposite; based on a trade-off between the need for inhibition during competitive retrieval, and the success of inhibition (Norman et al., 2007; Anderson and Levy, 2011; see experimental evidence, Keresztes and Racsmány, 2013) it can well be that strengthening items that later become competitors can render inhibitory processes ineffective—hence to no RIF. Similarly, results showing that retrieval of cue-item associations can decrease later interference generated by these associations (Szpunar et al., 2008; Halamish and Bjork, 2011) would suggest that an initial retrieval of competitors can decrease competition during later selective retrieval of related targets. Again, decreased competition would lead to decreased inhibition—hence to an attenuated RIF.

The first experiment reported here aimed to replicate previous findings of retrieval specificity of RIF. Then, using the same material and procedures, we investigated the effect of an initial retrieval of all items in the experiment on further effects of selective retrieval.

Experiment 1

Method

Participants

All four experiments were approved by the Ethical Committee of the Budapest University of Technology and Economics, and all participants gave their written informed consent.

Sixty[2] participants were recruited for Experiment 1 at the Budapest University of Technology and Economics. Outliers were defined as data points more than three standard deviations away from the group mean. We screened data for outliers for overall recall performance and recall in all four item types (see design section). Data for one participant was identified as outlier; and

[1]Although a few studies with clever designs (e.g.,Storm et al., 2006, 2008) did investigate the effect of iterating retrieval practice cycles and relearning cycles after the study phase in retrieval practice paradigms, these studies focused on the effect of adaptive forgetting on later relearning, and were not designed to look at the effect of an initial, non-selective retrieval on the negative effects of later selective retrieval.

[2]Since some of our hypotheses concerned finding null-effects (i.e., no RIF), we performed preliminary power calculations with G*Power (Faul et al., 2009) using earlier published (Racsmány et al., 2010) and unpublished data from a retrieval practice paradigm with similar material and population from our lab. To achieve a power of $(1 - \beta) = 0.8$ to detect RIF, with a two-sided paired-samples t-test, we needed to include 25 participants per condition. Therefore we settled on 30-participant samples for all our experiments—a number common in retrieval-practice paradigms.

excluded from further analyses. Therefore, the results section shows the data for 59 participants (26 men and 28 women), aged between 19 and 26 years ($M = 20.36$, SD $= 1.47$).

Design and Materials

We varied practice type (retest or restudy) between subjects, and item type within subjects. We used 10 categories and six words from each category, a total of 60 category-word pairs. To induce competitive retrieval supposed to be necessary to produce RIF, and to avoid moderation of the RIF effect (see Anderson, 2003), we followed strict selection criteria described in detail in Keresztes and Racsmány (2013). Briefly, we used neutral words of moderate frequency, based on the Frequency Dictionary of the Hungarian Webcorpus (Halácsy et al., 2004; Kornai et al., 2006). We used categories that were not associated to each other (either semantically or phonetically), and category members that were not associated to another member of another category.

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. Counterbalancing across all conditions was achieved by a full randomization procedure run by Presentation® software (Version 14.7, www.neurobs.com) for each participant separately. Briefly, four categories were selected randomly 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 served as baseline for Rp+ words, the other half (Nrp−) served as baseline for Rp− words.

Procedure

The experiment consisted of four phases: a study phase, a practice phase, a delay, and a final test phase. Restudy and retest conditions differed only in their practice phase.

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 centre 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 and reexposure blocks each consisted of 12 trials with Rp+ items and six trials with filler items. 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 retest 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, and 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 using a keyboard. 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. Each pair was shown once in the center of the screen with the category label on the left and the category member on the right, and participants were instructed to use these trials to restudy the category label—word pairs.

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. Note that whereas 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.

The three practice cycles (for both retest and restudy) followed each other in a repeated spaced retrieval schedule in order to enhance the effect of testing (see Karpicke and Bauernschmidt, 2011). We introduced 1, 3, and 6 min of delay filled with a two-back task, before the first, second, and third practice cycle, respectively.

After the practice phase participants performed a 5-min long two-back task, and then were introduced to the final test phase. In the 2-back task, participants saw a series of numbers, one at a time, in the middle of a computer screen, and for each trial they had to respond by pressing a button on the keyboard when the number in the current trial was the same as the one presented two trials before. In each trial, stimuli was sampled pseudorandomly from among five integers (1–5) so that the program selected the current number to be a target, i.e., the same as the number appearing to trials before, with a 25% probability. Trials were 2000 ms long (700 ms stimulus duration, 1300 ISI). Participants received a 2000 ms feedback for hits, misses, and false alarms.

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 et al., 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.

Randomization of trials, presentation of stimuli, response logging, and data preprocessing were performed by Presentation® software (Version 14.7, www.neurobs.com).

Results and Discussion

Throughout the manuscript, we report effect sizes using r for t-tests and η_p^2 for F-tests. Recall performance at the final test for the four item types are shown in **Figure 1**.

The Effect of Practice on Final Recall

We conducted a mixed design ANOVA on recall data with item type (Rp+, Rp−, Nrp+, Nrp−) as a repeated measures variable,

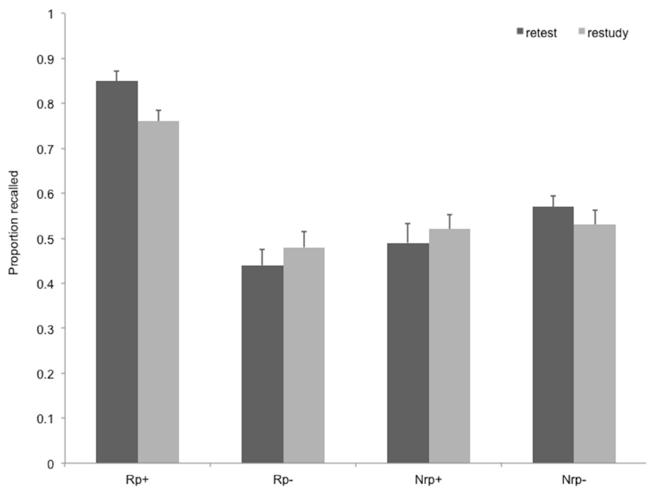


FIGURE 1 | Recall performance on the final test in Experiment 1, for the four item types in the two practice conditions. Rp+, Practiced words from practiced categories; Rp−, unpracticed words from practiced categories; Nrp+, words from unpracticed categories used as baseline for Rp+ words; Nrp−, words from unpracticed categories used as baseline for Rp− words. Error bars indicate standard error of the mean.

and practice type (retest vs. restudy) as a between subject variable. Item type had a significant main effect on final recall, $F(3,171) = 66.40$, $p < 0.001$, $\eta_p^2 = 0.54$, and there was a tendency toward an interaction of item type with practice type, $F(3,171) = 2.53$, $p = 0.058$, $\eta_p^2 = 0.04$. Retesting led to a similar overall recall as restudying, $F(1,57) = 0.26$, ns.

To detect RIF, we performed paired-samples t-tests for participants in the retest and the restudy condition separately, contrasting Rp− recall with Nrp− recall. The RIF effect was only significant in the retest condition, $t(28) = -3.13$, $p = 0.004$, $r = 0.37$, but no RIF was found in the restudy condition, $t(29) = 1.43$, $p = 0.16$. In brief, testing induced forgetting only when participants were retested during the practice phase, and not when they restudied the same material.

Retrieval practice led to enhancement of memory for practiced items (as compared to Nrp+ baseline items) in both conditions, $t(28) = 5.91$, $p < 0.001$, $r = 0.60$, in the restudy and $t(29) = 9.94$, $p < 0.001$, $r = 0.70$ in the retest condition.

In brief, the results of Experiment 1 replicated earlier findings: Selectively retrieving memories from a category induce forgetting of related, but non-retrieved memories from the same category, whereas selective restudy of memories does not lead to this type of forgetting. Importantly, *post hoc* power calculations on data from Experiment 1 showed that the paradigm was indeed well-powered to detect any differences between Rp− items and their Nrp− baselines $(1 - \beta) = 0.88$. It was crucial for us to have a well-powered paradigm in order to exclude Type II errors in the following experiments.

In Experiment 2 we manipulated the type of practice within subjects, and introduced an initial retrieval test immediately after the study phase to test whether an initial retrieval test able to eliminate the RIF effect. This procedure also introduced a novel baseline measure for each item type: the initial recall performance. Note that this experiment did not involve unpracticed items from unpracticed categories (NRP items) as a baseline.

Experiment 2

Method

Participants

Thirty participants were recruited at the Budapest University of Technology and Economics (15 men and 15 women), aged between 19 and 26 years ($M = 21.9$, SD $= 1.88$). None of them participated in Experiment 1.

Materials, Design, and Procedure

Materials were the same as in Experiment 1. Two differences were introduced in the design and procedure.

First, practice type (retest vs. restudy) was manipulated within subjects, so that half of the critical categories were randomly assigned to be retested and another half were assigned to be restudied. (Note that there were no categories that did not receive one kind of practice, i.e., Nrp categories were not used in this experiment.) Again, only half of the members from each category underwent practice. In the practice phase retest and restudy trials were run in separate blocks, with two blocks in each practice cycle. Within each cycle, the order of retest and restudy blocks was counterbalanced between subjects.

Second, participants were tested once for all word pair right after the study phase. Trials in this initial test phase were identical to trials in the final test phase (also identical to the test phase of Experiment 1). To our knowledge, this was the first experiment using the retrieval practice paradigm that measured baseline recall levels as the performance on the first retrieval attempt after an initial study. All other aspects of this experiment were the same as those in Experiment 1.

Results and Discussion

The Effect of Practice on Final Recall

We conducted a repeated measures ANOVA on final recall data with item type (practiced vs. unpracticed) and practice type (retest vs. restudy) as repeated measures variables. Item type had a significant main effect on final recall, $F(1,29) = 141.31$, $p < 0.001$, $\eta_p^2 = 0.83$. There was no main effect of practice type, $F(1,29) = 0.96$, ns., and no interaction, $F(1,29) = 0.004$, ns. The same ANOVA on initial recall performance revealed that baseline performance did not differ in the four conditions, i.e., no main effect of item type, $F(1,29) = 0.10$, ns., practice type, $F(1,29) = 0.25$, ns., and no interaction, $F(1,29) = 0.09$, ns., emerged.

To detect RIF, we performed paired-samples t-tests for unpracticed items vs. their own baselines, i.e., recall performance of the same items at the first retrieval attempt, in the retest and the restudy condition separately. Looking at the data in **Figure 2**, it is not surprising that we found no significant RIF in either the retest, $t(29) = 0.00$, ns., or the restudy condition, $t(29) = 0.72$, ns. In brief, practice, either through restudying or retesting, did not induce forgetting when items had been retrieved once after

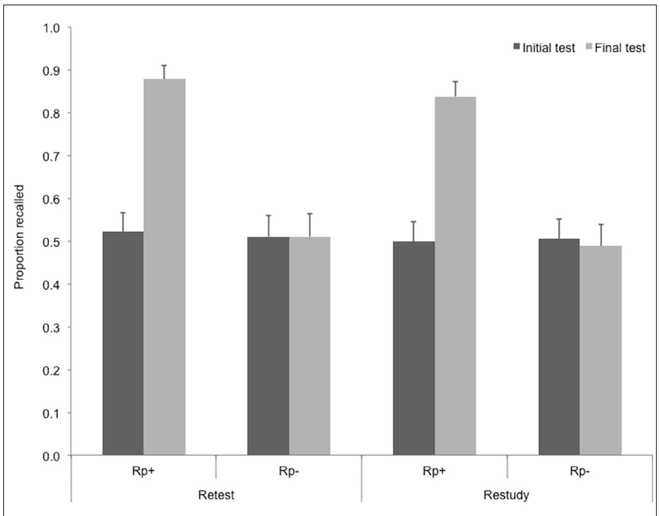


FIGURE 2 | Recall performance on the baseline test and the final test in Experiment 2, for the two item types in the two practice conditions. Rp+, Practiced words from practiced categories; Rp−, unpracticed words from practiced categories.

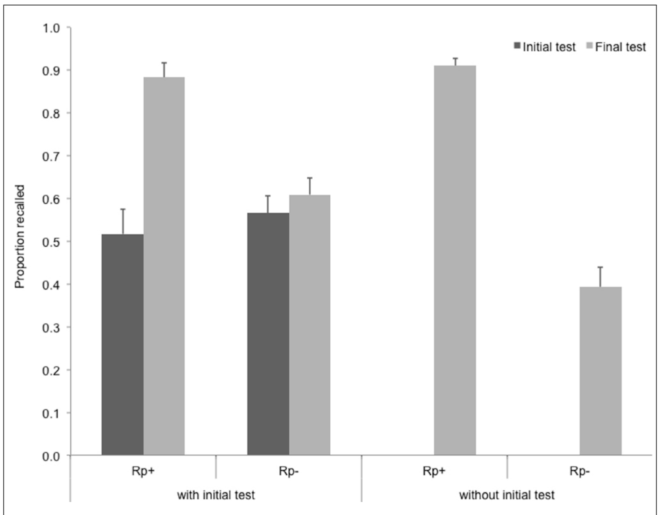


FIGURE 3 | Recall performance on the final test in Experiment 3, for the two item types in the two initial test conditions. Rp+ items with/without initial test: practiced words from categories that were/were not tested during an initial test, Rp− with/without initial test: unpracticed words from categories that were/were not tested during an initial test. Error bars indicate standard error of the mean.

the study phase. Retrieval practice led to enhancement of memory for practiced items (as compared to their own baselines, i.e., recall performance of the same items at the first retrieval attempt) in both conditions, $t(29) = 10.29$, $p < 0.001$, $r = 0.68$, in the restudy and $t(29) = 9.95$, $p < 0.001$, $r = 0.70$, in the retest condition.

In Experiment 2, we measured initial retrieval performance. Comparing the effect of selective retrieval and selective restudy to this initial retrieval performance, we found that practicing by means of both selective retrieval and selective restudy enhanced recall of practiced memories. We also found that neither type of practice (either retrieval or restudy) impaired accessibility of memories related to the cues associated to the practiced memories. This finding is not surprising in the condition where practice involved restudy—it is consistent with finding no RIF after selective restudy in Experiment 1, as well is many other experiments (Blaxton and Neely, 1983; Bäuml, 1996, 1997, 2002; Ciranni and Shimamura, 1999; Anderson et al., 2000; Anderson and Bell, 2001; Shivde and Anderson, 2001; Levy and Anderson, 2008). However, based on the predictions of inhibitory and interference explanations of RIF, the lack of RIF is indeed surprising in the condition where practice involved retrieval. In contrast, these results are in line with contextual accounts of RIF. These accounts predict that an initial retrieval of the entire learning set after the study phase will already have participants change their mental context and change contextual memory representation of studied items and later selective retrieval practice will cause no further change in this mental context and contextual memory representation (Racsmány and Conway, 2006; Jonker et al., 2013).

Although—as shown in Experiment 1—our paradigm was well-powered to detect a RIF effect if it existed, Experiment 2 did not allow directly testing the effect of initial testing, because it did not include a condition without initial testing. The goal of Experiment 3 was to allow for directly testing the impact of an initial test on the forgetting effect induced by later selective practice.

Experiment 3

Method

To analyze the effect of initial retrieval test in a single experiment, participants practiced word pairs during the practice phase through retrieval practice, and we varied whether categories received an initial test or not within subjects.

Participants

Thirty participants were recruited at the Budapest University of Technology and Economics. One participant's data were excluded from the analyses, because of a failure to type in the answers during the baseline test, therefore the final sample consisted of 29 individuals (16 men and 13 women), aged between 19 and 28 years ($M = 22.68$, SD $= 2.59$). None of them participated in previous experiments.

Materials, Design, and Procedure

Materials were the same as in Experiments 1 and 2. Two changes were introduced to the design and procedure of Experiment 2. First, initial test was administered only for half of the categories, and second, participants practiced by only retrieval and not by restudy for half of the words in all critical categories. Therefore this experiment did not involve a restudy condition, and item type (practiced vs. unpracticed) and initial test (administered vs. not administered) was varied within subjects. As in Experiment 2, each practice cycle contained two blocks. One block included trials with items that had received an initial test and the other block included trials with items that did not. Within each cycle, the order of blocks was counterbalanced between subjects.

All other aspects of this experiment were the same as those in Experiment 2.

Results and Discussion

The Effect of Practice on Final Recall

Recall performance at the final test for the four item types is shown in **Figure 3**, together with the initial test performance for items with initial test. A repeated measures ANOVA on final recall data with item type (practiced vs. unpracticed) and initial test (administered vs. not administered) as repeated measures variables. Item type had a significant main effect on final recall, $F(1,28) = 203.30$, $p < 0.001$, $\eta_p^2 = 0.88$, as well as the initial test, $F(1,28) = 17.05$, $p < 0.001$, $\eta_p^2 = 0.38$. Importantly, initial test interacted significantly with item type, $F(1,28) = 31.96$, $p < 0.001$, $\eta_p^2 = 0.53$. These effects occurred with initial performance not being different for practiced and unpracticed items, $t(28) = 1.11$, $p = 0.12$.

To assess RIF, we performed paired-samples t-tests contrasting unpracticed items and their available baselines. Note that in this experiment, we used initial test performance as a baseline—there were no categories that did not receive practice, i.e., Nrp categories were not used in this experiment. Therefore, final recall of Rp− items was contrasted with initial test performance of for Rp−, both for categories that received an initial test and for categories that did not. We found a significant RIF for Rp− items that did not receive an initial test, $t(28) = 4.23$, $p < 0.001$, $r = 0.41$, but no RIF for items that received an initial test, $t(28) = 1.61$, ns. Retrieval practice led to enhancement of memory for practiced items irrespective of whether initial test occurred or not, $t(28) = 9.04$, $p < 0.001$ for items without an initial test, $r = 0.72$, and $t(28) = 8.46$, $p < 0.001$ for items with an initial test, $r = 0.64$.

Using a within-subject design, Experiment 3 showed that initial testing can eliminate RIF due to later selective practice. However, one might argue that this experiment did not allow for calculating a classical RIF score ([Nrp−] − [Rp−]), as it did not include Nrp items. The goal of experiment 4 was to remedy this issue. This was important because the higher recall rate of Rp− items with initial test compared to recall rate of Rp− items without an initial test in Experiment 3 might have been the result of a higher rate of decay for initially non-tested information. The inclusion of unpracticed baseline items (Nrp) in Experiment 4 allowed us to test this alternative explanation.

Experiment 4

Method

This experiment was an extension of Experiment 1 with all items in the experiment receiving an initial test. This experiment allowed us to compare the effect of retrieval practice on the recall of different item types to a standard Nrp performance, i.e., items from categories with no selective retrieval practice, and also to an initial retrieval test performance measured for each item type.

Participants

Twenty-nine (13 men and 16 women), aged between 19 and 26 years ($M = 21.66$, SD = 2.16), participants were recruited at the Budapest University of Technology and Economics. None of them participated in previous experiments.

Materials, Design, and Procedure

Materials, design, and procedure were the same as in Experiment 1, with two major changes: First, an initial test was administered for all items. Second, participants practiced by only retrieval and not by restudy for half of the words from half the categories. The initial test phase was inserted immediately after the study phase. Trials in this initial test phase were identical to those in the same phase of Experiment 2.

Results and Discussion

The Effect of Practice on Final Recall

Recall performance at the final test as well as initial test performance for the four item types are shown in **Figure 4**. A repeated measures ANOVA on final recall data revealed a significant main effect of item type, $F(3,84) = 36.87$, $p < 0.001$, $\eta_p^2 = 0.57$. Importantly, in this experiment, we could assess RIF either by contrasting Rp− recall performance to Nrp− performance, i.e., items from categories with no selective retrieval practice, and also by contrasting initial test performance to final recall performance of the same Rp− items. Paired-samples t-tests showed that neither comparison yielded a significant RIF effect, $t(28) = 0.43$, ns., for the contrast with Nrp− items, and $t(28) = 0.24$, ns., for the contrast with initial Rp− recall performance. Similarly, retrieval practice led to enhancement of memory for practiced items both based on the contrast between Rp+ recall and Nrp+ recall, $t(28) = 11.91$, $r = 0.74$, and based on the contrast between Rp+ recall and initial Rp+ recall performance, $t(28) = 22.53$, $r = 0.86$.

Comparison of Retrieval-Induced Forgetting Across Experiments 1 and 4

Collapsing data across Experiments 1 and 4 also allowed us to compare classical RIF scores ([Nrp−] − [Rp−]) across two critical conditions—one with an initial test of all Nrp− and Rp− items, and one without it. Although this analysis was *post hoc*, it could provide converging evidence for the effect of initial test on RIF. To compare RIF scores across procedures with (Experiment 4) and without (Experiment 1) an initial test, we performed a mixed design ANOVA on recall data of Rp− and Nrp− items collapsing data from the two experiments (see **Figure 5**). This analysis revealed a tendency for an interaction of item type (Rp− vs. Nrp−) and initial test (with vs. without initial test), $F(1,56) = 3.842$, $p = 0.055$, $\eta_p^2 = 0.064$, indicating that initial testing of all items studied in the experiments reduced the RIF effect.

Importantly, we found no difference between the final recall of Nrp items with and without initial recall, $t(56) = 0.24$, ns., showing that the initial test itself did not change the studied items' forgetting rate. This also suggests that the different recall rates of Rp− items with and without initial test cannot be explained by faster forgetting without initial test.

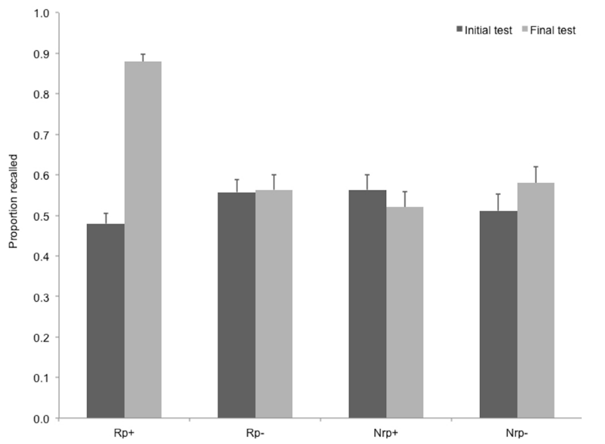


FIGURE 4 | Recall performance on the initial test and the final test in Experiment 4, for the four item types. Rp+, Practiced words from practiced categories; Rp−, unpracticed words from practiced categories; Nrp+, words from unpracticed categories used as baseline for Rp+ words; Nrp−, words from unpracticed categories used as baseline for Rp− words. Error bars indicate standard error of the mean.

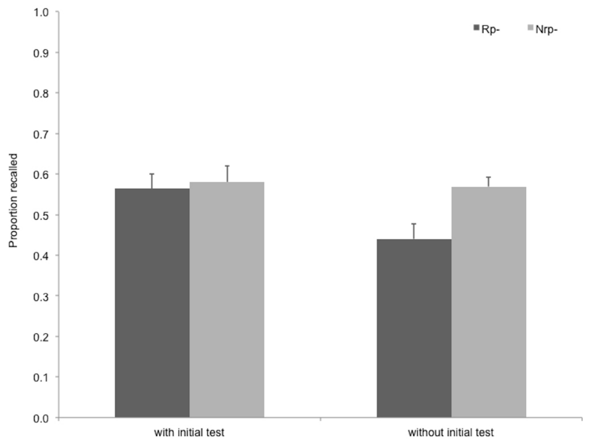


FIGURE 5 | Comparison of RIF as a function of whether an initial test took place or not. Data without initial test was collected in Experiment 1, data with initial test (testing all items shown in the study phase once, after the study phase, before the retrieval practice phase) was collected in Experiment 4. Rp−, unpracticed words from practiced categories; Nrp−, words from unpracticed categories used as baseline for Rp− words. Error bars indicate standard error of the mean.

Discussion

In four experiments we investigated the interactive effect of initial retrieval of the entire learning set and later selective practice of a part of studied items on final recall performance. We found that selectively practice—either by retrieval or by restudy—increased the recall probability of the practiced items on a final recall. However, only selective retrieval practice decreased final recall of the unpracticed members from the practiced categories in comparison with exemplars of unpracticed categories (Experiment 1). An initial retrieval of the learning set shielded against the adverse effect of retrieval practice; RIF was absent either when measured to baseline performance on the initial retrieval (Experiments 2 and 3), or to members of unpracticed categories (Experiment 4).

These results can be explained by assuming that selective retrieval, by shifting the context of the study phase to the context of retrieval practice phase, leads to RIF by generating a compound contextual episodic memory representation with a restricted and biased search set (Karpicke et al., 2014). In such contextualized memory sets, cue-item associations are biased toward increased recall probabilities for retrieved items from practiced categories and decreased recall probabilities for non-retrieved items from practiced categories (Racsmány and Conway, 2006; Jonker et al., 2013). In fact, these are genuine properties of episodic memories (Conway, 2009). On a more pragmatic point, these results also imply that the presence of RIF in any given experiment depends on the specific sequence of the experimental design—the selective practice phase must follow the study phase and no interim retrieval of studied items should take place in order to elicit RIF.

