Changing Unwanted Memories

A Consideration of Three Methods

Kevin van Schie

ISBN: 978-90-393-6906-7

Cover design: Evelien Jagtman
Printing: Ridderprint BV

Changing Unwanted Memories

A Consideration of Three Methods

Verandering van Ongewenste Herinneringen

Een Beschouwing van Drie Methoden

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit Utrecht op gezag van de rector magnificus, prof. dr. G.J. van der Zwaan, ingevolge het besluit van het college voor promoties in het openbaar te verdedigen op vrijdag 26 januari 2018 des middags te 12.45 uur

door

Kevin van Schie

geboren op 27 maart 1987 te Heemstede

Promotoren: Prof. dr. M. A. van den Hout
Prof. dr. I. M. Engelhard

Dit proefschrift werd (mede) mogelijk gemaakt met financiële steun van een TOP-subsidie (dossiernummer: 40-00812-98-12030) van de Nederlandse organisatie voor gezondheidsonderzoek en zorginnovatie (ZonMw).

“The truth is rarely pure and never simple.”

— Oscar Wilde, The Importance of Being Earnest

Table of Contents

CHAPTER 1

GENERAL INTRODUCTION

Changing Unwanted Memories

A Consideration of Three Methods

The majority of people experience a traumatic event at some point in their life (de Vries & Olff, 2009). This is an event in which an individual is exposed to actual or threatened death, serious injury, or sexual violence. Exposure to such an event can be via direct or indirect experience, witnessing the event, or learning about the event from someone else (DSM-5; American Psychiatric Association, 2013). While the majority of victims are able to cope with the event, a substantial minority (ranging between 5.6% and 22%) develop posttraumatic stress disorder (PTSD) in the months following the event (e.g., deRoon-Cassini, Mancini, Rusch, & Bonanno, 2010; Frans, Rimmö, Åberg, & Fredrikson, 2005; Kessler et al., 2005). This disorder is characterized by a number of debilitating symptoms that include persistent re-experiencing of the trauma, avoidance of trauma-related reminders, negative thoughts, and increased arousal or reactivity (DSM-5; American Psychiatric Association, 2013). Given post-traumatic stress disorder's prevalence, the disease already constitutes a heavy burden on society (e.g., productivity loss, healthcare costs) and on the individual (e.g., increased risks for suicide, school drop-out, and unemployment; Kessler, 2000). It has been proposed that this might increase in the coming years with increasing urbanity (Galea, Uddin, & Koenen, 2011) and increased (threat for) terrorism (Neria, Nandi, & Galea, 2008).

In clinical practice, PTSD can be treated with cognitive behavioral therapy (CBT), which is an evidence-based and generally considered efficacious treatment for this disorder[1] (Deacon & Abramowitz, 2004; Olatunji, Cisler, & Deacon, 2010). A frequently applied intervention in CBT is repeated confrontation with stimuli related to the traumatic event in real life (i.e., in vivo exposure) or in imagination (i.e., imaginal exposure) for prolonged periods of time (Arch & Craske, 2009; Rothbaum, Meadows, Resick, & Foy, 2000) to induce inhibitory learning (Craske, Treanor, Conway, Zbozinek, & Vervliet, 2014). However, research suggests that a significant number of PTSD patients may fail to improve (i.e., treatment-resistance) or show relapse after successful psychological treatment (Durham, Higgins, Chambers, Swan, & Dow, 2012; Hofmann & Smits, 2008; Loerinc et al., 2015). Moreover, some patients with PTSD may specifically be unwilling to undergo exposure, because they are reluctant to be confronted to what they fear most for longer periods of time (Arntz, Tiesema, & Kindt, 2007).

[1] In DSM-IV-TR PTSD was classified as an anxiety disorder (American Psychiatric Association, 2000). In DSM-V PTSD is classified as a trauma and stressor-related disorder (American Psychiatric Association, 2013). As a result, papers on (treatment of) anxiety disorders may also include PTSD.