An inhibitory explanation of RIF is at odds with these results at a first glance. It is because this theory assumes that the unpracticed competitors would compete for retrieval during practice and this competition is then resolved by active inhibition, which renders competitors less accessible for later recall (Anderson et al., 1994, 2000; Anderson and Bell, 2001). It is reasonable to assume that unpracticed items would compete for retrieval during practice independently whether these items were retrieved previously or not. Moreover, accepting that initial retrieval strengthened these items, it is plausible to assume that competitors compete even stronger following initial retrieval testing. Therefore, later selective retrieval should induce forgetting on related competitors with and without initial testing of the entire learning set.

However, recently Anderson and Levy (2011) described some fundamental prerequisite for applying inhibition as an explanation for the presence or the lack of RIF in a given experiment. This is the demand/success trade off principle that is proposed to apply inhibitory explanation in a functional theoretical frame for RIF. This principle holds that the relation between interference of competitors and the size of inhibition follows a non-monotonic function (Anderson and Levy, 2011; Detre et al., 2013; see experimental evidence in Keresztes and Racsmány, 2013). That is because inhibition is imperfect and failure of inhibition will influence final accessibility of competitors, therefore RIF reflects the joint influence of inhibition demand and failure rate. An inhibitory theory can explain the lack of RIF following initial retrieval by either assuming that inhibition of competitors failed because of earlier retrieval of these items or by assuming that initial testing of competitors decreased the demand of inhibition. If the failure of inhibition diminished RIF following initial retrieval in our experiments, then we should assume that the same failure of inhibition influenced the success of practice phase too. As it was described by Anderson and Levy (2011) the same inhibitory processes should be active during the practice and the final recall phases of the experiment. Accepting this, we should assume that inhibition failure decreased the success rate of the practice phase and the benefit of practice on final recall of practiced exemplars. This is certainly not the case, both the practice success and the benefit of practice on final recall were the same in conditions with and without initial testing. Therefore failure of inhibition cannot explain the lack of RIF following initial retrieval. The other possibility is that initial retrieval decreased the demand of inhibition during practice, as a consequence there was no need to elicit inhibition on competitors. There is no direct way to test this hypothesis in our experiments, however, there are indirect evidences underlying this assumption in the literature of retrieval-enhanced learning (Roediger and Karpicke, 2006b).

A couple of recent experiments have found the retrieval of cue-target associations decreased the interference of these associations with learning of new associations to the same cues (Szpunar et al., 2008; Halamish and Bjork, 2011). Based on this, it seems plausible to assume the initial testing of the entire learning test significantly decreased the interference between exemplars during practice, and the level of interference between target and competitor items did not trigger inhibition. Although the lack of inhibition demand can be used in explaining the results of the present study by inhibition, inhibitory theories can offer no mechanism to explain why initial retrieval decreased the later demand of inhibition.

Interference theories of RIF are at odds with our results because these theories do not predict that an initial retrieval attempt should modulate the effect of later selective practice. The latest version of interference models (Verde, 2013), assumes that RIF is the result of a sampling failure. In this model, retrieval strengthens the context-item associations, whereas restudy strengthens cue-item associations. Accepting this, we should assume that initial retrieval of the entire learning set strengthens the context-item associations equally for targets and competitors. As a consequence, relative association strengths, which determine the sampling process, remain unaffected; the following selective retrieval practice should still lead to RIF.

Context-based accounts suggest mechanisms inherent to episodic retrieval processes to explain the current pattern of results. Context-based accounts of RIF and retrieval-enhanced learning (Jonker et al., 2013; Karpicke et al., 2014) emphasize the role of context change between initial study of category-member pairs on the one hand, and selective retrieval and final recall on the other. These accounts predict that an initial retrieval of the entire learning set after the study phase will already have participants change their mental context and later selective retrieval practice will cause no further change in this mental context. As a consequence, the context of the initial retrieval will be the active context at final recall.

Altogether context-based accounts of RIF assume that in the retrieval practice paradigm, selective retrieval restricts the search set through encoding a biased contextual information into an episodic memory representation, but an initial, non-selective, retrieval of the entire learning set before the selective retrieval can hinder this search set restriction.

A recent account of the testing effect—the episodic context account of retrieval-enhanced learning (Karpicke et al., 2014)—can be regarded as an extension of episodic and context-based accounts of RIF to a broader range of episodic memory phenomena. This theory aims to explain a range of long-term changes that occur as a consequence of retrieval. Although a detailed presentation of this theory is beyond the scope of the present paper, one relevant suggestion of it is that whenever studying and retrieval take place in different temporal contexts, retrieval will reinstate and update the study context by encoding a composite of study and retrieval contexts (see Karpicke et al., 2014; Lohnas and Kahana, 2014). On a later test participants will use the updated compound context to restrict the search set—the group of items considered as candidates for retrieval (Karpicke et al., 2014). According to this account, the retrieval practice paradigm involves manipulations that produce different kinds of contexts for practiced and unpracticed categories. That is selectively practiced categories will have the compound context of the study and the practice phases, whereas the unpracticed categories will have solely the context of the study phase. Another specificity of the retrieval practice paradigm is that participants typically retrieve practiced items more than once (the most frequently applied procedure involves three retrieval practice cycles). This procedure enables participants to encode strong and detailed contextual information for the practiced sets. As a consequence, they probably will rely more on the context of retrieval practice than on the context of study phase during final recall, and this will bias the recall output in favor of practiced items over unpracticed ones, as unpracticed items have no associations to context features of the practice phase. In contrast, participants will reinstate the context of the study phase whenever they use an unpracticed category label as a retrieval cue.

In other words, according to this account—also in line with the context-based explanations of RIF–RIF is due to a core attribute of retrieval; it is present when the updated context of the selective retrieval allows the participants to restrict their search set mainly for the practiced items. The initial retrieval in our experiments let participants to update the context of the study phase with the context of the initial retrieval. As a consequence, receiving the category cue they could use the compound context of study and initial retrieval while attempting to retrieve unpracticed items from practiced categories at final recall. In this view, retrieval is the key process that enhances long-term accessibility of retrieved memories and it is the process that can hinder retrieval of items through search set restriction or can shield against the adverse effect of later selective retrieval.

Acknowledgments

We thank Dora Bakos for her help in data collection. This work was supported by the KTIA NAP Grant (ID: 13-2-2014-0020).

References

Abel, M., and Bäuml, K. H. T. (2012). Retrieval-induced forgetting, delay, and sleep. *Memory* 20, 420–428. doi: 10.1080/09658211.2012.671832

Anderson, M. C. (2003). Rethinking interference theory: executive control and the mechanisms of forgetting. *J. Mem. Lang.* 49, 415–445. doi: 10.1016/j.jml.2003.08.006

Anderson, M. C. (2005). "The role of inhibitory control in forgetting unwanted memories: a consideration of three methods," in *Dynamic Cognitive Processes*, eds N. Ohta, C. M. MacLeod, and B. Uttl (Tokyo: Springer), 159–189.

Anderson, M. C., and Bell, T. (2001). Forgetting our facts: the role of inhibitory processes in the loss of propositional knowledge. *J. Exp. Psychol. Gen.* 130, 544–570. doi: 10.1037/0096-3445.130.3.544

dc_1204_16

Anderson, M. C., Bjork, E. L., and Bjork, R. A. (2000). Retrieval-induced forgetting: evidence for a recall-specific mechanism. *Psychon. Bull. Rev.* 7, 522–530. doi: 10.3758/BF03214366

Anderson, M. C., Bjork, R. A., and Bjork, E. L. (1994). Remembering can cause forgetting: retrieval dynamics in long-term memory. *J. Exp. Psychol. Learn. Mem. Cogn.* 20, 1063–1087. doi: 10.1037/0278-7393.20.5.1063

Anderson, M. C., and Levy, B. J. (2007). "Theoretical issues in inhibition: insights from research on human memory," in *Inhibition in Cognition*, eds D. S. Gorfein and C. M. MacLeod (Washington, DC: American Psychological Association) 81–102.

Anderson, M. C., and Levy, B. J. (2011). "On the relationship between interference and inhibition in cognition," in *Successful Remembering and Successful Forgetting: A Festschrift in Honor of Robert A. Bjork*, ed. A. Benjamin (London: Psychology Press), 107–132.

Anderson, M. C., and McCulloch, K. C. (1999). Integration as a general boundary condition on retrieval-induced forgetting. *J. Exp. Psychol. Learn. Mem. Cogn.* 25, 608–629. doi: 10.1037/0278-7393.25.3.608

Anderson, M. C., and Spellman, B. A. (1995). On the status of inhibitory mechanisms in cognition: memory retrieval as a model case. *Psychol. Rev.* 102, 68–100. doi: 10.1037/0033-295X.102.1.68

Bäuml, K. H. (1996). Revisiting an old issue: retroactive interference as a function of the degree of original and interpolated learning. *Psychon. Bull. Rev.* 3, 380–384. doi: 10.3758/BF03210765

Bäuml, K. H. (1997). The list-strength effect: strength-dependent competition or suppression? *Psychon. Bull. Rev.* 4, 260–264. doi: 10.3758/BF03209403

Bäuml, K. H. (2002). Semantic generation can cause episodic forgetting. *Psychol. Sci.* 13, 356–360. doi: 10.1111/j.0956-7976.2002.00464.x

Bäuml, K. H., and Aslan, A. (2004). Part-list cuing as instructed retrieval inhibition. *Mem. Cogn.* 32, 610–617. doi: 10.3758/BF03195852

Bäuml, K. H., and Hartinger, A. (2002). On the role of item similarity in retrieval-induced forgetting. *Memory* 10, 215–224. doi: 10.1080/09658210143000362

Blaxton, T. A., and Neely, J. H. (1983). Inhibition from semantically related primes: evidence of a category-specific inhibition. *Mem. Cogn.* 11, 500–510.

Camp, G., Pecher, D., and Schmidt, H. G. (2007). No retrieval-induced forgetting using item-specific independent cues: evidence against a general inhibitory account. *J. Exp. Psychol. Learn. Mem. Cogn.* 33, 950–958. doi: 10.1037/0278-7393.33.5.950

Camp, G., Pecher, D., Schmidt, H. G., and Zeelenberg, R. (2009). Are independent probes truly independent? *J. Exp. Psychol. Learn. Mem. Cogn.* 35, 934–942. doi: 10.1037/a0015536

Carrier, M., and Pashler, H. (1992). The influence of retrieval on retention. *Mem. Cogn.* 20, 633–642. doi: 10.3758/BF03202713

Ciranni, M. A., and Shimamura, A. P. (1999). Retrieval-induced forgetting in episodic memory. *J. Exp. Psychol. Learn. Mem. Cogn.* 25, 1403–1414. doi: 10.1037/0278-7393.25.6.1403

Conway, M. A. (2009). Episodic memories. *Neuropsychologia* 47, 2305–2313. doi: 10.1016/j.neuropsychologia.2009.02.003

Detre, G. J., Natarajan, A., Gershman, S. J., and Norman, K. A. (2013). Moderate levels of activation lead to forgetting in the think/no-think paradigm. *Neuropsychologia* 51, 2371–2388. doi: 10.1016/j.neuropsychologia.2013.02.017

Faul, F., Erdfelder, E., Buchner, A., and Lang, A. G. (2009). Statistical power analyses using G* Power 3.1: tests for correlation and regression analyses. *Behav. Res. Methods* 41, 1149–1160.

Halácsy, P., Kornai, A., Németh, L., Rung, A., Szakadát, I., and Trón, V. (2004). "Creating open language resources for Hungarian," in *Proceedings of the 4th International Conference on Language Resources and Evaluation, LREC2004* (Lisbon: European Language Resources Association). p. 1201–1204.

Halamish, V., and Bjork, R. A. (2011). When does testing enhance retention? A distribution-based interpretation of retrieval as a memory modifier. *J. Exp. Psychol. Learn. Mem. Cogn.* 37, 801–812. doi: 10.1037/a0023219

Jakab, E., and Raaijmakers, J. G. (2009). The role of item strength in retrieval-induced forgetting. *J. Exp. Psychol. Learn. Mem. Cogn.* 35, 607–617. doi: 10.1037/a0015264

Jonker, T. R., and MacLeod, C. M. (2012). Retrieval-induced forgetting: testing the competition assumption of inhibition theory. *Can. J. Exp. Psychol.* 66, 204–211. doi: 10.1037/a0027277

Jonker, T. R., Seli, P., and MacLeod, C. M. (2013). Putting retrieval-induced forgetting in context: an inhibition-free, context-based account. *Psychol. Rev.* 120, 852–872. doi: 10.1037/a0034246

Karpicke, J. D., and Bauernschmidt, A. (2011). Spaced retrieval: absolute spacing enhances learning regardless of relative spacing. *J. Exp. Psychol. Learn. Mem. Cogn.* 37, 1250–1257.

Karpicke, J. D., Lehman, M., and Aue, W. R. (2014). "Retrieval-based learning: an episodic context account," in *The Psychology of Learning and Motivation*, Vol. 61, ed. B. H. Ross (San Diego, CA: Elsevier Academic Press), 238–284.

Karpicke, J. D., and Roediger, H. L. (2008). The critical importance of retrieval for learning. *Science* 319, 966–968. doi: 10.1126/science.1152408

Keresztes, A., Kaiser, D., Kovács, G., and Racsmány, M. (2013). Testing promotes long-term learning via stabilizing activation patterns in a large network of brain areas. *Cereb Cortex* 24, 3025–3035. doi: 10.1093/cercor/bht158

Keresztes, A., and Racsmány, M. (2013). Interference resolution in retrieval-induced forgetting: behavioral evidence for a nonmonotonic relationship between interference and forgetting. *Mem. Cogn.* 41, 511–518. doi: 10.3758/s13421-012-0276-3

Kornai, A., Halácsy, P., Nagy, V., Oravecz, C., Trón, V., and Varga, D. (2006). "Web-based frequency dictionaries for medium density languages," in *Proceedings of the 2nd International Workshop on Web as Corpus*, eds A. Kilgarriff and M. Baroni (East Stroudsburg, PA: ACL), 1–8.

Levy, B. J., and Anderson, M. C. (2008). Individual differences in the suppression of unwanted memories: the executive deficit hypothesis. *Acta Psychol.* 127, 623–635. doi: 10.1016/j.actpsy.2007.12.004

Lohnas, L. J., and Kahana, M. J. (2014). Compound cuing in free recall. *J. Exp. Psychol. Learn. Mem. Cogn.* 40, 12–24.

MacLeod, M. D., and Macrae, C. N. (2001). Gone but not forgotten: the transient nature of retrieval-induced forgetting. *Psychol. Sci.* 12, 148–152. doi: 10.1111/1467-9280.00325

Malmberg, K. J., and Shiffrin, R. M. (2005). The "one-shot" hypothesis for context storage. *J. Exp. Psychol. Learn. Mem. Cogn.* 31, 322–336. doi: 10.1037/0278-7393.31.2.322

Norman, K. A., Newman, E. L., and Detre, G. (2007). A neural network model of retrieval-induced forgetting. *Psychol. Rev.* 114, 887–953. doi: 10.1037/0033-295X.114.4.887

Raaijmakers, J. G., and Jakab, E. (2012). Retrieval-induced forgetting without competition: testing the retrieval specificity assumption of the inhibition theory. *Mem. Cogn.* 40, 19–27. doi: 10.3758/s13421-011-0131-y

Raaijmakers, J. G., and Shiffrin, R. M. (1981). Search of associative memory. *Psychol. Rev.* 88, 93–134. doi: 10.1037/0033-295X.88.2.93

Racsmány, M., and Conway, M. A. (2006). Episodic inhibition. *J. Exp. Psychol. Learn. Mem. Cogn.* 32, 44–57. doi: 10.1037/0278-7393.32.1.44

Racsmány, M., Conway, M. A., and Demeter, G. (2010). Consolidation of episodic memories during sleep long-term effects of retrieval practice. *Psychol. Sci.* 21, 80–85. doi: 10.1177/0956797609354074

Racsmány, M., Conway, M. A., Keresztes, A., and Krajcsi, A. (2012). Inhibition and interference in the think/no-think task. *Mem. Cogn.* 40, 168–176. doi: 10.3758/s13421-011-0144-6

Roediger, H. L., and Karpicke, J. D. (2006a). Test-enhanced learning taking memory tests improves long-term retention. *Psychol. Sci.* 17, 249–255. doi: 10.1111/j.1467-9280.2006.01693.x

Roediger, H. L., and Karpicke, J. D. (2006b). The power of testing memory: basic research and implications for educational practice. *Perspect. Psychol. Sci.* 1, 181–210. doi: 10.1111/j.1745-6916.2006.00012.x

Sahakyan, L., and Hendricks, H. E. (2012). Context change and retrieval difficulty in the list-before-last paradigm. *Mem. Cogn.* 40, 844–860. doi: 10.3758/s13421-012-0198-0

Sahakyan, L., and Kelley, C. M. (2002). A contextual change account of the directed forgetting effect. *J. Exp. Psychol. Learn. Mem. Cogn.* 28, 1064–1072.

Shivde, G., and Anderson, M. C. (2001). "The role of inhibition in meaning selection: insights from retrieval-induced forgetting," in *On the Consequences of Meaning Selection: Perspectives on Resolving Lexical Ambiguity*, ed. D. Gorfein (Washington, DC: American Psychological Association), 175–190.

Staudigl, T., Hanslmayr, S., and Bäuml, K. H. T. (2010). Theta oscillations reflect the dynamics of interference in episodic memory retrieval. *J. Neurosci.* 30, 11356–11362. doi: 10.1523/JNEUROSCI.0637-10.2010

Storm, B. C. (2011). "Retrieval-induced forgetting and the resolution of competition," in *Successful Remembering and Successful Forgetting: A Festschrift*

dc_1204_16

in Honor of Robert A. Bjork, ed. A. Benjamin (London: Psychology Press), 89–105.

Storm, B. C., Bjork, E. L., and Bjork, R. A. (2008). Accelerated relearning after retrieval-induced forgetting: the benefit of being forgotten. *J. Exp. Psychol. Learn. Mem. Cogn.* 34, 230–236.

Storm, B. C., Bjork, E. L., and Bjork, R. A. (2012). On the durability of retrieval-induced forgetting. *J. Cogn. Psychol.* 24, 617–629. doi: 10.1080/20445911.2012.674030

Storm, B. C., Bjork, E. L., Bjork, R. A., and Nestojko, J. F. (2006). Is retrieval success a necessary condition for retrieval-induced forgetting? *Psychon. Bull. Rev.* 13, 1023–1027. doi: 10.3758/BF03213919

Storm, B. C., and Nestojko, J. F. (2010). Successful inhibition, unsuccessful retrieval: manipulating time and success during retrieval practice. *Memory* 18, 99–114.

Szpunar, K. K., McDermott, K. B., and Roediger, H. L. III. (2008). Testing during study insulates against the buildup of proactive interference. *J. Exp. Psychol. Learn. Mem. Cogn.* 34, 1392–1399. doi: 10.1037/a0013082

Toppino, T. C., and Cohen, M. S. (2009). The testing effect and the retention interval. *Exp. Psychol.* 56, 252–257. doi: 10.1027/1618-3169.56.4.252

Verde, M. F. (2009). The list-strength effect in recall: relative-strength competition and retrieval inhibition may both contribute to forgetting. *J. Exp. Psychol. Learn. Mem. Cogn.* 35, 205–220. doi: 10.1037/a0014275

Verde, M. F. (2013). Retrieval-induced forgetting in recall: competitor interference revisited. *J. Exp. Psychol. Learn. Mem. Cogn.* 39, 1433–1448. doi: 10.1037/a0032975

Wheeler, M., Ewers, M., and Buonanno, J. (2003). Different rates of forgetting following study versus test trials. *Memory* 11, 571–580. doi: 10.1080/09658210244000414

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

VI.2. Long-term effects of selective retrieval (Study 8)

Paper: Consolidation of episodic memories during sleep: long-term effects of retrieval practice. Mihály Racsmány, Martin A Conway, Gyula Demeter, *Psychological Science* 01/2010; 21(1):80-5. DOI:10.1177/0956797609354074

dc_1204_16

Research Article

Psychological Science
21(1) 80–85

Reprints and permission: http://www.sagepub.com/journalsPermissions.nav
DOI: 10.1177/0956797609354074
http://pss.sagepub.com

Consolidation of Episodic Memories During Sleep: Long-Term Effects of Retrieval Practice

Mihály Racsmány[1,2], Martin A. Conway[3], and Gyula Demeter[4]
[1]Department of Cognitive Science, Budapest University of Technology and Economics; [2]Institute of Psychology, University of Szeged; [3]The Leeds Memory Group, Institute of Psychological Sciences, University of Leeds; and [4]Research Group on Cognitive Science, Hungarian Academy of Sciences and Budapest University of Technology and Economics

Abstract

Two experiments investigated the long-term effects of retrieval practice. In the retrieval-practice procedure, selected items from a previously studied list are repeatedly recalled. The typical retrieval-practice effects are considerably enhanced memory for practiced items accompanied by low levels of recall, relative to baseline, for previously studied items that are associated with the practiced items but were not themselves practiced. The two experiments demonstrated that the former effect persisted over 12 hr; the latter effect also persisted over 12 hr, but only if a period of nocturnal sleep occurred during the retention interval. We propose that consolidation processes occurring during sleep, and possibly featuring some form of off-line rehearsal, mediate these long-term effects of retrieval practice.

Keywords

retrieval practice, episodic memory, sleep, consolidation, rehearsal

Received 1/20/09; Revision accepted 4/29/09

It has long been thought that sleep plays a crucial role in the consolidation of recently formed memories. Current evidence shows that retention of procedural knowledge can be enhanced by a period of sleep (Stickgold & Walker, 2005), as can retention of motor skills (Walker, Brakefield, Morgan, Hobson, & Stickgold, 2002). In a recent programmatic series of studies, Gaskell, Dumay, and their coworkers demonstrated that sleep is critical to the retention of new vocabulary, and in particular to the integration of newly acquired words into the lexicon (see Dumay & Gaskell, 2007, for a review). Furthermore, it has been observed that relatively few episodic memories formed during a day are retained the following day, which suggests that only a minority of episodic memories are selected for enduring retention (Conway, 2009; Williams, Conway, & Baddeley, 2008). According to one view, consolidation processes operating during sleep mediate these effects. Reactivation of the medial temporal lobe memory system, and especially hippocampal circuits, may be the locus of sleep-mediated consolidation (Wilson & McNaughton, 1994). Other brain areas have been implicated too, and it seems that networks in medial prefrontal cortex, operating at faster processing rates during sleep than during awake periods, rapidly and repeatedly replay processing sequences featured in the immediately preceding awake period (Euston, Tatsuno, & McNaughton, 2007). This mechanism may consist of a sequence of speeded off-line rehearsal, and possibly it is these intense bursts of rehearsal that lead to the consolidation of recent experience in long-term memory.

Consolidation processes operating during a nocturnal sleep cycle should influence the retention of recently formed episodic memories, and we explored this idea in two experiments using the retrieval-practice procedure (Anderson, Bjork, & Bjork, 1994; Racsmány & Conway, 2006). In this procedure, participants first study a list of words and then selectively practice recalling a subset of the list. Memory is then tested, typically by cued recall. The retrieval-practice procedure is particularly suited to exploring the consolidation of episodic memories, as it is thought that the study phase gives rise to the formation of an episodic memory of learning the study list and that the later practice phase gives rise to a pattern of activation and inhibition over the contents of the episodic memory. It is this pattern of activation and inhibition that mediates later access to memory content and that gives rise to the characteristic pattern of recall seen on the memory test (Racsmány & Conway, 2006). By this view, retrieval practice should give rise to long-term patterns of activation and inhibition that are strengthened by consolidation during sleep. The sole previous

Corresponding Author:
Mihály Racsmány, Department of Cognitive Science, Budapest University of Technology and Economics, Stoczek u. 2., Budapest, 1111, Hungary
E-mail: racsmany@cogsci.bme.hu

study of long-term retrieval-practice effects indicates that this may indeed be the case (MacLeod & Macrae, 2001), although we acknowledge that other researchers consider the effects of retrieval practice to be more likely short-term than long-term (Saunders & MacLeod, 2002; but see Anderson, 2001), and at least one current model (Norman, Newman, & Detre, 2007) proposes that REM sleep may "reset" inhibitory patterns. The retrieval-practice procedure was well suited for our study because it easily allows a period of sleep or equivalent period of wakefulness to be interposed between practice and test.

Experiment 1

In the retrieval-practice procedure, exemplars from various categories are first studied. After the study phase, selected items from selected categories are then repeatedly recalled, typically three times, in response to cues consisting of a category name plus word fragment. For example, if "fruit-orange" is a studied item, "fruit-o_____?" might be a retrieval-practice cue. The three phases of study, practice, and test usually are separated only by the few minutes required to give the instructions for each phase. The design yields three types of items: items that have been practiced (Rp+), items that have not themselves been practiced but that originate from a category for which another item has been practiced (Rp–), and items from categories for which no items have been practiced (Nrp).