In light of this, a number of scientists made a call for a focus on *why* therapies and interventions work, and not simply on *whether* they work (Elsey & Kindt, 2017; Engelhard, 2012; Kazdin, 2001; McNally, 2007; van den Hout, Engelhard, & McNally, 2017). Uncovering the mechanisms underlying effective interventions in current treatments could transform present-day clinical science from a science that describes phenomena to a science that *explains* them (van den Hout et al., 2017). With this knowledge psychological treatments may be optimized (e.g., by adjusting treatment to individual differences). Simultaneous with putting effort in optimizing current treatments, some scientists have argued that development of new treatments by borrowing insights from neighboring fields (e.g., neuroscience), and scrutinously testing their working mechanisms may be an additional fruitful approach (Holmes, Craske, & Graybiel 2014).

One way to examine why existing or new interventions work is by taking an experimental approach to psychopathology. Experimental psychopathology involves manipulation of variables to induce time-limited analogous features of psychopathology in people (generally without known psychopathology), or the subsequent elimination of those induced features (e.g., Philippot & Hermans, 2006; van den Hout et al., 2017; Olatunji, Leen-Feldner, Feldner, & Forsyth, 2007; Vervliet & Raes, 2013; Zvolensky, Lejuez, Stuart, & Curtin, 2001). In line with experiments in other areas of research, the goal is to identify causal processes by means of an experimental method (Srinagesh, 2006). Similarly, experimental psychopathology research is usually performed in well-controlled lab environments, which allows for the precise manipulation of variables of interest. Its ultimate goal is to better understand the origin, nature, and treatment of psychopathology.

In this thesis an experimental psychopathology approach is applied to examine three different interventions that target unwanted memories: (1) taxing working memory by dual-tasks, (2) memory disruption and updating during reconsolidation via new learning, and (3) retrieval suppression. Some of these interventions are already used in the psychological treatment of patients with PTSD; others may be used in the future to fulfill that role. The general introduction continues with presenting these three lines of enquiry. Each section focusses on one of these three interventions. Finally, the introduction concludes with an outline of the thesis.

Taxing Working Memory by Dual-Task Interventions

The first intervention to target unwanted, aversive memories is by deliberately recalling the “hotspot” of an aversive memory, while simultaneously making horizontal eye movements. This intervention was discovered by chance by Francine Shapiro while she was walking in a park (Shapiro, 1989). She experienced that recurrent, disturbing thoughts seemed to disappear as a result of making involuntary multi-saccadic eye movements (Rosen, 1995). From this experience she developed Eye Movement Desensitization and Reprocessing (EMDR); a multiphase therapy for the treatment of post-traumatic stress disorder. The eye movement intervention figures prominently in EMDR; patients are instructed to repeatedly recall a traumatic memory while simultaneously making horizontal eye movements by following a therapist's finger. After making eye movements patients report affect, physiological states, and cognitive insights. This phase concludes when the patient reports no more distress for the traumatic memory.

Initially, the therapy had an obscure theoretical foundation and Shapiro (1995) stated that with EMDR the nervous system is rebalanced and that dysfunctionally locked information in the nervous system is shifted (Engelhard, 2012). Perhaps not surprisingly EMDR was received with great skepticism by the scientific community (e.g., Herbert et al., 2000; McNally, 1999; Muris & Merckelbach, 1999; Rosen, Lohr, McNally & Herbert, 1998; see Engelhard, 2012 for a brief overview). Because of the initial paucity of evidence a critical stance seemed more than understandable. However, slowly after its inception, evidence started to accumulate for EMDR. A first piece of evidence was provided by a randomized clinical trial in which EMDR was compared to waitlist control in treating PTSD. Patients treated with EMDR, but not waitlist controls, reported less anxiety and other PTSD associated symptoms (Wilson, Becker & Tinker, 1995). Subsequently, a number of meta-analyses showed that EMDR was equally effective in treating PTSD compared to CBT, which at that time was used to treat PTSD (Bisson et al., 2007; Bradley, Greene, Russ, Dutra, & Westen, 2005; Seidler & Wagner, 2006; van Etten & Taylor 1998). At this point, EMDR, together with CBT, is considered an evidence-based and first-choice treatment for PTSD (American Psychiatric Association, 2004; Dutch National Steering Committee Guidelines for Mental Health Care, 2003; National Institute for Clinical Excellence, 2011).