The typical finding is that memory for Rp+ items is highest, memory for Nrp items is at an intermediate level, and memory for Rp– items is poorest. This pattern is taken to indicate strong activation of Rp+ items resulting from retrieval practice making these items highly accessible to recall, weaker activation of Nrp items, and inhibition of Rp– items (Anderson & Spellman, 1995; Bjork, Bjork, & Anderson, 1998; Racsmány & Conway, 2006; Storm, Bjork, Bjork, & Nestojko, 2006). According to this explanation, practice recalling an item from a previously studied set of category exemplars induces inhibition of exemplars that are not practiced and that could potentially compete with and disrupt recall of the cued items (cf. Anderson & Levy, 2007). Thus, studying "apple," "pear," and "orange" and then repeatedly practicing recall of only "orange" induces inhibition of "apple" and "pear." The net result is that memory for "apple" and "pear" (Rp– items) is hurt, whereas memory for "orange" (an Rp+ item) is enhanced. Other interpretations of these effects of retrieval practice have emphasized the role of interference rather than inhibition (e.g., Camp, Pecher, & Schmidt, 2007; see also Mensink & Raaijmakers, 1988).

In our first experiment, participants were assigned to two groups: a sleep group and a no-sleep group. The sleep group studied and practiced the items in the evening; the following morning, some 12 hr later and after their usual period of nocturnal sleep, their memory for the items was tested. The no-sleep group studied and practiced the items in the morning; 12 hr later, in the evening, their memory was tested.[1] We expected that the no-sleep group would not show the typical retrieval-practice effect and instead would simply show forgetting of the items. In contrast, and assuming that consolidation can enhance retention, we expected the sleep group to show the usual retrieval-practice pattern.

Method

Participants. Sixty-four undergraduate Hungarian students from the Budapest University of Technology and Economics (32 females, 32 males) participated in return for partial credit in an introductory psychology course. Their ages ranged from 19 to 26 years. There were 32 participants each in the sleep and no-sleep groups (16 females and 16 males randomly assigned within gender to each group). Note that all participants in the sleep group were tested a minimum of 1 hr after awakening.

Materials. Following Anderson et al. (1994), we used 10 categories, 2 of which were fillers. Each target and filler category consisted of 6 exemplars. Exemplars were moderate- to high-frequency words drawn from two Hungarian word-frequency norms (Füredi & Kelemen, 1989; Kónya & Pintér, 1985). For each subject, 4 target categories were practiced and 4 were nonpracticed; across subjects, each target category was equally often practiced and nonpracticed. The practiced and nonpracticed exemplars from practiced categories were counterbalanced over participants. In sum, in each learning session, participants learned 60 exemplars from 10 categories (2 of which were fillers), practiced 18 exemplars (including 6 fillers) from 6 categories (including the 2 filler categories), and finally tried to recall 60 exemplars (including 12 fillers) from the original 10 categories. During both practice and final cued recall, items from filler categories were always in the first and last positions in order to avoid the confounding effect of category position.

Procedure. Participants were randomly assigned to either the sleep or the no-sleep group. All participants completed a short questionnaire about the length and quality of their sleep period prior to the experiment. Those who had slept less than 4 hr or used sleeping pills were excluded from the experiment. The no-sleep group completed the sleep questionnaire only on the day of the experiment, answering the questions with reference to the previous night's sleep, whereas the sleep group completed the same sleep questionnaire on the day of the study phase and also on the day of the recall test, in each case answering the questions with reference to the previous night's sleep.

At 8 p.m., the sleep group completed the study phase followed by the practice phase; these participants returned to the laboratory for the surprise delayed recall test at 8 a.m. the following morning. Note that in all cases the test was given a minimum of 1 hr after awakening. The no-sleep group completed the study phase and practice phase at 8 a.m. and the surprise recall test at 8 p.m. on the same day. Neither group knew that they were returning to take a memory test; rather, all participants were led to believe that they were returning to take part in a new and unrelated experiment.

In the study phase, participants were instructed that category-exemplar pairs would be presented on a computer screen and that they should study the pairs in preparation for a later memory test. Each category-exemplar pair was presented in uppercase letters in the center of the screen for 5 s. Order of presentation was semirandomized; exemplars from the same category did not appear on consecutive trials. When participants had completed the study phase, the experimenter distributed retrieval-practice booklets. Participants believed that this second phase was the memory test. Each page in the booklet showed one of the category names studied previously and the first two letters of one member of that category, also studied previously. Participants were instructed to complete the exemplar fragment with one of the words they had studied earlier. They were informed that some of the exemplars might be tested more than once and that in those cases they should respond with the remembered item. Rp+ items were repeated three times. At the end of the retrieval-practice phase, the booklets were collected, and participants were sent home for 12 hr. When they returned to the laboratory, they were given cued-recall booklets, in which the name of one of the previously studied categories appeared at the top of each page. Participants were instructed to recall as many examples as they could for each category in the 10-min period allocated for this test. They were instructed to complete the pages in order and not to return to a previous category once they had turned the page in the recall booklet. Order of presentation of the target categories was counterbalanced across participants.

Results

Planned comparisons revealed that the critical contrast of Nrp with Rp– items was reliable only in the sleep group, $t(1, 31) = -3.7$, $p_{rep} = .99$, $r = .55$ (for the no-sleep group, $t < 1$). Thus, the retrieval-practice effect was observed only in the sleep group (see Fig. 1 for mean percentages).[2] An independent t test revealed that there was no reliable difference between the two groups' recall of Rp+ items $t(62) = -1.12$; the long-term beneficial effect of selective practice (relative to baseline—i.e., Nrp items) was similar in the two groups. Debriefing interviews uncovered no evidence of conscious, intentional rehearsal in either group, and participants indicated that they were generally surprised by the delayed cued-recall test.

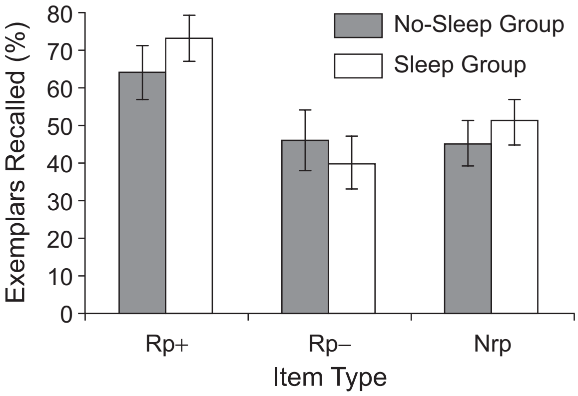


Fig. 1. Mean percentages of exemplars recalled by the sleep and no-sleep groups in Experiment 1. Results are shown separately for practiced items (Rp+), unpracticed items from practiced categories (Rp–), and unpracticed items from unpracticed categories (Nrp). Error bars show 95% confidence intervals.

Experiment 2

A problem with the retrieval-practice procedure is that although it may induce inhibition of Rp– items, performance on Rp– items must almost certainly also be impaired by output interference from Rp+ items. Given that we were primarily interested in the effects of sleep on memory performance in the retrieval-practice procedure, this was in some respects a secondary issue. Nevertheless, in order to reduce the potential effects of output interference, and also to further examine the effects of sleep on retrieval practice, we decided to run a replication of Experiment 1 in which output was more directly controlled. To achieve such control, we constructed a new study set in which the first letter of each word was unique within its category. At test, participants were cued with the category names and the first letters of studied items. Using these cues, we were able to control the order in which items were recalled. In addition, to control for potential time-of-day effects, we included a new control group who studied and practiced items at 8 a.m. and were then given the surprise recall test 1 hr later; we refer to this group as the morning no-sleep group. We reasoned that if the morning no-sleep group showed the retrieval-practice effect, then this effect might be attributable to the time of day of the test, rather than a period of sleep intervening between study and test.

Method

Participants. A new cohort of 96 undergraduate Hungarian students from the Budapest University of Technology and Economics (48 females, 48 males) participated in return for partial credit in an introductory psychology course. Their ages ranged from 20 to 28 years. There were 32 subjects in each of the three groups (16 females and 16 males randomly assigned within gender to each group). All participants in the sleep group and in the morning no-sleep group were tested a minimum of 1 hr after awakening.

Materials. Following Anderson et al. (1994), we used 10 categories, 2 of which were fillers. Each target and filler category consisted of 6 exemplars (as in Experiment 1). The exemplars were moderate- to high-frequency words drawn from two Hungarian word-frequency norms (Füredi & Kelemen, 1989; Kónya & Pintér, 1985). For each subject, 4 target categories were practiced and 4 were nonpracticed; across subjects, each

dc_1204_16

target category was equally often practiced and nonpracticed. The practiced and nonpracticed exemplars from practiced categories were counterbalanced over participants. The study list contained 6 words from each of the 10 categories, for a total of 60 words. Within each category, every word had a unique initial letter, so that a category name and first letter could serve as a specific cue for each target word.

Procedure. The procedure was the same as in Experiment 1 with the exception of changes in the cued-recall test. In that test, the cues appeared on a computer screen one at a time for 5 s each. Each cue consisted of a category name together with the first letter of one of the studied exemplars (e.g., "Fruit – O______"). Items for each category were presented as a block; the cues for the three Rp– words were always presented first, in random order, and the cues for the Nrp and Rp+ items were then presented ungrouped and in random order. Participants wrote their responses in a response booklet.

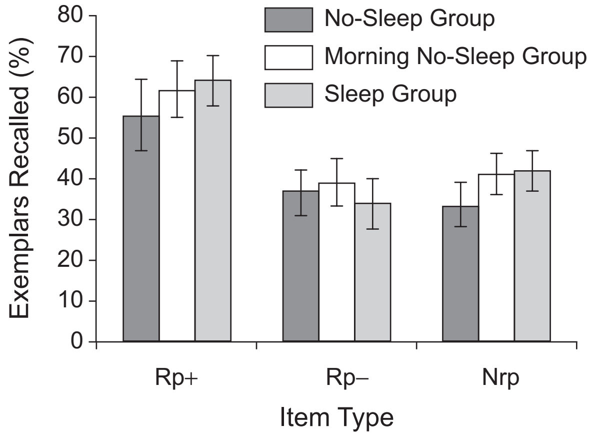


Fig. 2. Mean percentages of exemplars recalled by the sleep, no-sleep, and morning no-sleep groups in Experiment 2. Results are shown separately for practiced items (Rp+), unpracticed items from practiced categories (Rp–), and unpracticed items from unpracticed categories (Nrp). Error bars show 95% confidence intervals.

Results

Planned comparisons found that the critical contrast of Nrp with Rp– items was reliable in the sleep group, $t(1, 28) = -2.43$, $p_{rep} = .95$, $r = .43$, but not in the no-sleep group and the morning no-sleep group, *t*s < 1.2. Thus, the retrieval-practice effect was observed only in the sleep group (see Fig. 2). A one-way independent analysis of variance found no reliable difference between groups on recall of Rp+ items, $F < 1.2$, showing that the long-term beneficial effect of selective practice was present to the same degree in all groups (see Fig. 2). The debriefing interviews again indicated that participants did not rehearse items in the retention interval, and that all participants were surprised by the memory test. In sum, the overall pattern of findings replicated the pattern observed in Experiment 1 and indicates that output interference and time-of-day differences had little, or possibly no, influence in the two experiments.

Discussion

One account of the effects of retrieval practice posits that they are mediated by an episodic memory of the study phase (Racsmány & Conway, 2006). According to this view, which we term the *episodic-inhibition hypothesis* to distinguish it from accounts focusing on other types and sources of inhibition in long-term memory, retrieval practice establishes a pattern of activation and inhibition over the contents or features of an episodic memory of the study phase. As the episodic memory is consolidated in long-term memory, the pattern of activation and inhibition, which determines the accessibility of the contents of the memory, stabilizes and becomes resistant to further change. One major mechanism of this process of consolidation is rehearsal. According to the episodic-inhibition hypothesis, as a memory is repeatedly retrieved and its contents are accessed, its durability in long-term memory increases, and the accessibility levels of its contents become fixed (Racsmány & Conway, 2006; Racsmány, Conway, Garab, & Nagymáté, 2008).

The present findings suggest that sleep is important to this process of consolidation, as indeed other researchers using different procedures have also observed (e.g., Drosopoulos, Wagner, & Born, 2005). The findings of Experiment 2 indicate that retrieval-practice effects begin to dissipate after a retention interval of just 1 hr in the absence of rehearsal. Interestingly, in a related experiment not reported here, we found that if there is rehearsal in the retention interval, then retrieval-practice effects can be maintained over at least 12 hr (with no period of sleep). We use the term *rehearsal* here in a slightly nonstandard way, as according to our episodic-inhibition view, rehearsal occurs when a memory is activated and the pattern of activation and inhibition over its contents is instantiated. Such rehearsal does not have to occur consciously or intentionally, although, of course, it might. We suggest that when rehearsal occurs in this way, it approximates what has been termed *elaborative rehearsal* (Craik & Lockhart, 1972), and it promotes the integration of the memory with other memories and knowledge structures in autobiographical memory (see Conway, 2009). It is perhaps the degree and nature of the integration that determines the durability of access to a memory and its contents. Clearly, other memories formed during the retention interval may reduce integration, prevent it, or interfere with it in some other way. It seems likely that the opportunity for interference by new memories was greater in our no-sleep than in our sleep groups, and, consequently, integration may have been attenuated in the no-sleep relative to the sleep groups. (Note that this would not have been the case if rehearsal had been intentionally undertaken during the retention interval.)

dc_1204_16

According to this reasoning, the greater degree of integration of memories in the sleep groups underlies the long-term retrieval-practice effects we observed in these groups. This integration is, perhaps, similar in kind to the integration of new words with the lexicon found to occur after periods of nocturnal sleep (Dumay & Gaskell, 2007).

These novel long-term, sleep-related, retrieval-practice effects lend some support to suggestions that spontaneously occurring retrieval practice in everyday cognition may mediate aspects of remembering and forgetting (e.g., Anderson, 2001). But we can now add to this idea the notion that consolidation and integration processes occurring during sleep are also important in maintaining access to memories and their contents. The present findings demonstrate that consolidation of recently formed episodic memories during sleep may be integral to the normal functioning of episodic memory.

Declaration of Conflicting Interests

The authors declared that they had no conflicts of interests with respect to their authorship and/or the publication of this article.

Funding

Financial support for this research was provided by OTKA (Hungarian National Science Foundation) Grants K68463 and IN77932. Martin A. Conway is supported by a Professorial Fellowship (RES-051-27-0127) from the Economic and Social Research Council of Great Britain. Mihály Racsmány is grantee of the Bolyai János Research Scholarship of the Hungarian Academy of Science.

Notes

1. A very important design feature is that the memory test is unexpected. In a separate experiment not reported here, we found that long-term retrieval-practice effects can occur if participants rehearse the retrieval-practice items during the retention interval.
2. An alternative interpretation might focus on diurnal effects, such as the awakening cortisol response (ACR), which is thought to influence memory. However, as the ACR peaks and then begins to decline within 30 to 45 min following sleep (Clow, Thorn, Evans, & Hucklebridge, 2004), and all participants were tested at least 1 hr after awakening (and most were tested 90 to 120 min postsleep), it seems unlikely that the ACR could have directly influenced memory performance in the sleep group. Moreover, although cortisol levels begin to fall toward the onset of sleep and are at their lowest levels in the first 3 to 4 hr of sleep, all participants in the no-sleep group were tested several hours prior to sleep, and there is no reason to suppose that their cortisol levels had changed systematically at this point in the sleep/wake cycle. Thus, the sleep and no-sleep groups most likely had highly similar diurnal cortisol levels.

References

Anderson, M.C. (2001). Active forgetting: Evidence for functional inhibition as a source of memory failure. *Journal of Aggression, Maltreatment and Trauma*, *4*, 185–210.

Anderson, M.C., Bjork, E.L., & Bjork, R.A. (1994). Remembering can cause forgetting: Retrieval dynamics in long-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*, 1063–1087.

Anderson, M.C., & Levy, B.J. (2007). Theoretical issues in inhibition: Insights from research on human memory. In D.S. Gorfrein & C.M. MacLeod (Eds.), *Inhibition in cognition* (pp. 81–103). Washington, DC: American Psychological Association.

Anderson, M.C., & Spellman, B.A. (1995). On the status of inhibitory mechanisms in cognition: Memory retrieval as a model case. *Psychological Review*, *102*, 68–100.

Bjork, E.L., Bjork, R.A., & Anderson, M.C. (1998). Varieties of goal-directed forgetting. In J.M. Golding & C.M. MacLeod (Eds.), *Intentional forgetting: Interdisciplinary approaches* (pp. 103–139). Mahwah, NJ: Erlbaum.

Camp, G., Pecher, D., & Schmidt, H.G. (2007). No retrieval-induced forgetting using item-specific independent cues: Evidence against a general inhibitory account. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*, 950–958.

Clow, A., Thorn, L., Evans, P., & Hucklebridge, F. (2004). The awakening cortisol response: Methodological issues and significance. *Stress*, *7*, 29–37.

Conway, M.A. (2009). Episodic memories. *Neuropsychologia*, *47*, 2305–2313.

Craik, F.I.M., & Lockhart, R.S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, *11*, 671–684.

Drosopoulos, S., Wagner, U., & Born, J. (2005). Sleep enhances explicit recollection in recognition memory. *Learning & Memory*, *12*, 44–51.

Dumay, N., & Gaskell, M.G. (2007). Sleep-associated changes in the mental representation of spoken words. *Psychological Science*, *18*, 35–39.

Euston, D.R., Tatsuno, M., & McNaughton, B.L. (2007). Fast-forward playback of recent memory sequences in prefrontal cortex during sleep. *Science*, *318*, 1147–1150.

Füredi, M., & Kelemen, J. (1989). *A mai magyar nyelv szépprózai gyakorisági szótára* [A frequency dictionary of the literary language of Hungarian]. Budapest, Hungary: Akadémiai Kiadó.

Kónya, A., & Pintér, G. (1985). Kategória norma a verbális emlékezet vizsgálatához [Category norms for verbal memory research]. *Hungarian Psychological Review*, *2*, 93–111.

MacLeod, M.D., & Macrae, C.N. (2001). Gone but not forgotten: The transient nature of retrieval-induced forgetting. *Psychological Science*, *12*, 148–152.

Mensink, G.J.M., & Raaijmakers, J.W. (1988). A model of interference and forgetting. *Psychological Review*, *95*, 434–455.

Norman, K.A., Newman, E.L., & Detre, G. (2007). A neural network model of retrieval-induced forgetting. *Psychological Review*, *114*, 887–953.

Racsmány, M., & Conway, M.A. (2006). Episodic inhibition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*, 44–57.

Racsmány, M., Conway, M.A., Garab, E.A., & Nagymáté, G. (2008). Memory awareness following episodic inhibition. *Quarterly Journal of Experimental Psychology*, *61*, 525–534.

Saunders, J., & MacLeod, M.D. (2002). New evidence on the suggestibility of memory: The role of retrieval-induced forgetting in misinformation effects. *Journal of Experimental Psychology: Applied*, *8*, 127–142.

Stickgold, R., & Walker, M.P. (2005). Memory consolidation and reconsolidation: What is the role of sleep? *Trends in Neurosciences*, *28*, 408–415.

Storm, B.C., Bjork, E.L., Bjork, R.A., & Nestojko, J.F. (2006). Is retrieval success a necessary condition for retrieval-induced forgetting? *Psychonomic Bulletin & Review*, *13*, 1023–1027.

Walker, M.P., Brakefield, T., Morgan, A., Hobson, J.A., & Stickgold, R. (2002). Practice with sleep makes perfect: Sleep-dependent motor skill learning. *Neuron*, *35*, 205–211.

Williams, H.L., Conway, M.A., & Baddeley, A.D. (2008). The boundaries of episodic memories. In T.F. Shipley & J.M. Zacks (Eds.), *Understanding events: How humans see, represent, and act on events* (pp. 589–617). New York: Oxford University Press.

Wilson, M.A., & McNaughton, B.L. (1994). Reactivation of hippocampal ensemble memories during sleep. *Science*, *265*, 676–679.

VII. The role of episodic cues in retrieval-related inhibition and facilitation

VII.1. The alterations of episodic cue-item links following intentional inhibition of retrieval processes (Study 9)

Paper: Inhibition and interference in the think/no-think task. Mihály Racsmány, Martin A Conway, Attila Keresztes, Attila Krajcsi, *Memory & Cognition* 02/2012; 40(2):168-76. DOI:10.3758/s13421-011-0144-6

Mem Cogn (2012) 40:168–176
DOI 10.3758/s13421-011-0144-6

Inhibition and interference in the think/no-think task

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

Published online: 11 October 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

M. Racsmány (✉) · A. Keresztes
Department of Cognitive Science,
Budapest University of Technology and Economics,
Egry J. u. 1.,
Budapest 1111, Hungary
e-mail: racsmany@cogsci.bme.hu

M. Racsmány
Institute of Psychology, University of Szeged,
Szeged, Hungary

M. A. Conway
The Leeds Memory Group, Institute of Psychological Sciences,
University of Leeds,
Leeds, UK

A. Krajcsi
Institute of Psychology, Eötvös Loránd University,
Budapest, Hungary

Remembering 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, Wetzels, 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 target words 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. Only think and no-think words were used in this task; that is, 160 trials were performed altogether. Finally, participants took part in a cued recall test in which the cues of think, no-think, and baseline words were presented and participants were asked to recall the targets to each cue word. The procedure of this phase was identical to that of the test–feedback cycles of the initial learning phase. In sum, the task was to recall B items of the word pair cued A items.

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.

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.

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

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

*Standard deviations are shown in parentheses.

In Experiment 1, the association A→ 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→ 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→ 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← B is primed by the no-think trials but that the relation A→ 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 ($\varepsilon = .71$). Item type had a significant effect on recall performance, F(1.57, 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.

dc_1204_16

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.[2] 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 × 3 × 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 instruction × 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-

[2] Another way in which to obviate the problem of changing baselines over experiments would be to conduct a fully within-subjects design. But such a design requires considerably more items, leading to possible floor effects in some conditions and contamination between conditions at test. As a preliminary pilot study, we ran a within-subjects replication of Experiment 1 and 2—that is, using the suppress instruction only and involving cue type as a within-subjects variable. We ran two groups and tested with blocks of cues. One group had B cues first for half the items, followed by A cues for the other half of items, and the other group had the reverse of this order. The learning and practice phases were the same as in the present experiments. In this experiment, the group tested with A cues first corresponded to that in the standard TNT experiment (Experiment 1 in the present series), and the pattern of findings was as expected: think, .82; no-think, .62; and baseline, .70. For the group that then received B cues to recall A items, the pattern was as follows: think, .85; no-think, .70; and baseline, .65—also in line with the pattern of recall seen in Experiments 2 and 5 above. These differences, although consistent with the earlier patterns of means, did not reach statistical significance. Removing participants on the basis of the Bulevich et al. (2006) questionnaire and 2 participants with outlying scores, a reliable 2 (no-think vs. baseline) ×2 (A vs. B cuing) interaction is present. This interaction is due to a below-baseline performance for no-think B items when cued with A items (the TNT effect) and above-baseline performance for A items when cued with no-think B items—effects consistent with the findings of Experiments 1 and 2/5. Supportive though these findings are, we believe that this within-subjects type of design needs considerable development in order to find a procedure that does not increase individual differences and has the appropriate power.
In the presentation order condition in which participants were first tested with a block of B cues and subsequently with a block of A cues, the means were the following: B cues, think .85, no-think .69, and baseline .56; A cues, think .65, no-think .73, baseline .53. Interestingly, planned contrasts to baseline showed both think and no-think (irrespective of cue) to be reliably higher than baseline, $F(1, 10) = 6.60$ and $F(1, 10) = 7.64$, $p < .05$ in both cases. The pattern for B cues is consistent with Experiments 2/5. The pattern for A cues is, however, anomalous and is not consistent with any of the findings above–in particular, with those of Experiment 1 and the TNT effect. We tentatively suggest that this may reflect some type of carryover from having successfully used B no-think items in the first block of the test to recall A items. What form this might take will require further research. Clearly, a within-subjects and/or mixed approach is of value, and the preliminary findings reported here (full details are available from the first author) are broadly consistent with the findings using between-subjects designs (Experiments 1–5).

ipants 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)[3] 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).[4]

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.[5] 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-*

[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.

[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.

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.

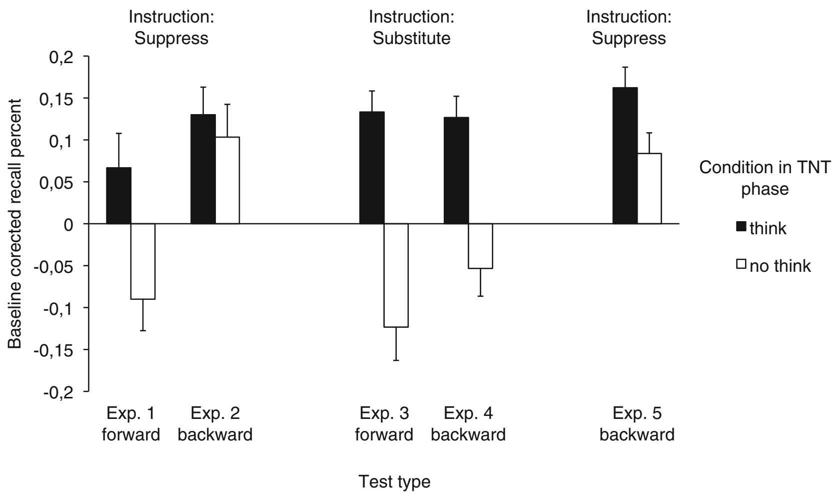
One further feature of the model is that because inhibition is assumed to be directed at associations between representations of A and B terms, it is possible for representations of the terms themselves to remain above some resting level of activation, as can other associations between them not targeted 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

Fig. 1 Baseline corrected recall performances for think and no-think items in 5 experiments using TNT



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 4, as compared with 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 Professorial Fellowship, RES-051-27-0127, from the Economic and Social Research Council of Great Britain.

References

Algarabel, S., Luciano, J. V., & Martínez, J. L. (2006). Inhibitory voluntary control of memory: Effect of stimulus onset asynchrony on reaction time to suppressed memories. *Psicológica, 27,* 35–56.

Anderson, M. C., & Green, C. (2001). Suppressing unwanted memories by executive control. *Nature, 410,* 366–369.

Anderson, M. C., Ochsner, K. N., Kuhl, B., Cooper, J., Robertson, E., Gabrieli, S. W., et al. (2004). Neural systems underlying the suppression of unwanted memories. *Science, 303,* 232–235.

Asch, S. E., & Ebenholtz, S. M. (1962). The principle of associative symmetry. *Proceedings of the American Philosophical Society, 106*, 135–163.

Baddeley, A. D. (1976). *The psychology of human memory.* New York: Basic Books.

Bergström, Z. M., de Fockert, J. W., & Richardson-Klavehn, A. (2009). ERP and behavioural evidence for direct suppression of unwanted memories. *NeuroImage, 48,* 726–737.

Bergström, Z. M., Velmans, M., de Fockert, J., & Richardson-Klavehn, A. (2007). ERP evidence for successful voluntary avoidance of conscious recollection. *Brain Research, 1151,* 119–133.

Bulevich, J. B., Roediger, H. L., III, Balota, D. A., & Butler, A. C. (2006). Failures to find suppression of episodic memories in the think/no-think paradigm. *Memory & Cognition, 34,* 1569–1577.

Carpenter, S. K., Pashler, H., & Vul, E. (2006). What types of learning are enhanced by a cued recall test? *Psychonomic Bulletin & Review, 13*, 826–830.

Conway, M. A. (2009). Episodic memories. *Neuropsychologia, 47,* 2305–2313.

Crowder, R. G. (1976). *Principles of learning and memory.* Hillsdale, NJ: Erlbaum.

Depue, B. E., Banich, M. T., & Curran, T. (2006). Suppression of emotional and nonemotional content in memory: Effects of repetition on cognitive control. *Psychological Science, 17,* 441–447.

Depue, B. E., Curran, T., & Banich, M. T. (2007). Prefrontal regions orchestrate suppression of emotional memories via a two-phase process. *Science, 317,* 215–219.

Grison, S., Tipper, S. P., & Hewitt, O. (2005). Long-term negative priming: Support for retrieval of prior attentional processes. *Quarterly Journal of Experimental Psychology, 58A,* 1199–1224.

Halácsy P., Kornai A., Németh L., Rung A., Szakadát I., & Trón V. (2004). Creating open language resources for Hungarian. *Proceedings of the 4th international conference on Language Resources and Evaluation, LREC2004*. European Language Resources Association.

Hanslmayr, S., Leipold, P., & Bauml, K. H. (2010). Anticipation boosts forgetting of voluntarily suppressed memories. *Memory, 18,* 252–257.

Hertel, P. T., & Calcaterra, G. (2005). Intentional forgetting benefits from thought substitution. *Psychonomic Bulletin & Review, 12,* 484–489.

Hertel, P. T., & Gerstle, M. (2003). Depressive deficits in forgetting. *Psychological Science*, 573–578.

Kahana, M. J., Howard, M. W., & Polyn, S. M. (2008). Associative retrieval processes in episodic memory. In J. Byrne (Series Ed.) & H. L. Roediger III (Vol. Ed.), *Cognitive psychology of memory*: Vol *2*. Learning and memory: A comprehensive reference (pp. 467–490). Oxford: Elsevier

Levy, B. J., & Anderson, M. C. (2008). Individual differences in the suppression of unwanted memories: The executive deficit hypothesis. *Acta Psychologica, 127,* 623–635.

Mecklinger, A., Parra, M., & Waldhauser, G. T. (2009). ERP correlates of intentional forgetting. *Brain Research, 1255,* 132–147.

Murdock, B. B. (1974). *Human memory: Theory and data*. Potomac, MD: Erlbaum.

Racsmány, M., & Conway, M. A. (2006). Episodic inhibition. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 32,* 44–57.

Thomson, D. M., & Tulving, E. (1970). Associative encoding and retrieval: Weak and strong cues. *Journal of Experimental Psychology, 86,* 255–262.

Tulving, E., & Osler, S. (1968). Effectiveness of retrieval cues in memory for words. *Journal of Experimental Psychology, 77,* 593–601.

Wegner, D. M. (1994). Ironic processes of mental control. *Psychological Review, 101,* 34–52.

Wessel, I., Wetzels, S., Jelicic, M., & Merckelbach, H. (2005). Dissociation and memory suppression: A comparison of high and low dissociative individuals' performance on the think–no think task. *Personality and Individual Differences, 39,* 1461–1470.

VII.2. Retrieval practice and long-term automatization of cued-recall (Study 10)

Paper: Testing promotes long-term learning via stabilizing activation patterns in a large network of brain areas. Attila Keresztes, Daniel Kaiser, Gyula Kovács, Mihály Racsmány, *Cerebral Cortex* 10/2014; 24:3025-3035. DOI:10.1093/cercor/bht158

Cerebral Cortex Advance Access published June 24, 2013

Cerebral Cortex
doi:10.1093/cercor/bht158

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

Attila Keresztes[1], Daniel Kaiser[2], Gyula Kovács[1,2,3,4] and Mihály Racsmány[1]

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

Address correspondence to M. Racsmány, Budapest University of Technology and Economics, Egry 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 (retest 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 could 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 precuneus (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 reconstructing all associated memories. According 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 3 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 retest 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, seated in front 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 pairs for the later test phase. They were also told that they would 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 retest, 1 restudy, and 1 feedback block. Unknown to the participants, half of the word pairs were assigned to the retest strategy condition and half to the restudy strategy condition. The retest–restudy words varied randomly across participants. In the retest blocks, all 30 word pairs assigned to the retest 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 (Roediger and Butler 2011). In each learning cycle, the order of the retest 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 retest 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 only 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 × 64 matrices; in-plane resolution: 3 × 3 mm; slice thickness: 3 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 × 2 × 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 × 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 ($P_{\text{uncorrected}} < 10^{-6}$; $T = 4.86$, 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; $P_{\text{uncorrected}} < 10^{-3}$; $T = 3.12$, 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

Table 1
Average MNI coordinates of regions of interest identified during the localizer run (2-back task)

	Mean cluster center		
	x	*y*	*z*
DLPFC, left (~BA9/45)	−47 (1)	28 (2)	32 (1)
DLPFC, right (~BA9/45)	42 (2)	37 (2)	32 (2)
Anterior prefrontal, left (~BA10)	−36 (1)	55 (1)	14 (2)
Anterior prefrontal, right (~BA10)	37 (1)	56 (1)	14 (2)
Posterior/dorsal prefrontal, left (~BA6)	−28 (1)	3 (1)	58 (1)
Posterior/dorsal prefrontal, right (~BA6)	34 (1)	5 (1)	60 (1)
Anterior cingulate, left (~BA32)	−11(1)	28 (2)	31 (1)
Anterior cingulate, right (~BA9/32)	8 (1)	34 (1)	31 (1)
Middle orbitofrontal, right (~BA11)	30 (1)	54 (2)	−11 (1)
Inferior parietal, left (~BA7)	−35 (1)	−54 (2)	48 (2)
Inferior parietal, right (~BA40)	43 (1)	−46 (1)	49 (1)
Superior parietal, left (~BA7)	−20 (1)	−71 (1)	55 (1)
Superior parietal, right (~BA7)	28 (1)	−66 (1)	57 (1)
Insula, left	−36 (1)	23 (1)	3 (1)
Insula, right	36 (1)	23 (1)	−1 (1)
Midbrain	2 (1)	−24 (1)	−11 (1)
Thalamus, left	−9 (1)	−13 (1)	11 (1)
Thalamus, right	12 (1)	−11 (1)	10 (1)
Fusiform gyrus, left	−42 (1)	−59 (3)	−17 (1)
Fusiform gyrus, right	39 (2)	−60 (3)	−17 (1)

Note: Standard errors are shown in brackets.

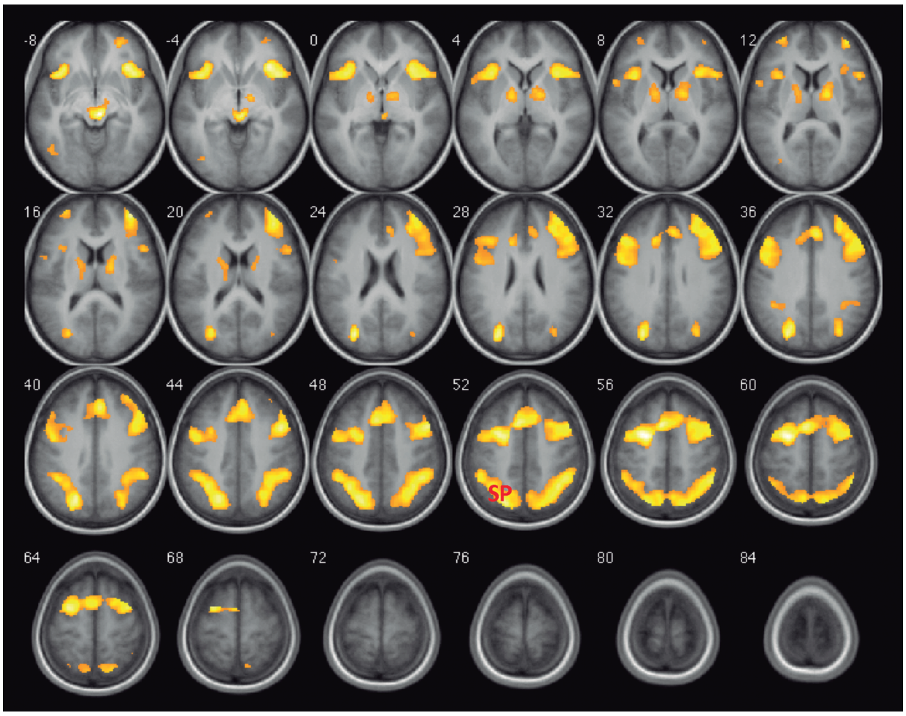


Figure 1. Activation maps for the functional localizer task (2-back vs. detection). Regions consistently activated across subjects are color-coded according to $P_{\text{FDR}} < 0.001$. The *z*-coordinate in Talairach space is indicated above each section. For anatomical details of the activations, see Table 1.

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 $Ps > 0.33$).

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 ($t_{(12)} = 2.33$, $P = 0.038$). However, there were no differences in the short retention interval condition, ($t_{(12)} = 0.92$, ns, $M = 67.44\%$, SE = 4.55% and $M = 70.77\%$, SE = 4.66% for retest and restudy, respectively). This result confirms previous findings in which repeated retrieval lead 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 = 182 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.

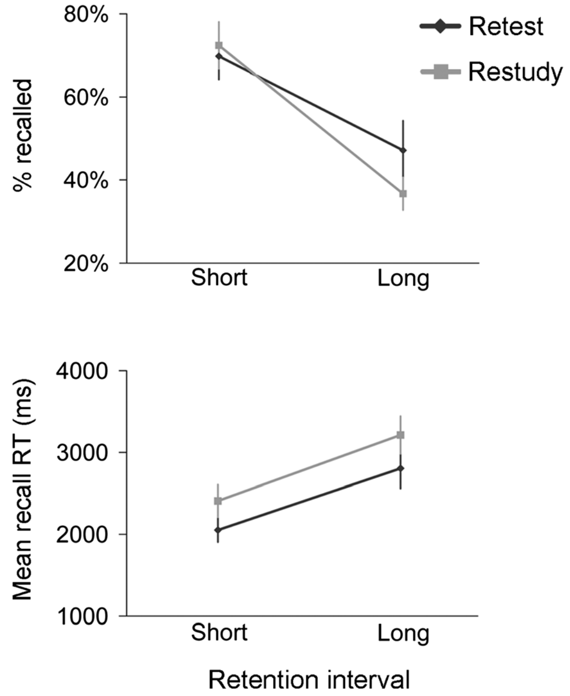


Figure 2. Accuracy (upper panel) and RTs (lower panel) of recall during the final test phase performed in the scanner as a function of retention interval and learning strategy. Error bars represent standard error of the mean.

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 $Ps > 0.33$), and only 1 region, the right thalamus showed a significant increase ($P < 0.026$, all other $Ps > 0.18$). 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

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

Effect	Area	*F*	df	*P*
Main effect of success	Inferior parietal, right (~BA40)	16.82	1.22	0.0001
	Thalamus, right	34.41	1.10	0.0001
	DLPFC, right (~BA9/45)	10.53	1.22	0.004
	Posterior/dorsal prefrontal, right (~BA6)	9.52	1.22	0.005
	Inferior parietal, left (~BA7)	8.05	1.23	0.009
	Anterior cingulate, left (~BA32)	8.57	1.20	0.008
	Thalamus, left	10.21	1.10	0.01
	Midbrain	7.95	1.9	0.02
	Anterior prefrontal, right (~BA10)	5.55	1.21	0.028
	Fusiform, left	5.5	1.16	0.032
	Insula, left	4.69	1.24	0.041
	Superior parietal, left (~BA7)	3.97	1.24	0.058
	Superior parietal, right (~BA7)	4.23	1.23	0.051
Main effect of retention interval	Middle orbitofrontal, right (~BA11)	6.3	1.22	0.02
Interaction: success × strategy	DLPFC, right (~BA9/45)	5.24	1.22	0.032
	Anterior prefrontal, right (~BA10)	5.28	1.21	0.032
Interaction: strategy × retention interval	Insula, left (~BA7)	19.86	1.24	0.0001
	Fusiform, left	13.65	1.16	0.002
	DLPFC, right (~BA9/45)	9.39	1.22	0.006
	Anterior cingulate, right (~BA9/32)	7.34	1.19	0.014
	Thalamus, right	7.63	1.10	0.02
	Insula, right	6.05	1.22	0.022
	Inferior parietal, left (~BA7)	5.62	1.23	0.026
	Superior parietal, right (~BA7)	5.14	1.23	0.033
	DLPFC, left (~BA9/45)	4.85	1.23	0.038
	Inferior parietal, right (~BA40)	4.32	1.22	0.05
	Middle orbitofrontal, right (~BA11)	4.21	1.22	0.052
	Posterior/dorsal prefrontal, right (~BA6)	4.12	1.22	0.055
Interaction: success × strategy × retention interval	Insula, left	9.14	1.24	0.006
	Anterior cingulate, left (~BA32)	9.08	1.20	0.007
	Thalamus, right	9.84	1.10	0.011
	Fusiform, left	7.77	1.16	0.013
	Fusiform, right	7.29	1.12	0.019
	Inferior parietal, right (~BA40)	5.35	1.22	0.03
	Anterior prefrontal, right (~BA10)	4.73	1.21	0.041
	Anterior prefrontal, left (~BA10)	4.27	1.23	0.05

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.

anterior PFC, the right middle orbitofrontal, and the right superior parietal cortex.

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 4*A* shows examples of the HRF for 2 such areas, the left insula (upper panel) and the left ACC (lower panel), while Figure 4*B* 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 in these areas, irrespective of final recall success. At long retention intervals, however, the result is different: previously restudied items elicit smaller 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 DLPFC, 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 retested 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.

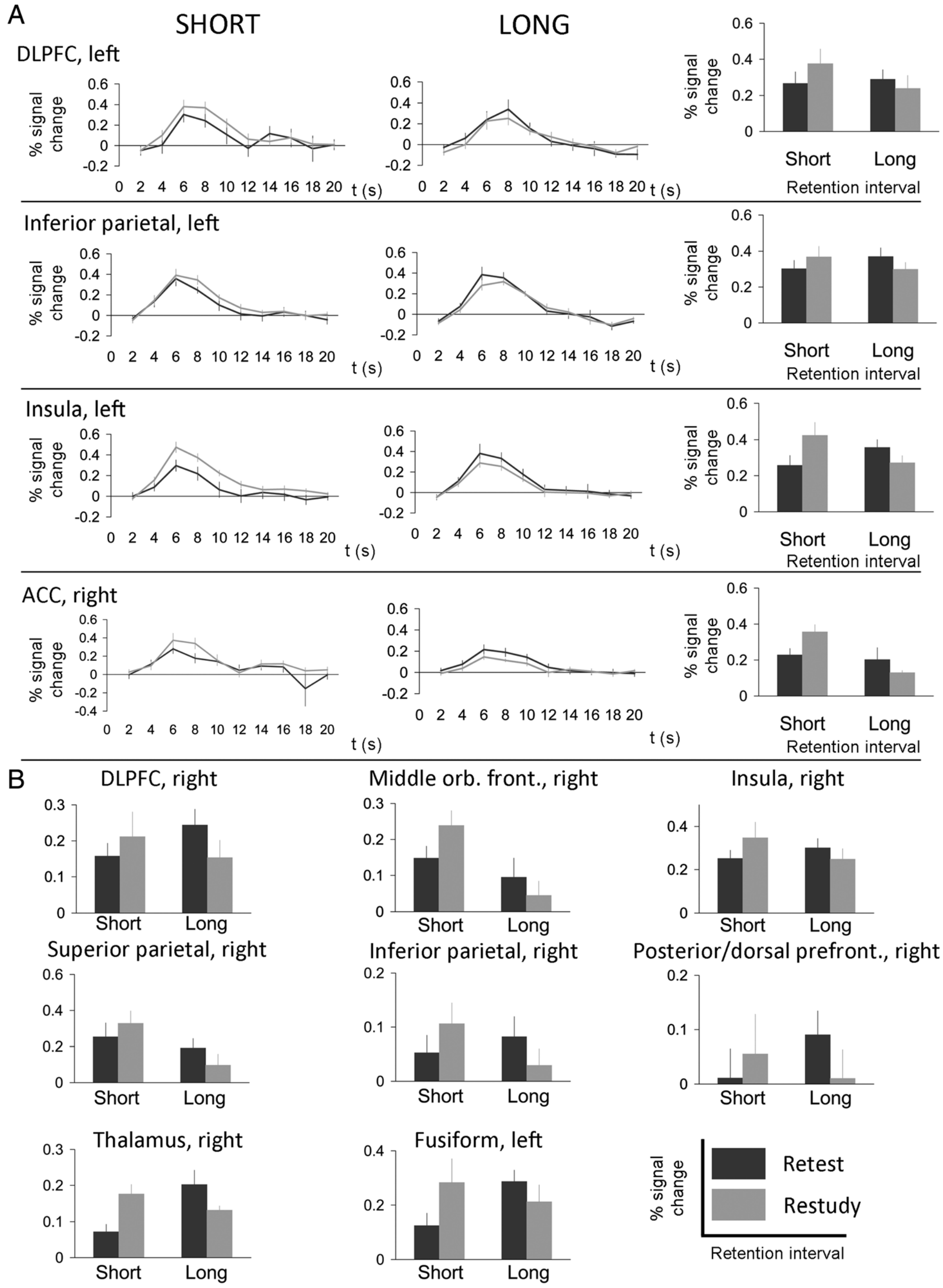


Figure 3. Activations during final recall in regions of interest where significant 2-way interactions between learning strategy (retest vs. restudy) × retention interval (short vs. long) were observed. (*A*) HRF functions (left column) and peak percent signal changes (right column) are shown for 4 representative areas. (*B*) Peak percent signal changes in 8 additional areas showing interaction of strategy and retention interval (see Table 2). Retested and restudied word pairs are shown separately for the short retention interval and long retention interval groups. Error bars represent standard error of the mean.

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 $P_{\mathrm{FWE}} < 0.05$ 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 $P_{\mathrm{uncorrected}} < 0.0001$ 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 5*A*, 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 5*B* 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.

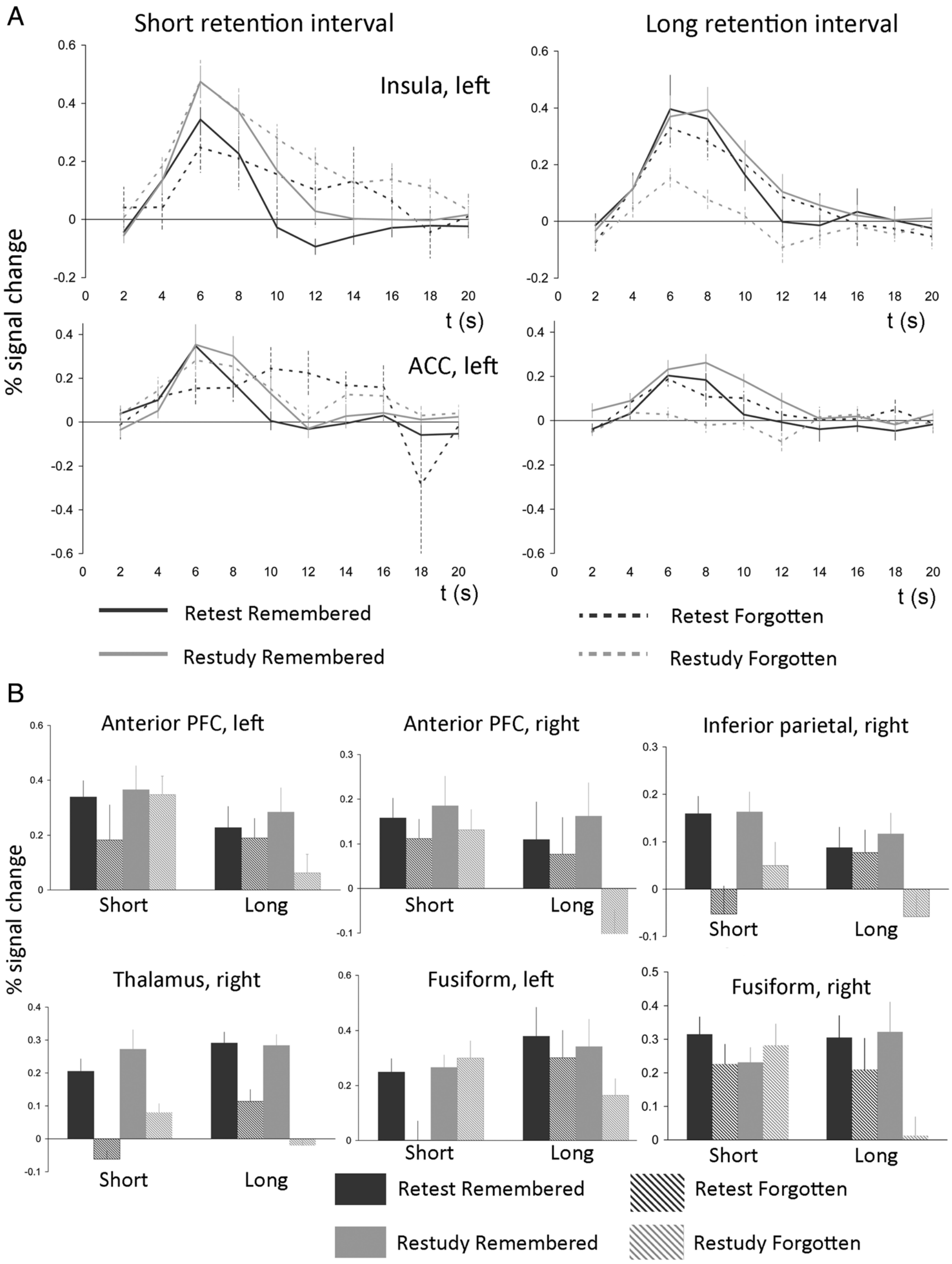


Figure 4. Activations during final recall in regions of interest where significant 3-way interactions (learning strategy × retention interval × final recall success) were observed. (*A*) HRF functions for the left insula (upper panel) and left ACC (lower panel). (*B*) Peak percent signal changes in 6 additional areas showing the 3-way interaction. Data are shown separately for the short retention interval and long retention interval groups. Error bars represent standard error of the mean.

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.