Yet, some controversy remained; questions arose whether EMDR's working mechanism might be nothing more than (imaginal) exposure (e.g., Davidson & Parker, 2001; Lee & Drummond, 2008; Rogers & Silver, 2002). Some considered making eye movements an unnecessary therapeutic element (e.g., Lohr, Tolin, & Lilienfeld, 1999). However, a comprehensive meta-analysis by Lee & Cuijpers (2013), taking into account clinical studies and analogue studies, showed that eye movements do add to EMDR's effect; thereby showing that EMDR is more than exposure, and that there should be a different mechanism of action.

So, how can the effects of these peculiar eye movements be explained? Working memory theory, which is grounded in cognitive psychology, may be able to explain the effects of making eye movements (e.g., Andrade Kavanagh, & Baddeley, 1997; Engelhard & van den Hout, 2012; Gunter & Bodner, 2008). According to working memory theory the effectiveness of eye movements in EMDR might be explained by dual taxation of limited resources of working memory (Andrade et al., 1997; Gunter & Bodner, 2008). It is hypothesized that simultaneously making eye movements with memory recall results in competition for these working memory resources, which impedes memory retrieval. This subsequently reduces memory vividness and/or emotionality. It is suggested that after this process the temporarily labile memory is again stored into long-term memory, and that during future recalls the changed memory will be retrieved (van den Hout & Engelhard, 2012). This mechanism has largely been tested with a laboratory model of EMDR. In this model a participant recalls two aversive autobiographical memories. One of these is recalled while simultaneously making horizontal eye movements by following a moving dot on screen (i.e., Recall+EM). The other memory is recalled without making these eye movements (i.e., Recall Only), and serves as a condition that controls for mere exposure to the memory.

By now, numerous (lab) studies with healthy participants have shown that making eye movements simultaneously with recall reduces self-reported ratings of vividness and/or emotionality compared to recall without eye movements. Reductions in vividness and/or emotionality have been observed for different types of memories, such as unpleasant autobiographical memories (e.g., Andrade et al., 1997; Engelhard, van Uijen, & van den Hout, 2010b; Kavanagh, Freese, Andrade, & May, 2001; Maxfield, Melnyk, & Hayman, 2008), neutral and negative pictures (Andrade et al., 1997, van den Hout, Bartelski, & Engelhard, 2013), prospective memories (e.g., Engelhard, van den Hout, Janssen, & van der Beek, 2010a), positive memories (van den Hout, Muris, Salemink, & Kindt, 2001), and memories acquired during fear conditioning (Leer, Engelhard, Altink, & van den Hout, 2013; Leer et al., 2017). There is some evidence that dual-tasks also affect more objective indices of memory change, such as memory accessibility (van den Hout et al., 2013), startle eye-blink (Engelhard et al, 2010b), and heart rate (Kearns & Engelhard, 2015). However, there is little evidence that dual-task interventions affect lab-analogues of posttraumatic stress symptoms (e.g., intrusions or avoidance). Moreover, how changes in vividness and/or emotionality eventually cascade into changes in the manifestation of PTSD remains unclear (Gunter & Bodner, 2009). One debilitating core symptom of PTSD that may be targeted with dual-task intervention is the re-experiencing of a traumatic memory. Research on patients with PTSD has shown that the majority of intrusive memories are related to the memory hotspot (Holmes, Grey, & Young, 2005; Grey & Holmes, 2008). Since dual-tasks are targeted at the memory hotspot and reduce the hotspot's vividness and/or emotionality, this may subsequently reduce the number of intrusive memories, by truncating the hotspot's excitable nature.