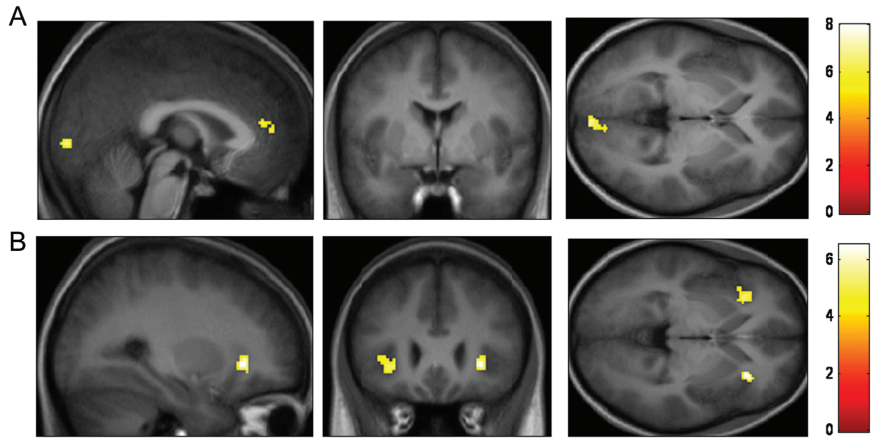


Figure 5. Results of the whole-brain analyses for the 2 contrasts with significant activations. Both panels show contrast images with activations (in white) significant at a threshold of $P < 0.0001$ (uncorrected) and an extended threshold of 50 voxels. (*A*) The effect of strategy at long retention interval (retest vs. restudy contrast) revealed significant activations in the medial frontal/anterior cingulate cortex and in an area in the occipital lobe at around the early visual cortex. (*B*) The interaction of strategy and retention interval ([(retest vs. restudy) long] vs. [(retest vs. restudy) short] contrast) revealed significant activations in the insula, bilaterally.

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 activations, associated with successfully recalled and forgotten words, were similar for both retested and restudied items. Similarly, after a week-long retention interval, these areas were activated for the previously retested 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 retested 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 retested 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 retest 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 retested 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 retested 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 retested items, that is, the emergence of the testing effect. In sum, we suggest that the average RT advantage of the retest 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 Science Foundation) (K84019 to M.R.), and by the New Hungary Development Plan (Project ID: TÁMOP-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 Márta Zimmer for her help in preparing the figures. *Conflict of Interest*: None declared.

References

Brett M, Anton JL, Valabregue R, Poline JB. 2002. Region of interest analysis using an SPM toolbox. Neuroimage. 16(2):372–373. In: Paper presented at 8th International Conference on Functional Mapping of the Human Brain, Sendai, Japan.

Buckner RL, Andrews-Hanna JR, Schacter DL. 2008. The Brain's default network: anatomy, function, and relevance to disease. Ann N Y Acad Sci. 1124:1–38.

Bunting M. 2006. The role of processing difficulty in the predictive utility of working memory span. Psychon Bull Rev. 13(6): 998–1004.

Cabeza R, Dolcos F, Graham R, Nyberg L. 2002. Similarities and differences in the neural correlates of episodic memory retrieval and working memory. Neuroimage. 16(2):317–330.

Cabeza R, Dolcos F, Prince SE, Rice HJ, Weissman DH, Nyberg L. 2003. Attention-related activity during episodic memory retrieval: a cross-function fMRI study. Neuropsychologia. 41(3):390–399.

Carrier M, Pashler H. 1992. The influence of retrieval on retention. Mem Cognit. 20(6):633–642.

Carpenter SK. 2009. Cue strength as a moderator of the testing effect: The benefits of elaborative retrieval. J Exp Psychol Learn Mem Cogn. 35:1563–1569.

Carpenter SK. 2011. Semantic information activated during retrieval contributes to later retention: Support for the mediator effectiveness hypothesis of the testing effect. J Exp Psychol Learn Mem Cogn. 37:1547–1552.

Cziraki C, Greenlee MW, Kovács G. 2010. Neural correlates of high-level adaptation-related aftereffects. J Neurophysiol. 103(3): 1410–1417.

Drobyshevsky A, Baumann SB, Schneider W. 2006. A rapid fMRI task battery for mapping of visual, motor, cognitive, and emotional function. Neuroimage. 31(2):732–744.

Ebbinghaus H. 1885/1964. Memory: a contribution to experimental psychology. New York (NY): Dover.

Eriksson J, Kalpouzos G, Nyberg L. 2011. Rewiring the brain with repeated retrieval: a parametric fMRI study of the testing effect. Neurosci Lett. 505(1):36–40.

Fletcher PC, Henson RNA. 2001. Frontal lobes and human memory. Brain. 124(5):849–881.

Fletcher PC, Shallice T, Frith CD, Frackowiak RS, Dolan RJ. 1998. The functional roles of prefrontal cortex in episodic memory. II. Retrieval. Brain. 121(7):1249–1256.

Gusnard DA, Raichle ME. 2001. Searching for a baseline: functional imaging and the resting human brain. Nat Rev Neuroscience. 2(10):685–694.

Henson RNA, Shallice T, Dolan RJ. 1999. Right prefrontal cortex and episodic memory retrieval: a functional MRI test of the monitoring hypothesis. Brain. 122(7):1367–1381.

Karpicke JD. 2012. Retrieval-based learning: active retrieval promotes meaningful learning. Curr Dir Psychol Sci. 21(3):157–163.

Karpicke JD, Blunt JR. 2011. Retrieval practice produces more learning than elaborative studying with concept mapping. Science. 331 (6018):772–775.

Karpicke JD, Roediger HL. 2008. The critical importance of retrieval for learning. Science. 319(5865):966–968.

Karpicke JD, Smith MA. 2012. Separate mnemonic effects of retrieval practice and elaborative encoding. J Mem Lang. 67:17–29.

Karpicke JD, Zaromb FM. 2010. Retrieval mode distinguishes the testing effect from the generation effect. J Mem Lang. 62:227–239.

Kim H. 2011. Neural activity that predicts subsequent memory and forgetting: a meta-analysis of 74 fMRI studies. Neuroimage. 54 (3):2446–2461.

Kovács G, Iffland L, Vidnyánszky Z, Greenlee MW. 2012. Stimulus repetition probability effects on repetition suppression are position invariant for faces. Neuroimage. 60(4):2128–2135.

Kovács G, Raabe M, Greenlee MW. 2008. Neural correlates of visually induced self-motion illusion in depth. Cereb Cortex. 18 (8):1779–1787.

Kuhl BA, Dudukovic NM, Kahn I, Wagner AD. 2007. Decreased demands on cognitive control reveal the neural processing benefits of forgetting. Nat Neurosci. 10(7):908–914.

Levy BJ, Kuhl BA, Wagner AD. 2010. The functional neuroimaging of forgetting. In: Della Sala S, editors. Forgetting (current issues in memory). New York (NY): Psychology Press. p. 135–164.

Mazoyer B, Zago L, Mellet E, Bricogne S, Etard O, Houdé O, Crivello F, Joliot M, Petit L, Tzourio-Mazoyer N. 2001. Cortical networks for working memory and executive functions sustain the conscious resting state in man. Brain Res Bull. 54:287–298.

Nelson TO, Dunlosky J. 1994. Norms of paired-associate recall during multitrial learning of Swahili-English translation equivalents. Memory. 2(3):325–335.

Ollinger JM, Corbetta M, Shulman GL. 2001. Separating processes within a trial in event-related functional MRI: II. analysis. Neuroimage. 13(1):218–229.

Owen AM, McMillan KM, Laird AR, Bullmore E. 2005. N-back working memory paradigm: a meta-analysis of normative functional neuroimaging studies. Hum Brain Mapp. 25(1):46–59.

Pyc MA, Rawson KA. 2010. Why testing improves memory: Mediator effectiveness hypothesis. Science. 330:335.

Ranganath C, Johnson MK, D'Esposito M. 2003. Prefrontal activity associated with working memory and episodic long-term memory. Neuropsychologia. 41(3):378–389.

Roediger HL, Butler AC. 2011. The critical role of retrieval practice in long-term retention. Trends Cogn Sci. 15(1):20–27.

Roediger HL, Karpicke JD. 2006b. The power of testing memory: basic research and implications for educational practice. Perspect Psychol Sci. 1(3):181–210.

Roediger HL, Karpicke JD. 2006a. Test-enhanced learning. Psychol Sci. 17(3):249–255.

Roediger HL, Weinstein Y, Agarwal PK. 2010. Forgetting: preliminary considerations. In: Della Sala S, editors. Forgetting (current issues in memory). New York (NY): Psychology Press. p. 1–22.

Rugg MD. 2004. Retrieval processing in human memory: electrophysiological and fMRI evidence. In: Gazzaniga MS, editors. The cognitive neurosciences. 3rd ed. Cambridge (MA): MIT Press. p. 727–738.

Rugg MD, Henson RNA. 2002. Episodic memory retrieval: an (event-related) functional neuroimaging perspective. In: Parker AE, Wilding EL, Bussey T, editors. The cognitive neuroscience of memory encoding and retrieval. Hove (UK): Psychology Press. p. 3–37.

Shulman GL, Fiez JA, Corbetta M, Buckner RL, Miezin FM, Raichle ME, Petersen SE. 1997. Common blood flow changes across visual tasks: II: decreases in cerebral cortex. J Cogn Neurosci. 9(5):648–963.
Spaniol J, Davidson PSR, Kim ASN, Han H, Moscovitch M, Grady CL. 2009. Event-related fMRI studies of episodic encoding and retrieval: meta-analyses using activation likelihood estimation. Neuropsychologia. 47(8–9):1765–1779.
Spitzer HF. 1939. Studies in retention. J Educ Psychol. 30(9):641–656.
Toppino TC, Cohen MS. 2009. The testing effect and the retention interval. Exp Psychol (formerly Z Exp Psychol). 56(4):252–257.
Tulving E. 1967. The effects of presentation and recall of material in free-recall learning. J Verb Learn Verb Behav. 6(2):175–184.
Uncapher MR, Wagner AD. 2009. Posterior parietal cortex and episodic encoding: insights from fMRI subsequent memory effects and dual-attention theory. Neurobiol Learn Mem. 91(2):139–154.
Unsworth N, Engle RW. 2007. Individual differences in working memory capacity and retrieval: a Cue-dependent search approach. In: Roediger HL, editors. The foundations of remembering: essays in Honor of Henry L. Roediger, III. New York (NY): Psychology Press. p. 241–258.
Unsworth N, Engle RW. 2006. A temporal–contextual retrieval account of complex span: an analysis of errors. J Mem Lang. 54(3):346–362.
Wheeler M, Ewers M, Buonanno J. 2003. Different rates of forgetting following study versus test trials. Memory. 11(6):571–580.
Wiklund-Hörnqvist C, Karlsson L, Eriksson J, Jonsson B, Nyberg L. 2012. The neural mechanisms underlying test-enhanced learning: an event-related fMRI study. Talk presented at the EARLI SIG 22: Neuroscience and Education, London, England.
Wimber M, Rutschmann RM, Greenlee MW, Bäuml KH. 2009. Retrieval from episodic memory: neural mechanisms of interference resolution. J Cogn Neurosci. 21(3):538–549.
Yarkoni T, Poldrack RA, Nichols TE, Van Essen DC, Wager TD. 2011. Large-scale automated synthesis of human functional neuroimaging data. Nat Methods. 8(8):665–670.

VIII. General Discussion

Hereinafter I shortly summarize the main findings and conclusions of the presented studies. The summary chapter of the dissertation organized around 10 thesis points.

VIII.1. Thesis (T) 1. The concept of episodic inhibition

The central idea of the presented experimental works is the concept of 'episodic inhibition'. Originally this theoretical account proposed that representation of episodic memories preserves a pattern of activation/inhibition encoded from the original experience or generated in it by subsequent retrieval acts (Racsmány and Conway, 2006). Thus, an item inhibited in episodic memory may nonetheless be activated in a conceptual knowledge structure. This proposal was supported by the findings of six experiments (Racsmány and Conway, 2006). Experiment 1 and 2 applied the list method of directed forgetting paradigm with some modifications. In both experiments all studied list items were involved in a lexical decision tasks as stimuli. An interesting aspect of this experimental design is its capability to separate suppression effects of directed forgetting instruction in episodic retrieval performance from other form of memory access, such as lexical and semantic access in lexical decision task. In line with episodic inhibition account, directed forgetting instruction produced suppression effect in episodic list-cued retrieval performance, whereas this effect was absent in lexical decision performance. Similarly, in Experiment 3-5 using the retrieval practice paradigm by Anderson and colleagues (1994) it was found that retrieval practice induced suppression on target-related items when access of these items were initiated by the studied (episodic) cues, and not when access of these items was guided by non-episodic, e.g. semantic cues. In Experiment 6 it was found that episodic retrieval was a central prerequisite for any suppression effect in the retrieval practice paradigm. If selective practice phase of the paradigm was replaced with a category generation task related to the same cues, no suppression effect of related items emerged. According to the episodic inhibition account, the diminished suppression in this procedure is due to the specificity of the design that participants could recall of generated items without accessing the episodic representation of the study phase. Altogether, these experiments underlie the concept of episodic inhibition which proposes that retrieval related suppression effects are guided by access of episodic representations. Participants may encode informational content of a given event in different representation forms, and as such, the same nominal item of a study event can show suppression when it is accessed by an episodic retrieval cue and activation when it is accessed by lexical or semantic cues.

VIII.2. T2. Differences between intentional and retrieval-induced forgetting

Beside the use of episodic cues, another important aspect of episodic retrieval is that it is accompanied by the experience of recollection (Tulving, 1985; Conway, 2005), whereas the access of conceptual system is accompanied by other state of memory awareness, such as the feeling of familiarity. According to the concept of episodic inhibition, retrieval-related suppression effects are mediated by the access of episodic representations, therefore it is assumed that episodic suppression effects will solely be present in retrieval performances accompanied by recollective consciousness. Using a list-method directed forgetting paradigm we found a powerful suppression effect in list-cued recall, moreover the effect was present in recognition performance accompanied by recollective performance, however suppression effect was absent in recognition performance accompanied by the feeling of familiarity (Racsmány et al., 2008). Experiment 2 and 3 of this study using retrieval practice paradigm demonstrated that retrieval-induced forgetting is present in voluntary cued recall, however there is no interaction between the selective practice effect and recollective judgement of the practiced items in a recognition task. These findings point to a conclusion that suppression effects in directed forgetting and retrieval practice procedures are due to partly different processes. It was suggested that following the directed forgetting instruction, episodic representation of the study phase marked as irrelevant and to-be-forgotten information. The main point is that this process is shaped by the goal of the participants to acquire only that information during the study phase that is relevant for the goal of the learning situation (i.e., to retrieve a high number of studied items at final recall test). This aspect of the intentional forgetting procedure was investigated in the third paper of the dissertation.

VIII.3. T3. Retrieval goal and forgetting

Concerning the goal-related relevance explanation of directed forgetting, we can assume when people observe another person with the same intention to learn, and see that this person is instructed to forget previously studied information, then they will produce the same intentional forgetting effect as the person they observed. This seems to be an important aspect of human learning: if we can understand the goal of an observed person and this is in line with our behavioural goals then our learning performance will mirror the learning performance of the model. In a series of directed forgetting experiments we investigated the relationship between the goal of the participants and the effect of forget instruction. However, in a standard list-method directed forgetting procedure it is difficult to independently manipulate the goal of the participants and the type of the instruction they receive, as they are strongly associated in this paradigm, so it was detailed in Racsmány et al. (2012) the goal of the participants and the instruction they receive is hardly dissociable in the standard DF procedure. A possible way to circumvent of this problem if participants are not directly instructed, but observe another person, who receive either a 'forget' or a 'remember' instruction. This was done in 4 experiments using a modified directed forgetting procedure. In Experiment 1 and 2 participants watched a movie of a directed forgetting experiments. They either receive a 'simple observation' or a 'goal sharing' instruction, the latter instruction make the participants to share the learning goal of the actor in the movie. In Experiment 3 and 4 followed the same logic than the previous two experiments with the exception that this time the observed model was a real participant in the experiment. The results were straightforward, participants produced directed forgetting effect only with the goal sharing instruction, and not with the simple observation instruction. In Experiment 3 and 4 the recall performance of the observer mirrored the performance of the model. Our results support the assumption that suppression of episodic memories is not automatically generated by environmental cues but depends on the goals of the person who encodes and retrieves them.

This assumption is further supported by still unpublished findings from a recent study of my laboratory. In this study we aimed to investigate the role of post-instruction encoding and pre-instruction relevant information in successful directed forgetting. Using a 2-list directed forgetting procedure and designated the second list as to-be-forgotten items (in Experiment 3) we were able to formulate different hypotheses based on the retrieval inhibition and context-change accounts. Retrieval inhibition theory assumes that forgetting of the to-be-forgotten items serves the adaptive goal of the participants to facilitate the learning of relevant information throughout the experiment (Bjork, 1989). Based on this assumption, we should not assume that the relevant information must follow the F-instruction. The F-instruction designates the to-be-forgotten items as irrelevant information, however, the participants need to learn some relevant information in order to being an adaptive behavior disregarding irrelevant information. From this point of view, it is not necessary to assume that the relevant information should follow the F-instruction, the only requirement for them is to be encoded before the final recall of all studied items. The reset-of-encoding hypothesis (Pastötter & Bäuml, 2010) would hypothesize that using two study lists and designated List 2 items as irrelevant by the F-instruction, there will be significant directed forgetting cost on List 2 items due to retrieval inhibition, whereas there will be no benefit on List 1 items. Considering, that according to the reset-of-encoding hypothesis, the benefit in the standard directed forgetting procedure is due to the improved encoding of the post-instruction items, improvement in recall of the relevant information when these items were encoded before the to-be-forgotten-items is not expected. The context-change account (Sahakyan & Delaney, 2003), however, would hypothesize different results, as this account assumed that the F-instruction changed the internal context of the participants and this internal context-change caused forgetting all pre-instruction information. Consequently, based on the context-change account it was hypothesized that both List 1 and List 2 would be forgotten if the F-instruction followed the study of List 2 items.

It is detected that the size of DF cost was sensitive for the specific form of the F-instruction, and for the stimuli presented in the two study lists (see MacLeod, 1998; see also Sahakyan & Delaney, 2003). Therefore, first, we wanted to find significant directed forgetting cost and benefit with the specific design and procedure that were applied in our laboratory. Then, using this procedure, we aimed to replicate the lack of DF cost of the F-instruction without consecutive second list learning. This could be important because the only published study in this regard was Pastötter and Bäuml (2007), as Gelfand and Bjork (1985) was a poster presentation. Then, in the third experiment using the same instructions and same stimuli then in Experiment 1 and 2, we aimed to find a significant cost of F-instruction without post-instruction encoding.

VIII.3.1. Experiment 1.

The aim of Experiment 1 was to detect significant cost and benefit of F-instruction in a list-method directed forgetting paradigm with two study lists.

Method

Participants

Altogether 60 (34 men and 26 women) Hungarian undergraduate students (native Hungarian speakers) participated in the first experiment between the ages of 18 and 28 years. They were recruited at the Budapest University of Technology and Economics, Hungary and received extra course credits for their participation. All participants gave written informed consent. In each of the three experiments, participants were randomly assigned into either a 'forget' or a 'remember' group. Based on the post-experiment debriefing, the data of three participants were excluded from the analysis, as they were informed of the goal of the experiment.

Stimuli and Design

Two lists (List A and List B) of Hungarian words with moderate recall probabilities were used as stimuli (12 items/list). Half of the participants in both groups (forget, remember) were first presented with the List A items, whereas the remaining participants were first presented with the List B words.

Procedure

All experiments (Experiment 1-3) were conducted under the same conditions: in the same lab with the same experimenter. Participants seated at 60 cm from a computer display. Stimuli appeared in random order on the computer screen (2000 msec/word) with a 1000-msec delay between them.

Participants were informed that they would see a list of words on the computer screen and were asked to memorize as many words, as they could. Immediately after they were given the List 1 items, participants in the forget group were given a between-list forget-instruction (F-instruction). They were told that the previously studied words were presented by a mistake and they were asked to forget the List 1 items. Before the presentation of the List 2 items, they were instructed to concentrate on the upcoming list and were asked to memorize as many List 2 words, as they could. Immediately after the presentation of the List 1 items, participants in the remember group were told that they had completed studying the first list and that they would receive a second list of words that had to be remembered as well (between-list remember-instruction [R-instruction]).

Immediately after studying the second list, participants in both experimental groups were exposed to an 8-minute arithmetic distractor task that was followed by a free recall test. Participants were first asked to recall the List 1 items, and then, the List 2 words. Participants were directed to recall first list items right away in order to eliminate output interference (see Anderson, 2005).

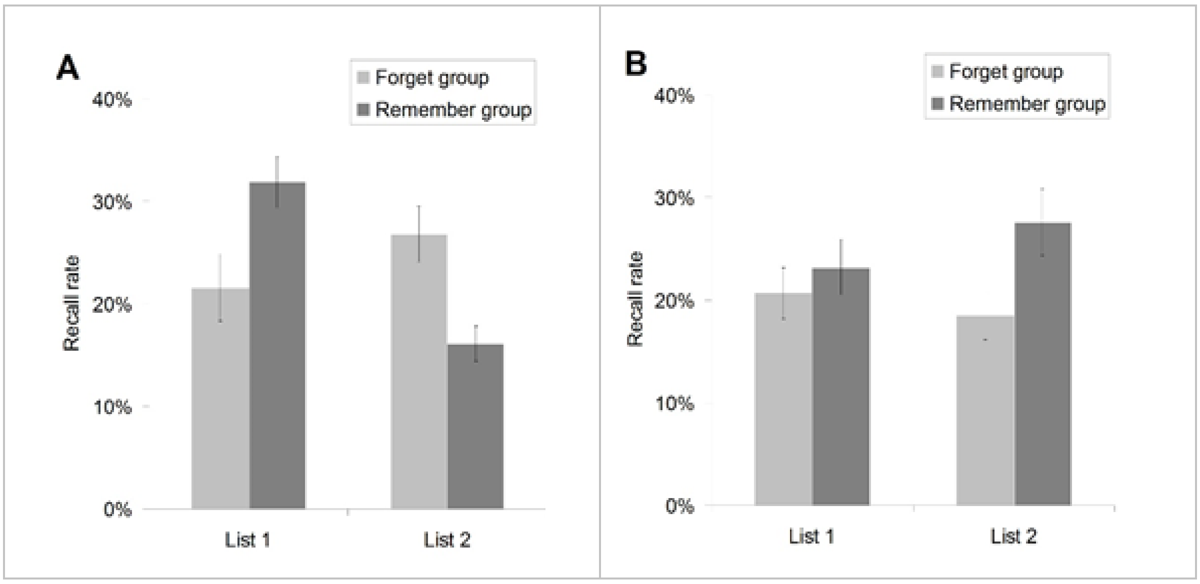
Results and Discussion

A mixed-design ANOVA was conducted on recall rates with Group (Forget/Remember) as a between-subjects variable and List (List 1/List 2) as a within-subjects factor. During post hoc analyses, independent-samples *t*-tests were used for the between-subjects factor.

The ANOVA revealed the predicted Group x List interaction ($F[1, 57] = 15.52$, $p < .001$, $\eta^2_p = .21$). Participants in the forget group recalled fewer List 1 words ($t[57] = -2.52$, $p < .05$, $r = -.32$) and more List 2 items ($t[57] = 3.31$, $p < .01$, $r = .40$) than participants in the remember group. In brief, a between-list F-instruction impaired memory for the List 1 items (the cost of the F-instruction) and improved memory for the List 2 words (the benefit of the F-instruction), in comparison with a baseline condition when a between-list R-instruction was given for the participants (see Figure 1A).

Significant cost and benefit of F-instruction were found in Experiment 1. This is an important starting point for Experiment 2 and 3, as based on the dominant theoretical account of the field, the lack of directed forgetting effects (cost and/or benefit) was hypothesized in the following experiments, and this experiment gave evidence that the lack of directed forgetting effect could not be due to the specific procedure or to any item features of the present study.

Figure 1. *Comparison of Recall Rates between the Forget and the Remember Groups in Experiment 1 and 3*



Notes. (A) Experiment 1: in a standard list-method directed forgetting task, a significant cost of a between-list F-instruction was seen for the first study list (fewer recalled List 1 words in the forget group than in the remember group) with a benefit of the F-instruction for the second study list (more recalled List 2 words in the forget group than in the remember group). (B) Experiment 3: when the F-instruction was given for the second study list after its presentation, a significant cost of the forget-instruction was found for the second list without a benefit of the F-instruction for the first list when participants in both groups were first asked to recall the List 2 items, and then, the List 1 words.

Error bars represent standard error of the mean (SEM).

VIII.3.2. Experiment 2.

The aim of Experiment 2 was to replicate the main results of Gelfand and Bjork (1985) and Pastötter and Bäuml (2007), as in these studies were found that there were no cost of F-instruction if the to-be-forgotten list was the only study list in the forget condition.

Method

Participants

Altogether 60 (30 men and 30 women) Hungarian undergraduate students (native Hungarian speakers) participated in the second experiment between the ages of 18 and 28 years. They were recruited at the Budapest University of Technology and Economics, Hungary and received extra course credits for their participation. All participants gave written informed consent. Based on the post-experiment debriefing, the data of three participants were excluded from the analysis, as they were informed of the goal of the experiment.

Procedure

We used the same procedure as in the first experiment until the point when participants were given a second study list in Experiment 1. In Experiment 2, after the presentation of the List 1 items, participants were given either an F- or an R-instruction, but there was no second list following the F-/R-instruction. Instead, participants in both groups were given the 8-minute arithmetic distractor task, and then, the free recall test.

Results and Discussion

In Experiment 2, recall rates were compared between the groups (forget vs. remember) by conducting an independent-samples *t*-test. When we compared the forget group's recall rate (M = 35.2%, SD = 15.7) to the remember group's performance (M = 38.7%, SD = 18.6), we did not find any reliable difference between them (t[58] = -0.77, n.s.). These results support that an F-instruction after the presentation of a list without a consecutive study list is not enough to the suppression of the to-be-forgotten items.

VIII.3. Experiment 3.

Experiment 2 replicated the results of Gelfand and Bjork (1985) and also of Pastötter and Bäuml (2007), as it was found that without a post-cue study list no cost of F-instruction emerged. Experiment 3 examined whether the F-instruction for List 2 items could be successful without consecutive learning, if there was another list (List 1) acquired right before the to-be forgotten list.

Method

Participants

Altogether 60 (38 men and 22 women) Hungarian undergraduate students (native Hungarian speakers) participated in the third experiment between the ages of 18 and 28 years. They were recruited at the Budapest University of Technology and Economics, Hungary and received extra course credits for their participation. All participants gave written informed consent. Based on the post-experiment debriefing, the data of two participants were excluded from the analysis, as they were informed of the goal of the experiment.

Procedure

In Experiment 3, we used the same procedure as in Experiment 1 with two modifications: (1) the forget-instruction was given for the second list after its presentation and not for the first list; (2) in the free recall phase, participants were first asked to recall the List 2 items, and then, the List 1 words. In other words, participants in both groups were presented with the List 1 items. Then, they were informed that a second list would be presented on the computer screen that had to be remembered as well. In the forget group, following the presentation of List 2, participants were told that the previously studied words were presented by a mistake and they were asked to forget the List 2 items. In the remember group immediately after the presentation of the List 2 items, participants were told that they had completed studying the second list that had to be remembered as well. Following the 8-minute arithmetic distractor task, participants in both groups were exposed to the test phase (the free recall of List 2 words, and then, the free recall of List 1 words). Note that in this Experiment forget instruction was given on List 2 items, therefore participants were directed to recall second list items right away in order to eliminate output interference (see Anderson, 2005).

Results and Discussion

In Experiment 3, participants in the forget group recalled fewer List 2 items than in the remember group ($t[58] = -2.28$, $p < .05$, $r = -.29$), but we did not find any group difference in the mean number of recalled List 1 items ($t[58] = -0.70$, n.s.). In sum, a significant cost of an F-instruction was seen for the recalled words from the second list without a benefit for the recalled List 1 words (see Figure 1B).

In Experiment 3, a significant cost of F-instruction was found without the encoding of a post-instruction study list. This result is in contrast with the result of Experiment 2, where no significant cost emerged without a post-instruction study list. Because the only difference between Experiment 2 and 3 is the presence of a pre-instruction study list (List 1) in Experiment 3, these results point to an interpretation that instead of post-instruction encoding of a new information, the presence of any relevant information during the entire experimental session is the critical factor in eliciting directed forgetting cost. A further result of Experiment 3 is that despite the successful suppression of List 2 items in the forget group, no benefit in recall of List 1 items emerged. This result is consistent with the reset-of-encoding hypothesis (Pastötter & Bäuml, 2010), because this theory suggests that the benefit of F-instruction is due to the improved encoding of the post-instruction items, therefore improvement in recall of to-be-remembered items when these items were encoded before the to-be-forgotten-items is not expected.

Altogether, here we presented evidence showing that post-instruction encoding of a new study list is not a necessary factor for eliciting directed forgetting cost. In three experiments it was found that with a procedure and experimental stimuli that was appropriate to find significant directed forgetting cost and benefit in a standard design (Experiment 1), there was no directed forgetting cost when the to-be-forgotten list was the only study list for the participants (Experiment 2), however, it was revealed a significant directed forgetting cost when a relevant to-be-remembered study list was presented for the participants before the F-instruction. To our knowledge this is the first directed forgetting study which found significant directed forgetting cost in free recall without post-instruction encoding of new information. The directed forgetting procedure is a model case of intentional learning, where a participant has to keep in an active form of relevant information while has to suppress irrelevant information. From the perspective of an adaptive cognitive system we can assume, that participants are able to produce an intentional suppression of successfully studied information by being informed which information is relevant and which is irrelevant from all the information they met during the entire experiment. In the present study, it was shown that participants used the F-instruction to suppress to-be-forgotten information without any post-instruction learning, if they were presented with relevant information to learn, the benefit of which made sensible to suppress irrelevant information.

VIII.4. T4. and T5. Executive system and retrieval-induced forgetting

A crucial issue regarding intentional forgetting and retrieval induced-forgetting whether executive control system is involved in interference resolution and memory suppression effects. Earlier we demonstrated (see T1 and T2) that intentional forgetting and retrieval-induced forgetting are different in a range of attributes of declarative memory, such as the involvement of the participants' goal or the conscious state accompanied retrieval processes. In further two studies we investigated memory functions of patients with executive disorders (patients were diagnosed with schizophrenia and obsessive-compulsive disorder) in order to reveal the involvement of executive processes in DF and RIF. In the first study it was found that patients diagnosed with schizophrenia and characterized with severe executive disorders produced no directed forgetting effect, whereas produced intact RIF (Racsmány et al., 2008). Our results indicate that possible disrupted executive functions may considerably weaken the ability of patients to intentionally avoid recent memories. This can occur even when other incidentally initiated inhibitory processes appear to function relatively normally. However, RIF may depend less on executive control, and even with serious symptoms of perseverative disorders, a symptom usually regarded as the consequence of inhibitory deficit, following repeated retrieval practice the suppression of target-related items remained intact. This result is in line with earlier findings of Conway and Fthenaki (2003), who found diminished DF and intact RIF in patients with frontal lobe injures.

In another study, using retrieval practice paradigm, there was demonstrated that retrieving memories did not induce forgetting of related memories among participants with OCD (Demeter et al., 2014). Lack of forgetting in OCD occurred in spite of the fact that overall memory and the mnemonic effect of practicing memories was almost identical to that among healthy controls. This result suggests that suppression of irrelevant, interfering memories during competitive recall is impaired in OCD. The lack of retrieval-induced forgetting (RIF) among OCD patients is not related to overall recall performance and working memory functions. A possible interpretation of this result that due to the lack of inhibitory deficit widely documented in OCD, the lack of RIF is another example of deficient inhibitory functions in this disorder (Chamberlain et al., 2005; Harsányi et al., 2014). However, we suggest that the lack of RIF in OCD is the consequence of interference insensitivity also documented in OCD, and as a consequence RIF is related to differences in resolving interference during competitive retrieval. Our findings suggest that the detected level of interference of target-related items is not high enough to kick in interference resolution processes and suppression of target-related memories in OCD. From another point of view, our results support the assumption that a certain amount of detected interference during retrieval is needed to make memory suppression an adaptive process. The relationship between the level of interference and the suppression of target-related memories in retrieval practice paradigm was further investigated in further studies with retrieval practice paradigm which is the subject of the next thesis point.

VIII.5. T6.-T8. The boundary conditions of retrieval-induced forgetting

In three studies with altogether seven experiments we investigated the prerequisite of suppression of target-related memory items in retrieval practice paradigm. Based on the findings of the first study it was proposed that RIF occurs only when interference during competitive retrieval reaches moderate levels, but not when it is too low or too high (Keresztes and Racsmány, 2013). This proposal indicates that low levels of interference do not trigger interference resolution, whereas interference resolution can fail when the interference reaches extremely high levels. Based on the findings of the second study, it was proposed that an initial retrieval of the learning set shields against the adverse effect of retrieval practice; RIF is absent either when measured by a comparison to baseline performance on the initial retrieval or to members of unpracticed categories (Racsmány and Keresztes, 2015). It is suggested that retrieval is the key process that enhances long-term accessibility of retrieved memories and it is the process that can hinder retrieval of items through search set restriction or can shield against the adverse effect of later selective retrieval. Based on the findings of the third study a revised form of the episodic inhibition account is proposed: retrieval practice establishes a pattern of activation and inhibition over the contents or features of an episodic memory of the study phase (Racsmány et al., 2010). As the episodic memory is consolidated in long-term memory, the pattern of activation and inhibition, which determines the accessibility of the contents of the memory, stabilizes and becomes resistant to further change. As a memory is repeatedly retrieved and its contents are accessed, its durability in long-term memory increases, and the accessibility levels of its contents become fixed (see also Szőllősi et al., 2016). Our findings suggest that sleep is important to this process of consolidation. It is proposed that consolidation processes occurring during sleep, and possibly featuring some form of offline rehearsal, mediate these long-term effects of retrieval practice.

VIII.6. T9.-T11. How to make skill from memory: an automatization account of retrieval practice effects

In two published studies and one still unpublished study with altogether 8 experiments we investigated the effect of repeated retrieval on cue-target item relationship. Episodic retrieval is controlled and channeled by cues, the role of cues in its simplest form can be investigated by using cued recall of paired associates (Baddeley, 1976; Crowder, 1976). In the first study the so-called Think/No-Think paradigm was applied in five experiments, in which participants learned a list of paired associates, then they were cued to repeatedly retrieve or repeatedly try to avoid the target elements. Based on our experimental findings we proposed that recalling two associated items can be simultaneously attenuated or primed depending on how the association is accessed (Racsmány et al., 2012). Furthermore, not thinking about a target item, as compared with thinking about an alternative, can produce the same decrements in cued recall or, sometimes, differences. Our findings suggest that the locus of suppression in the Think/No-think Task (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 (Racsmány et al., 2012). In other words, retrieval practice changes the strength of cue-item relationship, and the cue involved in this process represents a specific route to an episodic representation.

Interestingly, as it was introduced earlier, repeated retrieval practice reduces long-term forgetting and promotes better long-term retention than a restudy practice (i.e., repeated study), has been termed the "the testing effect" (Roediger and Buttler, 2011). Two influential theoretical accounts of this phenomenon exist. 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). In contrast, the search set constraining theoretical framework (Karpicke and Smith, 2012; Karpicke and Zaromb, 2010; Karpicke, 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 & Blunt, 2011; Karpicke & Zaromb, 2010; Karpicke, 2012). Based on the experimental findings of Racsmány and Keresztes (2015) and also of Racsmány et al. (2012), we assume that retrieval practice changes the cue-item association to an episodic link between episodic cues and target items, as it is suggested by the episodic inhibition account. We agree with Karpicke (2012) that 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, retrieval prompts a process which narrows the cue-related semantic network, and even a single retrieval can decrease the number of potentially retrievable items in response to a specific retrieval cue. However, to understand the effect of repeated retrieval practice on memory representations we turned not to the literature of episodic retrieval, but to the concept and experimental investigation of automaticity (Logan, 1988; Mours and De Hower, 2006). A complete theoretical and experimental review of the research on automaticity is beyond the scope of this dissertation, however, following Logan (1988), we assume that practice makes the development of automaticity as a transition from algorithm, or multistep memory retrieval to single-step memory retrieval. As it was detailed by Logan (1988) “Automaticity is memory retrieval: Performance is automatic when it is based on single-step direct-access retrieval of past solutions from memory. The theory assumes that novices begin with a general algorithm that is sufficient to perform the task. As they gain experience, they learn specific solutions to specific problems, which they retrieve when they encounter the same problems again.” (Logan, 1988., p.493). According to Logan (1988) we can quantitatively measure the level of automatization by the speed-up of the specific process, “the speed-up follows a regular function, characterized by substantial gains early in practice that diminish with further experience. More formally, the speed-up follows a power function…” (Logan, 1988, p.495). Based on this, we assume that retrieval and study practice fundamentally differ in how they involve the process of automatization and how they speed-up retrieval process. In a functional neuroimaging study applying retest vs. restudy paradigm, following repeated practice we measured the speed of retrieval along with retrieval success. There were two kinds of delay between practice and final retrieval, 20 minutes vs. one week. Retrieval practice produced faster retrieval reaction times than restudy practice both following 20 minutes and seven days delay. The shorter reaction time was accompanied with higher retrieval performance only at the longer delay, whereas restudy and retest practice produced equally high percentage of retrieval at 20 minutes delay. Although automatized behaviour has a series of diagnostic features, theorists of the field share the idea that practice changes task-related attention, awareness, control, speed and accuracy (see Moors and Houwer, 2006). The literature of attention and memory usually defined controlled behaviour as a goal-related process where one engage in, alter, stop or avoid the act when the goal and the effects are present (see Moors and Houwer, 2006). The involvement of executive control typically investigated with one of the following task, usually labelled as executive or working memory task: shifting tasks, fluency tasks, Stroop-like tasks, Go/NoGo and n-back tasks. Following this line in an fMRI experiment we aimed to analyse the role of control-related neuronal network in the testing effect (Keresztes et al., 2014). Based on the findings of this neuroimaging study it was proposed that the long-term behavioral advantage of repeated retrieval over repeated study is due to the differential activation of a large network. Specifically, when the retention interval is long, participants cannot effectively process the cue and a large percentage of retrieval attempts fail. Thus, the so-called 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. Considering the findings that repeated retrieval resulted in faster, more accurate retrieval performance with higher resistance to forgetting, which is related a decreased activation of the control-network short after the practice phase, and the activation pattern was stabile even after a seven days delay, we formulated an automatization account of the testing effect. This theory assume that the long-term memory advantage of repeated retrieval over repeated study is due to the automatization of cue-item retrieval association following repeated retrieval. Recent and still unpublished experimental findings from our laboratory underlie this concept showing a relationship between decreased retrieval reaction time following retrieval practice and long term memory performance. Here it is proposed that retrieval practice decreases the involvement of control processes in episodic retrieval, makes the mobilization of episodic cue-target relationship faster, more effective in long-term, thereby turning memory to skill.

VIII.6.1. Experiment 4.

Method

Participants

Twenty-nine Hungarian undergraduate students (native Hungarian speakers) participated in Experiment 4 (12 men and 17 women; age range: 19-27 years; M_{age} = 23.0 years, *SD* = 2.2). Subjects were recruited at different universities in Budapest. They received money for their participation. All participants gave written informed consent. The study was approved by the Ethical Committee of the Budapest University of Technology and Economics.

Materials and procedure

Participants were presented with a computer-controlled learning paradigm, while seated at approximately 70 cm from a computer display. The experiment was created in Matlab using Psychtoolbox 3.0. Stimuli were 40 neutral Swahili-Hungarian word pairs translated from Nelson and Dunlosky (1994). The memory task consisted of three main phases: an initial learning phase, a practice phase, and a final test phase.

In the initial learning phase, participants were presented with all word pairs in random order (5000 msec/word pair; inter-stimulus interval [ISI]: 500 msec), with the Swahili word on the left and its Hungarian equivalent on the right. Participants were instructed to memorize as many word pairs as possible.

Immediately after the initial learning phase, participants practiced the word pairs in six cycles (practice phase). Word pairs were randomly assigned into a restudy (20 word pairs) or a retest condition (20 word pairs). Each cycle began with a restudy or a retest block (the order of the restudy and retest blocks varied randomly across the learning cycles). Each restudy-retest block was followed by a feedback block. In the *restudy* blocks, participants saw 20 Swahili words together with their Hungarian meanings in random order (8000 msec/word pair; ISI: 500 msec). In the *retest* blocks, 20 Swahili words were presented in random order on the computer screen. Participants were instructed to press the space button on a standard keyboard of the computer when the right answer came to their mind. Participants were allowed to type the Hungarian meanings of the Swahili words only when they pressed the space button. They had a maximum of 8000 msec to complete one word pair – 8000 msec after the onset of the stimulus (Swahili word), the next stimulus was presented automatically even if the subject gave an answer and even if not. In the feedback blocks, all 40 word pairs were presented randomly for the participants (1500 msec/word pair; ISI: 500 msec).

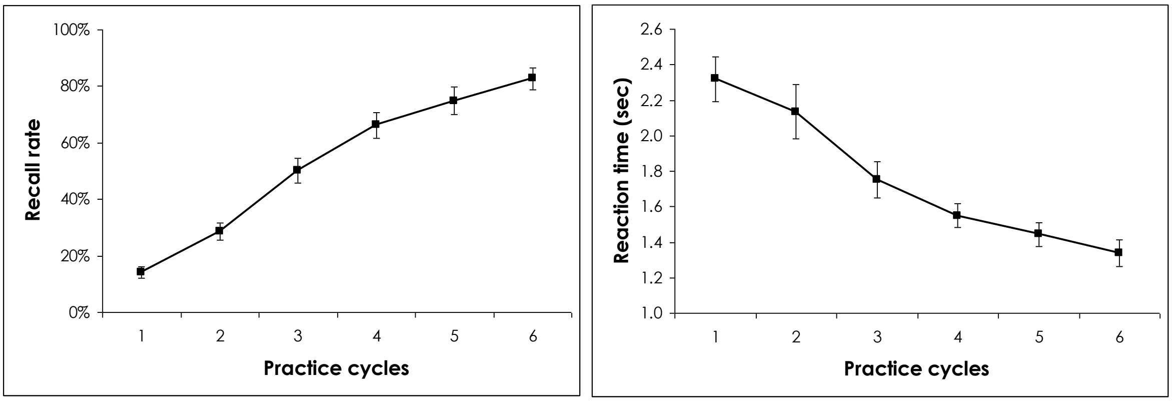
Following a 7-day retention interval, participants' memory for all 40 word pairs was tested in the final test phase. Swahili words were presented in random order. Similarly as in the practice phase, participants were instructed to press the space button when the right answer came to their mind. Participants were allowed to type the Hungarian meanings of the Swahili words only when they pressed the space button (8000 msec after the onset of the stimulus, the next stimulus was presented automatically).

Results

For each practice cycle, mean reaction times were calculated for the correct responses (i.e., time interval between stimulus onset and press of the space button). Recall rates and reaction times were compared between the practice cycles by conducting repeated measures ANOVAs with six levels, and then, by conducting simple contrasts with the last (sixth) practice cycle as a reference point. The ANOVA indicated a significant effect for the recall rate ($F[5,28] = 187.64$, $p < .001$, $\eta^2_p = .87$) and also for the reaction time ($F[5,25] = 27.75$, $p < .001$, $\eta^2_p = .53$), see Figure 2. The contrast analysis established that recall rate in the last practice cycle was higher than in all previous cycles (cycle 1: $F[1,28] = 425.08$, $p < .001$, $\eta^2_p = .94$; cycle 2: $F[1,28] = 308.94$, $p < .001$, $\eta^2_p = .92$; cycle 3: $F[1,28] = 169.70$, $p < .001$, $\eta^2_p = .86$; cycle 4: $F[1,28] = 82.23$, $p < .001$, $\eta^2_p = .75$; cycle 5: $F[1,28] = 27.23$, $p < .001$, $\eta^2_p = .49$). Furthermore, mean reaction time obtained in the last cycle was lower than reaction times in the first four practice cycles (cycle 1: $F[1,25] = 47.53$, $p < .001$, $\eta^2_p = .66$; cycle 2: $F[1,25] = 46.07$, $p < .001$, $\eta^2_p = .65$; cycle 3: $F[1,25] = 15.93$, $p < .001$, $\eta^2_p = .39$; cycle 4: $F[1,25] = 10.36$, $p < .01$, $\eta^2_p = .29$). Reaction times did not differ significantly between the last and the fifth practice cycles ($F[1,25] = 2.05$, $p = .16$, $\eta^2_p = .08$). In brief, whereas recall success increased during the practice cycles, reaction times decreased from cycle 1 to cycle 5.

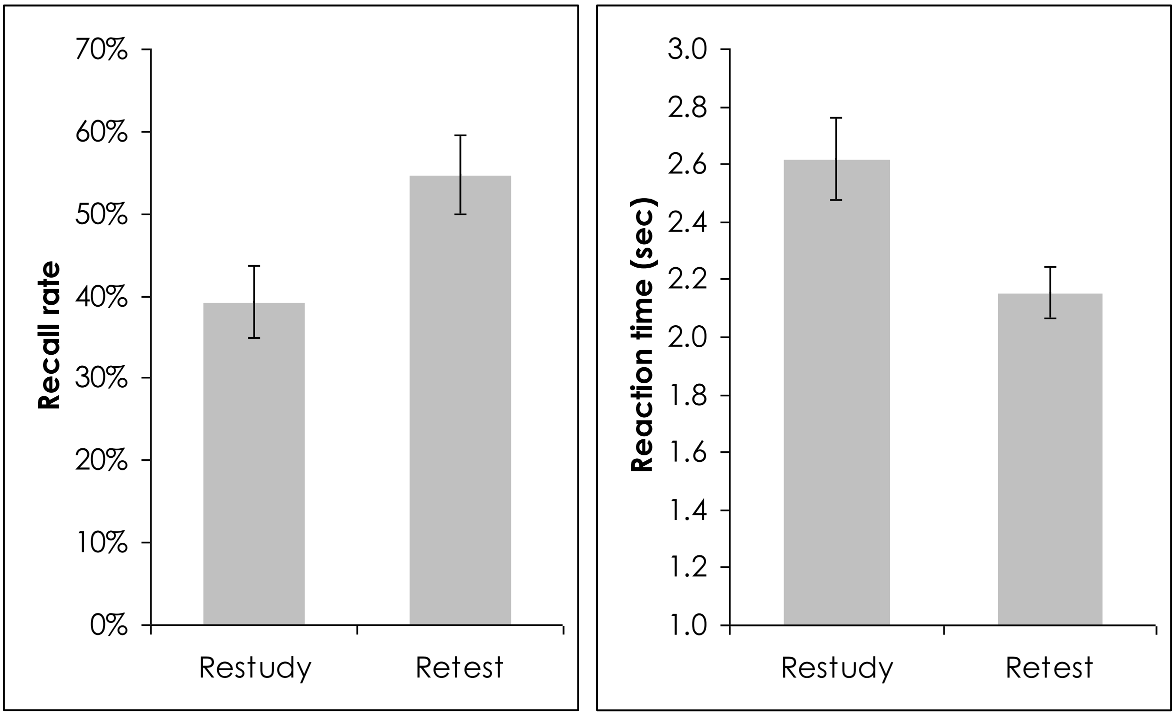
Recall rates and reaction times obtained in the final test phase were compared between the Restudy and the Retest conditions, see Figure 3. As expected, recall rate for the retested words was higher than for the restudied items ($t[28] = -4.85$, $p < .001$), whereas mean reaction time for the correct responses was lower in the Retest condition than in the Restudy condition ($t[28] = 2.78$, $p < .01$). Importantly, a significant negative correlation was found between the last practice cycle's reaction time and recall rate for the retested word pairs in the final test phase ($r[29] = -.40$, $p < .05$).

Figure 2. *Practice phase of Experiment 4: (A) recall rates and (B) reaction times.*



Note. Error bars represent standard error of the mean.

Figure 3. *Final test phase of Experiment 4: comparison of (A) recall rates and (B) reaction times between the Restudy and Retest conditions.*



Note. Error bars represent standard error of the mean.

VIII.6.2. Experiment 5.

Method

Participants

Twenty-four Hungarian undergraduate students (native Hungarian speakers) participated in Experiment 5 (11 men and 13 women; age range: 17-24 years; M_{age} = 22.3 years, *SD* = 6.2). Subjects were recruited at different universities in Budapest. They received money for their participation. All participants gave written informed consent. The study was approved by the Ethical Committee of the Budapest University of Technology and Economics.

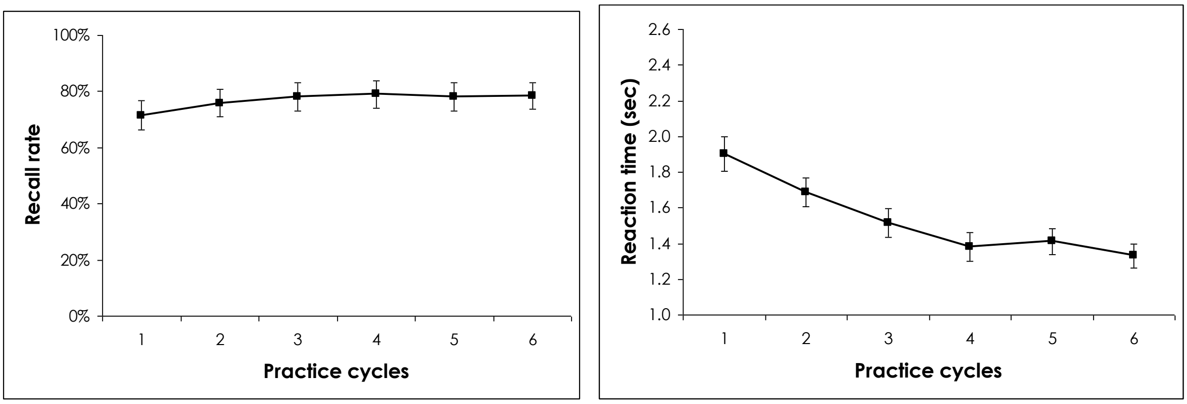
Materials and Procedure

A similar memory paradigm was used as in Experiment 4 with three modifications. First, in the initial learning phase of the experiment, participants were presented with the stimuli in 5 cycles. In one learning cycle, all 40 word pairs were presented randomly for the participants. There was no delay between the learning cycles. Second, there was a 5-min delay between the initial learning phase and the practice cycles. During the delay, participants were presented with simple arithmetic tasks. Finally, we used no feedback during the practice cycles; therefore, each practice cycle consisted of a restudy block (20 word pairs) and a retest block (20 word pairs). All other parameters remained the same as in Experiment 1.

Results

Recall rates and reaction times obtained in the practice phase were analyzed in a similar way as in Experiment 4. The ANOVAs revealed significant effects for recall rates ($F[5,23] = 11.02$, $p < .001$, $\eta^2_p = .32$) and reaction times ($F[5,23] = 28.40$, $p < .001$, $\eta^2_p = .55$), see Figure 3. Compared to the last practice cycle, recall rates were significantly lower in cycle 1 ($F[5,23] = 19.16$, $p < .001$, $\eta^2_p = .45$) and cycle 2 ($F[5,23] = 11.60$, $p < .001$, $\eta^2_p = .34$). Besides, mean reaction time obtained in the last practice cycle was significantly lower than in the first three cycles (cycle 1: $F[5,23] = 54.60$, $p < .001$, $\eta^2_p = .70$; cycle 2: $F[5,23] = 53.50$, $p < .001$, $\eta^2_p = .70$; cycle 3: $F[5,23] = 16.88$, $p < .001$, $\eta^2_p = .42$).

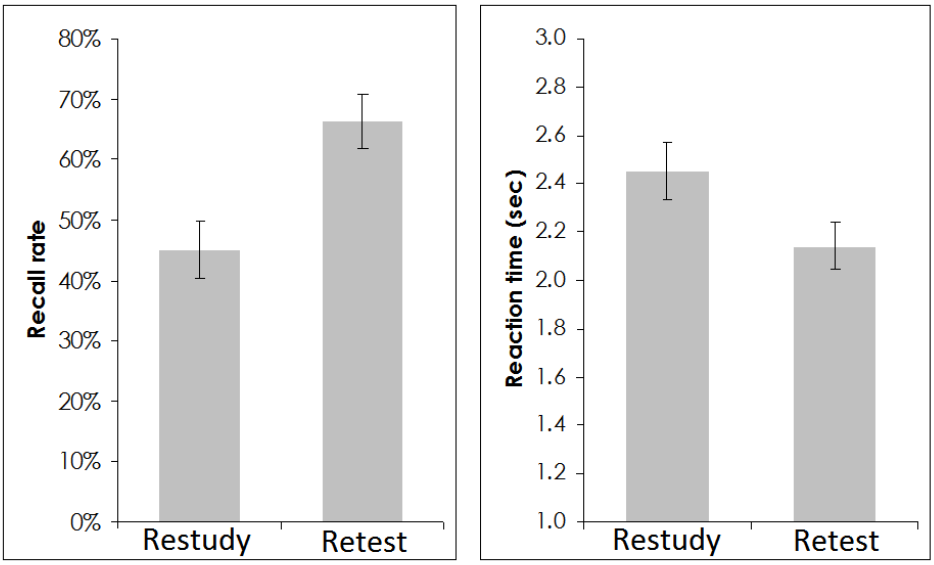
Figure 3. *Practice phase of Experiment 5 (A) recall rates and (B) reaction times.*



Note. Error bars represent standard error of the mean.

In the final test phase, a similar pattern of results was found as in Experiment 4. Recall accuracy in the Retest condition was higher than in the Restudy condition ($t[23] = -5.19$, $p < .001$), whereas reaction time was lower for the retested word pairs than for the restudied items ($t[23] = 2.80$, $p < .05$), see Figure 3. As in Experiment 4, there was a significant negative correlation between reaction time of the last practice cycle and recall rate for the retested word pairs in the final test phase ($r[24] = -.48$, $p < .05$).

Figure 4. *Final test phase of Experiment 5: comparison of (A) recall rates and (B) reaction times between the Restudy and Retest conditions.*



Note. Error bars represent standard error of the mean.

In sum, Experiment 6.1 and 6.2 demonstrated that retrieval practice significantly speeded up long-term access of target memories in comparison with restudy practice. Logarithmic change of reaction times during retrieval practice was present either at low and high initial learning criterion that is suggesting that faster reaction times is not a side effect of increasing learning performance. The moderate and significant correlation between short term reaction times of retrieval and long-term retrieval performance, together with the findings of Keresztes et al. (2014) underlie the assumption that long-term advantage of retrieval practice over restudy mediated by automatization of episodic cued-recall. Due to the repeated retrieval practice of paired-associates, at final recall episodic cues accessed faster and without control the episodic representation of the learned material. Here an automatization account of the testing effect is proposed suggesting that the long-term memory advantage of repeated retrieval over repeated study is due to the automatization of cue-item retrieval association following repeated retrieval. Experimental findings from our laboratory underlie this concept showing a relationship between decreased retrieval reaction time following retrieval practice and long term memory performance. We suggest that retrieval practice decreases the involvement of executive control in retrieval processes, thereby turning memory to skill. These findings are introduced in the discussion part of the dissertation.

VIII.7. Summary

One of the central issues in memory research since the pioneering work of Ebbinghaus (1885) is the seemingly complex relationship between retrieval and forgetting. More specifically, why some memories that were successfully accessed in one occasion failed to be remembered in another retrieval occasion. Note in this theoretical frame improper encoding processes do not belonging to the issue of forgetting, which concept by definition concerns with the loss of once retrievable memories. Until the end of the last century three widely accepted accounts emerged and became widely investigated in the literature of human memory. As it was detailed in the introduction chapter of this dissertation, three influential family of theories emerged in the literature of human forgetting: inhibitory control-based accounts, interference based-accounts, and context-based accounts (Anderson & Bell, 2001;

Anderson, 2005; Raaijmakers & Shiffrin, 1981; Verde, 2013; Jonker, Seli, and MacLeod, 2013; Sahakyan & Delaney, 2003; Sahakyan & Kelley, 2002).

The concept of episodic inhibition is a kind of hybrid account melting together many attributes of inhibitory and context-based accounts of human forgetting (Racsmány and Conway, 2006). This theoretical explanation pinpoints episodic retrieval as the source of memory suppression effects, suggesting that selective retrieval creates and reshapes highly contextualized episodic memory representations (Conway, 2009; Racsmány & Conway, 2006; Racsmány, Conway, Keresztes, & Krajcsi, 2012 see also Karpicke, Lehman, & Aue, 2014). This account assumes that episodic memory sets contain context, cue, and item features (Conway, 2009; Racsmány & Conway, 2006). In this framework, the act of selective episodic retrieval of a studied memory set in the retrieval practice paradigm transcribes the contextual features and the current ratios of cue-item associations of the learnt memory set into a constrained episodic representation, and RIF occurs whenever these association strengths are reestablished through reinstatement of contextual episodic memory sets of the latest retrieval phase (Racsmány & Conway, 2006; Racsmány et al., 2010). This account was supported a series of experimental results presented earlier in this dissertation. The most important findings were summarized in 10 thesis points of this dissertation.

Importantly, derived from the results of Racsmány and Keresztes (2015) it was suggested that an initial retrieval attempt of the entire learning set can eliminate the adverse effect of later selective retrieval. This is because an initial retrieval can already transcribe the entire learning set into an episodic memory representation (see Conway, 2009; Racsmány & Conway, 2006), and establish the episodic context for the rest of the experiment (see Jonker, et al., 2013; Karpicke et al., 2014). This way, final recall will bias the retrieval process to mimic the pattern of the initial retrieval and grant access to items not selectively practiced as well. The presented results, e.g. initial retrieval shields against RIF, can be explained by assuming that selective retrieval leads to RIF by generating a compound contextual episodic memory representation with a restricted and biased search set (Karpicke et al., 2014). In such episodic memory sets, cue-item associations are biased towards increased recall probabilities for retrieved items from practiced categories and decreased recall probabilities for non-retrieved items from practiced categories (Racsmány & Conway, 2006; Jonker et al, 2013). In fact, these are genuine properties of episodic memories (Conway, 2009). Episodic inhibition and context-based accounts suggest mechanisms inherent to episodic retrieval processes to explain most of the findings in the literature concerning selective forgetting. Context-based accounts of RIF and retrieval-enhanced learning (Jonker et al., 2013; Karpicke et al., 2014) emphasize the role of context change between initial study of category-member pairs on the one hand, and selective retrieval and final recall on the other. These accounts predict that an initial retrieval of the entire learning set after the study phase will already have participants change their mental context and later selective retrieval practice will cause no further change in this mental context. As a consequence, the context of the initial retrieval will be the active context at final recall. Beyond an emphasis on a passive contextual shift, episodic accounts of forgetting phenomena (Conway, 2009; Racsmány & Conway, 2006; Racsmány et al., 2012) highlight the active role of retrieval processes in creating and reshaping episodic memory representations. According to the episodic inhibition accounts, episodic retrieval transcribes current contextual information and cue-item association strength ratios of a learning set into an episodic representation. Whenever the same episodic representation is accessed through episodic cues, the encoded cue-item association strength ratios are reinstated. Initial retrieval of the entire learning set can eliminate the adverse effect of later selective retrieval because it transcribes the entire learning set into an episodic representation. When – in the absence of an initial test – the first retrieval is in the practice phase, then final retrieval using the episodic context of the practice phase restricts the search set to practiced items and some arbitrarily activated competitors, whereas other competitors are not involved into the search set at all. That is why RIF is a long-term phenomenon, if final retrieval can reinstate the context of practice phase RIF will be detected following longer delay (Racsmány et al., 2010, see also Abel & Bäuml, 2012). Altogether episodic and context-based accounts of RIF assume that in the retrieval practice paradigm, selective retrieval restricts the search set through encoding a biased contextual information into an episodic memory representation, but an initial, non-selective, retrieval of the entire learning set before the selective retrieval can hinder this search set restriction.

Interestingly, the hypotheses derived from episodic inhibition account recently turned to be suitable to explain not only suppression effects (such as directed forgetting, retrieval-induced forgetting and the forgetting of No-think items in the TNT paradigm). As it was introduced earlier a recent account of the testing effect – the episodic context account of retrieval-enhanced learning (Karpicke et al., 2014) – can be regarded as an extension of episodic and context-based accounts of RIF to a broader range of episodic memory phenomena. This theory aims to explain a range of long-term changes that occur as a consequence of retrieval. Although a detailed presentation of this theory is beyond the scope of this summary chapter, one relevant suggestion of it is that whenever studying and retrieval take place in different temporal contexts, retrieval will reinstate and update the study context by encoding a composite of study and retrieval contexts (see Karpicke et al., 2014; Lohnas & Kahana, in press). On a later test participants will use the updated compound context to restrict the search set – the group of items considered as candidates for retrieval (Karpicke et al., 2014). According to this account, the retrieval practice paradigm involves manipulations that produce different kinds of contexts for practiced and unpracticed categories. That is selectively practiced categories will have the compound context of the study and the practice phases, whereas the unpracticed categories will have solely the context of the study phase. Another specificity of the retrieval practice paradigm is that participants typically retrieve practiced items more than once (the most frequently applied procedure involves three retrieval practice cycles). This procedure enables participants to encode strong and detailed contextual information for the practiced sets, and as it was demonstrated in the last thesis point of this dissertation, the process of cued recall became automatized, in other words episodic cues could access episodic representation without a controlled search process. As a consequence, they probably will rely more on the context of retrieval practice than on the context of study phase during final recall, and this will bias the recall output in favor of practiced items over unpracticed ones, as unpracticed items have no associations to context features of the practice phase. In contrast, participants will reinstate the context of the study phase whenever they use an unpracticed category label as a retrieval cue.

In other words, according to this account – also in line with the episodic inhibition explanations of RIF – RIF is due to a core attribute of retrieval; it is present when the updated context of the selective retrieval allows the participants to restrict their search set mainly for the practiced items. The initial retrieval in our experiments let participants to update the context of the study phase with the context of the initial retrieval. As a consequence, receiving the category cue they could use the compound context of study and initial retrieval while attempting to retrieve unpracticed items from practiced categories at final recall. In this view, retrieval is the key process that enhances long-term accessibility of retrieved memories and it is the process that can hinder retrieval of items through search set restriction or can shield against the adverse effect of later selective retrieval.

REFERENCE LIST

Abel, M., & Bäuml, K. H. T. (2012). Retrieval-induced forgetting, delay, and sleep. *Memory*, *20*, 420-428.

Allen, S. W., & Vokey, J. R. (1998). Directed forgetting and rehearsal on direct and indirect memory tests. In J. M. Golding & C. M. MacLeod (Eds.), *Intentional forgetting: Interdisciplinary approaches* (pp. 173-197). Mahwah, NJ: Lawrence Erlbaum Associates.

Anderson, J. R. (1972). FRAN: A simulation model of free recall. In G. H. Bower (Ed.), *The psychology of learning and motivation, Vol. 5* (pp. 315-378). New York: Academic Press.

Anderson, J. R. (1974). Retrieval of propositional information from long-term memory. *Cognitive Psychology*, *6*, 541-474.

Anderson, J. R. (1986). *The architecture of cognition.* Cambridge, Massachusetts: Harvard University Press.

Anderson, M. C. (2001). Active forgetting: Evidence for functional inhibition as a source of memory failure. *Journal of Aggression, Maltreatment and Trauma*, *4*, 185-210.

Anderson, M. C. (2003). Rethinking interference theory: Executive control and the mechanisms of forgetting. *Journal of Memory and Language*, *49*, 415-445.

Anderson, M. C., & Bell, T. (2001). Forgetting our facts: The role of inhibitory processes in the loss of propositional knowledge. *Journal of Experimental Psychology: General*, *130*, 544-570.

Anderson, M. C., & Green, C. (2001). Suppressing unwanted memories by executive control. *Nature*, *410*, 366-369.

Anderson, M. C., & Hanslmayr, S. (2014). Neural mechanisms of motivated forgetting. *Trends in Cognitive Sciences*, *18*, 279-292.

Anderson, M. C., & Huddleston, E. (2012). Towards a cognitive and neurobiological model of motivated forgetting. In R. F. Belli (Ed.), *True and false recovered memories* (pp. 53-120). New York: Springer.

Anderson, M. C., & Levy, B. J. (2007). Theoretical issues in inhibition: Insights from research on human memory. In D. S. Gorfein & C. M. MacLeod (Eds.), *Inhibition in cognition* (pp. 81-102). Washington, DC: American Psychological Association.

Anderson, M. C., & Levy, B. J. (2011). On the relationship between interference and inhibition in cognition. In A. S. Benjamin (Ed.), *Successful remembering and successful forgetting: A festschrift in honor of Robert A. Bjork* (pp. 107-132). London: Psychology Press.

Anderson, M. C., & McCulloch, K. C. (1999). Integration as a general boundary condition on retrieval-induced forgetting. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*, 608-629.

Anderson, M. C., & Neely, J. H. (1996). Interference and inhibition in memory retrieval. In E. L. Bjork & R. A. Bjork (Eds.), *Memory: Handbook of perception and cognition* (pp. 237-313). San Diego: Academic Press.

Anderson, M. C., & Spellman, B. A. (1995). On the status of inhibitory mechanisms in cognition: Memory retrieval as a model case. *Psychological Review, 102*, 68-100.

Anderson, M. C., Bjork, E. L., & Bjork, R. A. (2000). Retrieval-induced forgetting: Evidence for a recall-specific mechanism. *Psychonomic Bulletin and Review*, *7*, 522-530.

Anderson, M. C., Bjork, R. A., & Bjork, E. L. (1994). Remembering can cause forgetting: Retrieval dynamics in long-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*, 1063-1087.

Anderson, M. C., Bunce, J. G., & Barbas, H. (in press). Prefrontal-hippocampal pathways underlying inhibitory control over memory. *Neurobiology of Learning and Memory*.

Anderson, M. C., De Kok, D., & Child, C. (1997). Retrieval-induced forgetting on a test of recognition memory. *Abstracts of the Psychonomic Society*, *2*, Volume 2.

Anderson, M. C., Ochsner, K. N., Kuhl, B., Cooper, J., Robertson, E., Gabrieli, S. W., & Gabrieli, J. D. (2004). Neural systems underlying the suppression of unwanted memories. *Science*, *303*, 232-235.

Baddeley, A. D., & Della Sala, S. (1998). Working memory and executive control. In A. C. Roberts, T. W. Robbins, & L. Weiskrantz (Eds.), *The prefrontal cortex: Executive and cognitive functions* (pp. 9-22). Oxford: Oxford University Press.

Baddeley, A. D., Bressi, S., Della Sala, S., Logie, R., & Spinnler, H. (1991). The decline of working memory in Alzheimer's disease: A longitudinal study. *Brain*, *114*, 2521-2542.

Baldo, J. V., & Shimamura, A. P. (1998). Letter and category fluency in patients with frontal lobe lesions. *Neuropsychology*, *12*, 259-267.

Barnier, A. J., Conway, M. A., Mayoh, L., & Speyer, J. (in press). Directed forgetting of autobiographical memory: Repression reconceived.

Barnier, A. J., Wright, J., & McConkey, K. M. (in press). Posthypnotic amnesia for autobiographical episodes: Influencing memory accessibility and quality. *International Journal of Clinical and Experimental Hypnosis.*

Basden, B. H., & Basden, D. R. (1996). Directed forgetting: A further comparison of the list and item methods. *Memory*, *4*, 633-653.

Basden, B. H., & Basden, D. R. (1998). Directed forgetting: A contrast of methods and interpretations. In J. M. Golding & C. M. MacLeod (Eds.), *Intentional Forgetting: Interdisciplinary Approaches* (pp. 139-173). Mahwah, NJ: Lawrence Erlbaum Associates.

Basden, B. H., Basden, D. R., & Gargano, G. J. (1993). Directed forgetting in implicit and explicit memory tests: A comparison of methods. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*, 603-616.

Basden, D. R., Basden, B. H., & Galloway, B. C. (1977). Inhibition with part-list cueing: Some tests of the item strength hypothesis. *Journal of Experimental Psychology: Human Learning and Memory*, *3*, 10-108.

Bäuml, K. H. (1998). Strong items get suppressed, weak items do not: The role of item strength in output interference. *Psychonomic Bulletin and Review, 5*, 459-463.

Bäuml, K. H. (2002). Semantic generation can cause episodic forgetting. *Psychological Science*, *13*, 356-360.

Bäuml, K. H., & Hartinger, A. (2002). On the role of item similarity in retrieval-induced forgetting. *Memory*, *10*, 215-224.

Bäuml, K. H., & Schlichting, A. (2014). Memory retrieval as a self-propagating process. *Cognition*, *132*, 16-21.

Bäuml, K. H., Pastötter, B., & Hanslmayr, S. (2010). Binding and inhibition in episodic memory - Cognitive, emotional, and neural processes. *Neuroscience and Biobehavioral Reviews, 34*, 1047-1054.

Benoit, R. G., & Anderson, M. C. (2012). Opposing mechanisms support the voluntary forgetting of unwanted memories. *Neuron*, *76*, 450-460.

Bjork, E. L., & Bjork, R. A. (1996). Continuing influences of to-be-forgotten information. *Consciousness and Cognition*, *5*, 176-196.

Bjork, E. L., Bjork, R. A., & Anderson, M. C. (1998). Varieties of goal-directed forgetting. In J. M. Golding & C. M. MacLeod (Eds.), *Intentional forgetting: Interdisciplinary approaches* (pp. 103-139). Mahwah, NJ: Lawrence Erlbaum Associates.

Bjork, R. A. (1970). Positive forgetting: The noninterference of items intentionally forgotten. *Journal of Verbal Learning and Verbal Behavio*r, *9*, 255-268.

Bjork, R. A. (1989). Retrieval inhibition as an adaptive mechanism in human memory. In H. L. Roediger & F. I. M. Craik (Eds.), *Varieties of memory and consciousness: Essays in honour of Endel Tulving* (pp. 309-330). Hillsdale, NJ: Lawrence Erlbaum Associates.

Block, R. A. (1971). Effects of instructions to forget in short-term memory. *Journal of Experimental Psychology*, *89*, 1-9.

Bodenhause, G. V., & Macrae, C. N. (1998). Stereotype activation and inhibition. In R. S. Wyer (Ed.), *Stereotype activation and inhibition: Advances in social cognition* (pp. 1-52). New Jersey: Erlbaum.

Brewer, W. F. (1986). What is autobiographical memory? In D. C. Rubin (Ed.), *Autobiographical Memory* (pp. 25-49). Cambridge: Cambridge University Press.

Bunge, S. A., Oschner, K. N., Desmond, J. E., Glover, G. H., & Gabrieli, J. D. E. (2001). Prefrontal regions involved in keeping information in and out of mind. *Brain*, *124*, 2074-2086.

Burwitz, L. (1974). Proactive interference and directed forgetting in short-term motor memory. *Journal of Experimental Psychology, 102*, 799-805.

Bush, G., Vogt, B. A., Holmes, J., Dale, A. M., Greve, D., Jenike, M. A., & Rosen, B. R. (2002). Dorsal anterior cingulate cortex: A role in reward-based decision making. *Proceedings of the National Academy of Sciences*, *99*, 523-528.

Butler, K. M, Williams, C. C., Zacks, R. T., & Maki, R. H. (2001). A limit on retrieval-induced forgetting. Journal of Experimental Psychology*: Learning, Memory and Cognition*, *27*, 1314-1318.

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

Carpenter, S. K. (2011). Semantic information activated during retrieval contributes to later retention: Support for the mediator effectiveness hypothesis of the testing effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*, 1547-1552.

Carr, T. H. (1986). Perceiving visual language. In K. Boff, L. Kauffman, & J. Thomas (Eds.), *Handbook of perception and human performance* (pp. 1-92). New York: Wile.

Carr, T. H., & Dagenbach, D. (1990). Semantic priming and repetition priming from masked words: Evidence for a center-surround attentional mechanism in perceptual recognition. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *16*, 341-350.

Ciranni, M. A., & Shimamura, A. P. (1999). Retrieval-induced forgetting in episodic memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*, 1403-1414.

Cloitre, M. (1998). Intentional forgetting and clinical disorders. In J. M. Golding & C. M. MacLeod (Eds.), *Intentional forgetting: Interdisciplinary approaches* (pp. 395-413). London: Lawrence Erlbaum Associates.

Cloitre, M., Cancienne, J., Brodsky, B., Dulit, R., & Perry, S. W. (1996). Memory performance among women with parental abuse histories: Enhanced directed forgetting or directed remembering? *Journal of Abnormal Psychology*, *105*, 204-211.

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review, 82*, 407-428.

Conway, M. A. (1988). Images in autobiographical memory. In M. Denis, J. Engelkamp, & J. T. E. Richardson (Eds.), *Cognitive and neuropsychological approaches to mental imagery* (pp. 337-346). Dordrecht, the Netherlands: Martinus Nijhoff.

Conway, M. A. (1990). *Autobiographical memory: An introduction.* Buckingham: Open University Press.

Conway, M. A. (2001). Repression revisited. *Nature*, *410*, 319-320.

Conway, M. A. (2001). Sensory perceptual episodic memory and its context: Autobiographical memory. In A. D. Baddeley, J. P. Aggleton, & M. A. Conway (Eds.), *Episodic memory: New directions in research* (pp. 53-71). Oxford: Oxford University Press.

Conway, M. A. (2005). Memory and the self. *Journal of Memory and Language*, *53*, 594-628.

Conway, M. A. (2009). Episodic memories. *Neuropsychologia*, *47*, 2305-2313.

Conway, M. A., & Pleydell-Pearce, C. W. (2000). The construction of autobiographical memories in the self memory system. *Psychological Review*, *107*, 261-288.

Conway, M. A., Harries, K., Noyes, J., Racsmány, M., & Frankish, C. (2000). The disruption and dissolution of directed forgetting: Inhibitory control of memory. *Journal of Memory and Language, 43*, 409-430.

Conway, M. A., Pleydell-Pearce, C. W., & Whitecross, S. (2001). The neuroanatomy of autobiographical memory: A slow cortical potential study (SCP) of autobiographical memory retrieval. *Journal of Memory and Language*, *45*, 493-524.

Conway, M., Howell, A., & Giannopoulos, L. (1991). Dysphasia and thought suppression. *Cognitive Therapy and Research*, *15*, 153-166.

Conway, M. A., & Fthenaki, A. (2003). Disruption of inhibitory control of memory following lesions to the frontal and temporal lobes. *Cortex, 39*, 667-686.

Crowder, R. G. (1976). *Principles of learning and memory: Classic edition*. New York: Psychology Press.

Cruse, D., & Jones, R. A. (1976). Intentional forgetting of tones in a choice recognition-time task. *Journal of Experimental Psychology: Human Learning and Memory*, *2*, 577-585.

Dagenbach, D., & Carr, T. H. (1994). Inhibitory processes in perceptual recognition. In D. Dagenbach & T. H. Carr (Eds.), *Inhibitory processes in attention, memory and language* (pp. 265-327). San Diego: Academic Press.

Davis, J. C., & Okada, R. (1971). Recognition and recall of positively forgotten items. *Journal of Experimental Psychology*, *89*, 181-186.

Dempster, F. N. (1991). Inhibitory processes: A neglected dimension of intelligence. *Intelligence*, *15*, 157-193.

Depue, B. E., Curran, T., & Banich, M. T. (2007). Prefrontal regions orchestrate suppression of emotional memories via a two-phase process. *Science*, *317*, 215-219.

Desimone, R., & Duncan, J. (1994). Neural mechanisms of selective visual attention. *Annual Review of Neurosciences*, *18*, 193-222.

Duncan, J., & Owen, A. M. (2000). Common regions of the human frontal lobe recruited by diverse cognitive demands. *Trends in Neurosciences*, *23*, 475-483.

Ebbinghaus, H. (1885). *Uber das Gedachtnis*. Leipzig, Germany: K. Buehler.

Elmes, D. G., Adams, C. A., & Roediger, H. L. (1970). Cued forgetting in short-term memory: Response selection. *Journal of Experimental Psychology*, *86*, 103-107.

Engle, R. W., Conway, A. R. A., Tuholski, S. W., & Shisler, R. J. (1995). A resource account of inhibition. *Psychological Science*, *6*, 122-125.

Epstein, W. (1969). Poststimulus output specification and differential retrieval from short-term memory. *Journal of Experimental Psychology*, *82*, 168-174.

Erdelyi, M. H., & Goldberg, B. (1979). Let's not sweep repression under the rug: Toward a cognitive psychology of repression. In J. F. Kihlstrom & F. J. Evans (Eds.), *Functional disorders of memory* (pp. 355-401). Hillsdale, NJ: Lawrence Erlbaum Associates.

Estes, W. K. (1955). Statistical theory of spontaneous recovery and regression. *Psychological Review*, *62*, 145-154.

Estes, W. K. (1972). An associative basis for coding and organization in memory. In A. W. Melton & E. Martin (Eds.), *Coding processes in human memory* (pp. 161-190). Washington, DC: Winston.

Euston, D. R., Tatsuno, M., & McNaughton, B. L. (2007). Fast-forward playback of recent memory sequences in prefrontal cortex during sleep. *Science*, *318*, 1147-1150.

Finkelhor, D., & Browne, A. (1985). The traumatic impact of child sexual abuse: A conceptualization. *American Journal of Orthopsychiatry*, *55*, 530-541.

Fleck, D. E., Berch, D. B., Shear, P. K., & Strakowski, S. M. (2001). Directed forgetting in explicit and implicit memory: The role of encoding and retrieval mechanisms. *The Psychological Record*, *5*, 207-220.

Freyd, J. J. (1996). *Betrayal trauma: The logic of forgetting childhood abuse.* Cambridge, MA: Harvard University Press.

Füredi, M., & Kelemen, J. (Szerk.), *A mai magyar nyelv szépprózai gyakorisági szótára* [A frequency dictionary of the literary language of Hungarian]. 1989. Budapest: Akadémiai Kiadó.

Gardiner, J. M., & Richardson-Klavehn, A. (2000). Remembering and knowing. In E. Tulving (Ed.), *The Oxford handbook of memory* (pp. 229-244). Oxford: University Press.

Gardiner, J. M., Gawlik, B., & Richardson-Klavehn, A. (1994). Maintenance rehearsal affects knowing, not remembering: Elaborative rehearsal affects remembering, not knowing. *Psychonomic Bullettin and Review*, *1*, 107-110.

Gardiner, J. M., Ramponi, C., & Richardson-Klavehn, A. (1998). Experiences of remembering, knowing, and guessing. *Consciousness and Cognition*, *7*, 1-26.

Geiselman, R. E., & Bagheri, B. (1985). Repetition effects in directed forgetting: Evidence for retrieval inhibition. *Memory and Cognition*, *13*, 57-62.

Geiselman, R. E., & Richle, J. P. (1975). The fate of to-be-forgotten sentences in semantic positive forgetting. *Bulletin of the Psychonomic Society*, *6*, 19-21.

Geiselman, R. E., Bjork, R. A., & Fishman, D. L. (1983). Disrupted retrieval in directed forgetting: A link with posthypnotic amnesia. *Journal of Experimental Psychology: General*, *112*, 58-72.

Geiselman, R. E., Rabow, W. E., Wachtel, S. L., & MacKinnon, D. P. (1985). Strategy control in intentional forgetting. *Human Learning*, *4*, 169-178.

Gernsbacher, M. A. (1990). *Language comprehension as structure building*. Hillsdale, NJ: Erlbaum.

Gernsbacher, M. A., & Faust, M. F. (1991). The mechanism of suppression: A component of general comprehension skill. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17*, 245-262.

Gernsbacher, M. A., & Faust, M. F. (1995). Skilled suppression. In F. N. Dempster & C. J. Brainerd (Eds.), *Interference and inhibition in cognition* (pp. 295-327). San Diego: Academic Press.

Glucksberg, S., & King, L. J. (1967). Motivated forgetting mediated by implicit verbal chaining: A laboratory analogue of repression. *Science*, *158*, 517-518.

Graf, P., & Mandler, G. (1984). Activation makes words more accessible, but not necessarily more retrievable. *Journal of Verbal Learning and Verbal Behavior*, *23*, 553-568.

Guthrie, E. R. (1935). *The psychology of learning.* New York: Harper.

Hamm, V. P., & Hasher, L. (1992). Age and availability of inferences. *Psychology and Aging*, *7*, 56-64.

Hartman, M., & Hasher, L. (1991). Aging and suppression: Memory for previously relevant information. *Psychology and Aging, 6*, 587-594.

Hasher, L., Stoltzfus, E. R., Zacks, R. T., & Rypma, B. (1991). Age and inhibition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17*, 163-169.

Hauselt, J. (1998). An illusion of retrieval inhibition: Directed forgetting and implicit memory. In J. M. Golding & C. M. MacLeod (Eds.), *Intentional forgetting: Interdisciplinary approaches* (pp. 197-219). Mahwah, NJ: Lawrence Erlbaum Associates.

Hebb, D. O. (1949). *The organization of behavior: A neuropsychological theory.* New York: Wiley.

Herman, J. L. (1981). *Father-daughter incest.* Cambridge, Mass: Harvard University Press.

Horton, K. D., & Petruk, R. (1980). Set differentation and depth of processing in the directed forgetting paradigm. *Journal of Experimental Psychology: Human Learning and Memory, 6*, 599-610.

Hulbert, J. C., Shivde, G., & Anderson, M. C. (2012). Evidence against associative blocking as a cause of cue-independent retrieval-induced forgetting. *Experimental Psychology, 59*, 11-21.

Jackson, J. L., & Michon, J. A. (1984). Effects of item concreteness on temporal coding. *Acta Psychologica*, *57*, 83-95.

Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language*, *30*, 513-541.

Jacoby, L. L., & Hollingshead, A. (1990). Toward a generate/recognize model of performance on direct and indirect tests on memory. *Journal of Memory and Language*, *29*, 433-454.

Jenkins, J. G., & Dallenbach, K. M. (1924). Oblivescence during sleep and waking. *American Journal of Psychology*, *35*, 605-612.

Jonker, T. R., Seli, P., & MacLeod, C. M. (2013). Putting retrieval-induced forgetting in context: An inhibition-free, context-based account. *Psychological Review*, *120*, 852-872.

Karpicke, J. D. (2012). Retrieval-based learning active retrieval promotes meaningful learning. *Current Directions in Psychological Science*, *21*, 157-163.

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

Karpicke, J. D., Lehman, M., & Aue, W. R. (2014). Retrieval-based learning: An episodic context account. In B. H. Ross (Ed.), *Psychology of learning and motivation*, *Vol. 61* (pp. 237-284). San Diego: Elsevier Academic Press.

Keele, S. W., & Neill, W. T. (1978). Mechanisms of attention. In E. C. Carterete & M. P. Friedman (Eds.), *Handbook of perception IX* (pp. 3-47). New York: Academic Press.

Kihlstrom, J. F. (2002). No need for repression. *Trends in Cognitive Sciences*, *6*, 502.

Kimberg, D. Y., & Farah, M. J. (1993). A unified account of cognitive impairments following frontal lobe damage: The role of working memory in complex organized behavior. *Journal of Experimental Psychology: General*, *122*, 411-428.

Kónya, A., & Pintér, G. (1985). Kategória norma a verbális emlékezet vizsgálatához [Category norms for verbal memory research]. *Hungarian Psychological Review*, *2*, 93-111.

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

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

Lesh, T. A., Niendam, T. A., Minzenberg, M. J., & Carter, C. S. (2011). Cognitive control deficits in schizophrenia: Mechanisms and meaning. *Neuropsychopharmacology*, *36*, 316-338.

Levy, B. J., & Anderson, M. C. (2002). Inhibitory processes and the control of memory retrieval. *Trends in Cognitive Sciences*, *6*, 299-305.

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

MacDonald, A. W., Cohen, J. D., Stenger, V. A., & Carter, C. S. (2000). Dissociating the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control. *Science*, *288*, 1835-1838.

MacLeod, C. M. (1989). Directed forgetting affects both direct and indirect tests of memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*, 13-21.

MacLeod, C. M. (1998). Directed forgetting. In J. M. Golding & C. M. MacLeod (Eds.), *Intentional forgetting: Interdisciplinary approaches* (pp. 1-59). Mahwah, NJ: Lawrence Erlbaum Associates.

MacLeod, C. M. (2007). The concept of inhibition in cognition. In D. S. Gorfein & C. M. MacLeod (Eds.), *Inhibition in cognition* (pp. 3-23). Washington, DC: American Psychological Association.

MacLeod, C. M. (2012). Directed forgetting. In N. M. Seel (Ed.), *Encyclopedia of the sciences of learning* (pp. 993-995). New York: Springer.

MacLeod, M. (2002). Retrieval-induced forgetting in eyewitness memory: Forgetting as a consequence of remembering. *Applied Cognitive Psychology*, *16*, 135-149.

MacLeod, M. D., & Macrae, C. N. (2001). Gone but not forgotten: The transient nature of retrieval-induced forgetting. *Psychological Science*, *12*, 148-152.

Macrae, C. N., & MacLeod, M. D. (1999). On recollection lost: When practice makes imperfect. *Journal of Personal and Social Psychology, 77*, 463-473.

Martin, E. (1971). Verbal learning theory and independent retrieval phenomena. *Psychological Review, 78*, 314-332.

Mathews, A., & MacLeod, C. (1994). Cognitive approaches to emotional disorders. *Annual Review of Psychology*, *45*, 25-50.

McClelland, J. L., & Rumelhart, D. E. (1986). *Parallel distributed processing. Vol. 2.* Cambridge, Mass: MIT Press.

McDowd, J. M., & Oseas-Kreger, D. M. (1991). Aging, inhibitory processes, and negative priming. *Journal of Gerontology*, *46*, 340-345.

McGeoch, J. A. (1936). Studies of retroactive inhibition: VII. Retroactive inhibition as a function of the length and frequency of presentation of the interpolated lists. *Journal of Experimental Psychology, 19*, 674-693.

McGeoch, J. A. (1942). *The Psychology of Human Learning*. New York: Longmans Green.

McGeoch, J. A., & Irion, A. L. (1942). *The psychology of learning*. New York: McKay.

McNally, R. J., Metzger, L. J., Lasko, N. B., Clancy, S. A., & Pitman, R. K. (1998). Directed forgetting of trauma cues in adult survivors of childhood sexual abuse with and without posttraumatic trauma. *Journal of Abnormal Psychology, 107*, 596-601.

Melton, A. W., & Irwin, J. M. (1940). The influence of degree of interpolated learning on retroactive inhibition and the overt transfer of specific responses. *American Journal of Psychology*, *3*, 173-203.

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neurosciences*, *21*, 167-202.

Minami, H., & Dallenbach, K. M. (1946). The effect of activity upon learning and retention in the cockroach. *American Journal of Psychology*, *59*, 1-58.

Minzenberg, M. J., Laird, A. R., Thelen, S., Carter, C. S., & Glahn, D. C. (2009). Meta-analysis of 41 functional neuroimaging studies of executive function in schizophrenia. *Archives of General Psychiatry*, *66*, 811-822.

Moulin, C. J. A., Perfect, T. J., Conway, M. A., North, A. S. Jones, R. W., & Niamh, J. (2002). Retrieval-induced forgetting in Alzheimer's disease. *Neuropsychologia*, *40*, 862-867.

Mueller, C. W., & Watkins, M. J. (1977). Inhibition from part-set cueing: A cue-overload interpretation. *Journal of Verbal Learning and Verbal Behavior, 16*, 699-710.

Murdock, B. B. (1962). The serial position effect of free recall. *Journal of Experimental Psychology: General*, *64*, 482-488.

Neely, J. H. (1977). Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited-capacity attention. *Journal of Experimental Psychology: General*, *106*, 226-254.

Neely, J. H. (1991). Semantic priming effects in visual word recognition. A selective review of current findings and theories. In D. Besner & G. Humphreys (Eds.), *Basic processes in reading: Visual word recognition* (pp. 264-336). Hillsdale, NJ: Lawrence Erlbaum.

Neill, W. T. (1977). Inhibitory and facilitatory processes in selective attention. *Journal of Experimental Psychology: Human Perception and Performance*, *3*, 444-450.

Nickerson, R. S. (1984). Retrieval inhibition from part-set cuing: A persisting enigma in memory research. *Memory and Cognition*, *12*, 531-552.

Niendam, T. A., Laird, A. R., Ray, K. L., Dean, Y. M., Glahn, D. C., & Carter, C. S. (2012). Meta-analytic evidence for a superordinate cognitive control network subserving diverse executive functions. *Cognitive, Affective, and Behavioral Neuroscience*, *12*, 241-268.

Norman, K. A., Newman, E. L., & Detre, G. (2007). A neural network model of retrieval-induced forgetting. *Psychological Review*, *114*, 887-953.

Paller, K. A. (1990). Recall and stem completion priming have different electrophysiological correlates and are modified differentially by directed forgetting. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*, 1021-1032.

Park, J., & Kanwisher, N. (1994). Negative priming for spatial location: Identity mismatching, not distractor inhibition. *Journal of Experimental Psychology: Human Perception and Performance*, *20*, 613-623.

Pastötter, B., Hanslmayr, S., & Bäuml, K. H. T. (2010). Conflict processing in the anterior cingulate cortex constrains response priming. *NeuroImage*, *50*, 1599-1605.

Perfect, T., Tree, J., Moulin, C. J. A., Hutter, R., & Stark, L. (2001). What causes retrieval-induced forgetting? *British Experimental Psychology Society, Annual Meeting.* Bristol, UK, April 5-7. Abstract.

Posner, M. I., & Snyder, C. R. R. (1975). Facilitation and inhibition in the processing of signals. In P. M. Rabbit & S. Dornic (Eds.), *Attention and Performance X* (pp. 669-682). Hillsdale, NJ: Lawrence Erlbaum.

Postman, L. (1971). Organization and interference. *Psychological Review*, *78*, 290-302.

Postman, L., & Underwood, B. J. (1973). Critcal issues in interference theory. *Memory and Cognition*, *1*, 19-40.

Postman, L., Stark, K., & Fraser, J. (1968). Temporal changes in interference. *Journal of Verbal Learning and Verbal Behavior, 7*, 672-694.

Raaijmakers, J. G., & Shiffrin, R. M. (1981). Search of associative memory. *Psychological Review*, *88*, 93-134.

Reitman, W., Malin, J. T., Bjork, R. A., & Higman, B. (1973). Strategy control and directed forgetting. *Journal of Verbal Learning and Verbal Behavio*r, *12*, 140-149.

Richardson-Klavehn, A., & Gardiner, J. M. (1996). Cross-modality priming in stem completion reflects conscious memory, but not voluntary memory. *Psychonomic Bulletin and Review*, *3*, 238-244.

Roberts, R. J., Hager, L. D., & Heron, C. (1994). Prefrontal cognitive processes: Working memory and inhibition in the antisaccade task. *Journal of Experimental Psychology*, *General, 123*, 374-393.

Roediger, H. L. (1973). Inhibition in recall from cueing with recall targets. *Journal of Verbal Learning and Verbal Behavior*, *12*, 644-657.

Roediger, H. L. (1974). Inhibiting effects of recall. *Memory and Cognition*, *2*, 261-296.

Roediger, H. L. (1978). Recall as a self-limiting process. *Memory and Cognition*, *6*, 54-63.

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

Roediger, H. L., & Guynn, M. J. (1996). Retrieval processes. In E. L. Bjork & R. A. Bjork (Eds.), *Human memory* (pp. 197-236). San Diego: Academic Press.

Roediger, H. L., & Karpicke, J. D. (2006). Test-enhanced learning taking memory tests improves long-term retention. *Psychological Science*, *17*, 249-255.

Roediger, H. L., & Neely, J. H. (1982). Retrieval blocks in episodic and semantic memory. *Canadian Journal of Psychology*, *36*, 213-242.

Roediger, H. L., & Schmidt, S. R. (1980). Output interference in the recall of categorized and paired associate lists. *Journal of Experimental Psychology: Human Learning and Memory, 6*, 91-105.

Roediger, H. L., Stellon, C., & Tulving, E. (1977). Inhibition from part-list cues and rates of recall. *Journal of Experimental Psychology: Human Learning and Memory*, *3*, 174-188.

Rosen, V. M., & Engle, R. W. (1998). Working memory capacity and suppression. *Journal of Memory and Language*, *39*, 418-436.

Rundus, D. (1973). Negative effects of using list items as recall cues. *Journal of Verbal Learning and Verbal Behavior*, *12*, 43-50.

Russel, D. E. H. (1986). *The secret trauma: Incest in the lives of girls and women.* New York: Basic Books.

Russo, R., & Andrade, J. (1995). The directed forgetting effect in word-fragment completion: An application of the process dissociation procedure. *Quarterly Journal of Experimental Psychology*, *48*, 405-423.

Sahakyan, L., & Delaney, P. F. (2003). Can encoding differences explain the benefits of directed forgetting in the list method paradigm? *Journal of Memory and Language*, *48*, 195-206.

Sahakyan, L., & Hendricks, H. E. (2012). Context change and retrieval difficulty in the list-before-last paradigm. *Memory and Cognition*, *40*, 844-860.

Sahakyan, L., & Kelley, C. M. (2002). A contextual change account of the directed forgetting effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*, 1064-1072.

Schacter, D. L. (1996). *Searching for memory: The brain, the mind, and the past.* New York: Basic Books.

Shaw, J. S, Bjork, R. A., & Handal, A. (1995). Retrieval-induced forgetting in an eyewitness-memory paradigm. *Psychonomic Bulletin and Review*, *2*, 249-253.

Shebilske, W., Wilder, L., & Epstein, W. (1971). Forget instructions: Effect of selective rehearsal and categorical distinctiveness. *Journal of Experimental Psychology*, *89*, 372-378.

Shimamura, A. P. (1995). Memory and frontal lobe function. In M. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 803-813). Cambridge, MA: MIT Press.

Shivde, G., & Anderson, M. C. (2001). The role of inhibition in meaning selection: Insights from retrieval-induced forgetting. In D. S. Gorfein (Ed.), *On the consequences of meaning selection:*

Perspectives on resolving ambiguity (pp. 175-190). Washington, DC: American Psychological Association.

Slamecka, N. J. (1968). An examination of trace storage in free recall. *Journal of Experimental Psychology*, *76*, 504-513.

Slamecka, N. J. (1969). Testing for associative storage in multitrial free recall. *Journal of Experimental Psychology*, *91*, 557-560.

Slamecka, N. J. (1972). The question of associative growth in the learning of categorized materials. *Journal of Verbal Learning and Verbal Behavior, 11*, *324-332.*

Sloman, S. A., Bower, G. H., & Roher, D. (1991). Congruency effects in part-list cueing inhibition. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *17*, 974-982.

Smith, A. D. (1971). Output interference and organized recall from long-term memory. *Journal of Verbal Learning and Verbal Behavior*, *10*, 400-408.

Smith, A. D. (1973). Input order and output interference in organized recall. *Journal of Experimental Psychology*, *100*, 147-150.

Spinnler, H., Della Sala, S., Bandera, R., & Baddeley, A. D. (1988). Dementia, ageing and the structure of human memory. *Cognitive Neuropsychology*, *5*, 193-211.

Squire, L. R., Shimamura, A. P., & Graf, P. (1987). Strength and duration of priming effects in normal subjects and amnesic patients. *Neuropsychologia*, *25*, 195-210.

Storm, B. C. (2011). Retrieval-induced forgetting and the resolution of competition. In A. Benjamin (Ed.), *Successful remembering and successful forgetting: A festschrift in honor of Robert A. Bjork* (pp. 89-105). London: Psychology Press.

Storm, B. C., & Nestojko, J. F. (2010). Successful inhibition, unsuccessful retrieval: Manipulating time and success during retrieval practice. *Memory*, *18*, 99-114.

Storm, B. C., Bjork, E. L., Bjork, R. A., & Nestojko, J. F. (2006). Is retrieval success a necessary condition for retrieval-induced forgetting? *Psychonomic Bulletin and Review*, *13*, 1023-1027.

Timmins, W. K. (1974). Varying processing time in directed forgetting. *Journal of Verbal Learning and Verbal Behavio*r, *13*, 539-544.

Tipper, S. P. (1985). The negative priming effect: Inhibitory priming by ignored objects. *Quarterly Journal of Experimental Psychology*, *37*, 571-590.

Tipper, S. P. (1991). Less attentional selectivity as a result of declining inhibition in older adults. *Bulletin of the Psychonomic Society*, *29*, 45-47.

Tolin, D. F., Hamlin, C., & Foa, E. B. (2002). Directed forgetting in obsessive-compulsive disorder: Replication and extension. *Behaviour Research and Therapy*, *40*, 793-803.

Toth, J. P., Reingold, E. M., & Jacoby, L. L. (1994). Toward a redefinition of implicit memory: Process dissociations following elaborative processing and self-generation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*, 290-303.

Tulving, E. (1983). *Elements of episodic memory.* New York: Oxford University Press.

Tulving, E. (1985). Memory and consciousness. *Canadian Psychology*, *26*, 1-12.

Tulving, E., & Arbuckle, T. Y. (1963). Sources of intratrial interferences in paired-associate learning. *Journal of Verbal Learning and Verbal Behavior*, *1*, 321-334.

Tulving, E., & Pearlstone, Z. (1966). Availability versus accessibility of information in memory for words. *Journal of Verbal Learning and Verbal Behavio*r, *5*, 381-391.

Tzeng, O. J. L., Lee, A. T., & Wetzel, C. D. (1979). Temporal coding in verbal information processing. *Journal of Experimental Psychology: Human Learning and Memory, 5*, 52-64.

Verde, M. F., & Perfect, T. J. (2011). Retrieval-induced forgetting in recognition is absent under time pressure. *Psychonomic Bulletin and Review*, *18*, 1166-1171.

Watkins, M. J. (1975). Inhibition in recall with extralist 'cues'. *Journal of Verbal Learning and Verbal Behavior, 14*, 294-303.

Wegner, D. M. (1994). Ironic processes of mental control. *Psychological Review*, *101*, 34-52.

Weiner, B. (1968). Motivated forgetting and the study of repression. *Journal of Personality*, *36*, 213-234.

Weiner, B., & Reed, H. (1969). Effects of the instructional sets to remember and to forget on short-term retention: Studies of rehearsal control and retrieval inhibition (repression). *Journal of Experimental Psychology, 79*, 226-232.

Wenzlaff, R. M., Wegner, D. M, & Roper, D. (1988). Depression and mental control. The resurgence of unwanted negative thoughts. *Journal of Personality and Social Psychology*, *55*, 882-892.

Wenzlaff, R. M., Wegner, D. M., & Klein, S. B. (1991). The role of thought suppression in the bonding of thought and mood. *Journal of Personality and Social Psychology*, *60*, 500-508.

Wetzel, C. D. (1975). Effects of orienting tasks and cue timing on the free recall of remember and forget-cued words. *Journal of Experimental Psychology: Human Learning and Memory, 3*, 233-245.

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

Wilhelm, S., McNally, R. J., Baer, L., & Florin, I. (1996). Directed forgetting in obsessive-compulsive disorder. *Behaviour Research and Therapy*, *8*, 633-641.

Williams, C. C., & Zacks, R. T. (2001). Is retrieval-induced forgetting an inhibitory process? *American Journal of Psychology*, *3*, 329-354.

Williams, L. M. (1994). Recall of childhood trauma: A prospective study of women's memories of childhood sexual abuse. *Journal of Consulting and Clinical Psychology*, *62*, 1167-1176.

Wimber, M., Alink, A., Charest, I., Kriegeskorte, N., & Anderson, M. C. (2015). Retrieval induces adaptive forgetting of competing memories via cortical pattern suppression. *Nature Neuroscience*, *18*, 582-589.

Woodward, A. E. Jr., & Bjork, R. A. (1971). Forgetting and remembering in free recall: Intentional and unintentional. *Journal of Experimental Psychology*, *89*, 109-116.

Woodward, A. E., Jr., Bjork, R. A., & Jongeward, R. H., Jr. (1973). Recall and recognition as a function of primary rehearsal. *Journal of Verbal Learning and Verbal Behavio*r, *12*, 608-617.

Yehuda, R., Keefe, R. S. E., Harvey, P. D., Levengood, R. A., Gerber, D. K., Geni, J., & Siever, L. J. (1995). Learning and memory in combat veterans with post traumatic stress disorders. *American Journal of Psychiatry*, *152*, 137-139.

Zacks, R. T., & Hasher, L. (1994). Directed ignoring: Inhibitory regulation of working memory. In D. Dagenbach & T. H. Carr (Eds.), *Inhibitory processes in attention, memory, and language* (pp. 241-263). San Diego: Academic Press.