There is a lively debate in the literature on whether visual processing specifically increases the effect of working memory loads on visual imagery (e.g., Andrade et al., 1997; Kavanagh et al., 2001), or whether load per se (Gunter & Bodner, 2008) is the most important factor. There is evidence for both accounts; some studies (only) show effects for visuospatial dual-tasks (e.g., Andrade, Pears, May, & Kavanagh, 2012; May, Andrade, Panabokke, & Kavanagh, 2010), while others show it is not modality per se, but general WM taxation (e.g., Engelhard et al., 2011; Tadmor, McNally, and Engelhard, 2017; van den Hout et al., 2010). This latter account proposes that a dual-task does not necessarily need to be visuospatial in nature; rather a dual-task simply needs to be sufficiently taxing on WM, regardless of modality. However, one account may not rule out the other. One study has shown general effects of dual-task on visual and auditory images (Baddeley & Andrade, 2000). Another study showed evidence for modality specific effects; a visual dual-task reduced ratings of vividness and/or emotionality more when the memory was also visual in nature compared to auditory. For the auditory dual-task this effect was mirrored. However, compared to a control condition performing either dual-task intervention produced the largest change in memory ratings (Kemps & Tiggeman, 2007). These effects may be explained by how modality specific dual-tasks differentially tax components of working memory (Matthijssen, van Schie, & van den Hout, 2017).

According to working memory theory, the effectiveness of the dual-task is not solely determined by the dual-task's modality. After all, competition in working memory is dependent on three general elements: (1) taxation of the dual-task (which includes the modality), (2) an individual's working memory capacity, and (3) taxation of the memory. From this follows that individuals with a large working memory capacity (compared to individuals with a small working memory capacity) may be able to perform both tasks without there being (much) competition. Without competition, there is no impediment on memory recall, and therefore no change in memory vividness and/or emotionality. A number of studies, indeed, seem to suggest that working memory capacity is an important factor to take into account in research on dual-tasks; higher scores on tasks that measure working memory capacity correlate negatively with decreases in memory vividness and/or emotionality (Engelhard et al., 2010b; Gunter & Bodner, 2008; van den Hout et al., 2010; van den Hout et al., 2011).

Creating competition in individuals with a relatively large working memory capacity may be easily done by increasing taxation of a dual-task by increasing its complexity, for instance by speeding up the eye movements (Maxfield et al., 2008), or counting backwards in steps of 7 compared to steps of 2 (Engelhard et al., 2011; van den Hout et al., 2010). However, the relationship between a dual-task and its effects on vividness and emotionality is not necessarily linear in nature (Gunter & Bodner, 2008). According to working memory theory competition needs to be attained with a dual-task that is just right. It should not be too easy to do, which would leave many resources for memory recall, but also should not be too hard, which would impede memory recall too much. Preliminary evidence suggests there may be truth in this inverted U-curve shape (though only for emotionality, and not for vividness ratings; Engelhard et al., 2011). When applied to working memory capacity, this means that how taxing a dual-task is, needs to be adjusted to an individual's working memory capacity. Low (compared to high) taxing dual-tasks should reduce memory ratings more in individuals with small working memory capacity, while high (compared to low) taxing dual-tasks should be more effective in individuals with large working memory capacity.

From the inverted U-curve shape hypothesis follows a second prediction; one that is related to adjusting dual-tasks to how taxing a memory is. It is argued that a memory's taxation depends on how vivid the memory originally is. Memories high in vividness presumably tax working memory more than memories low in vividness (Baddeley & Andrade, 2000). It follows that if a memory is high in vividness, a relatively low taxing dual-task may not result in sufficient competition, and therefore precludes changes in the memory. This means that memories high in vividness are best targeted with a high taxing dual-task, while memories low in vividness are best targeted with low taxing dual-task.

In conclusion, dual-tasks have proven to be effective in changing self-reported ratings of vividness and/or emotionality, but there is room for improvement. Adjusting the dual-task intervention to individual differences (e.g., memory vividness and working memory capacity) may increase its effectiveness. Furthermore, there is a need for studies that move beyond the realm of self-reported vividness and emotionality, to methods that objectively measure memory accessibility, or measure lab-analogs of PTSD symptoms (e.g., intrusions).

Memory Disruption and Updating During Reconsolidation

Since the introduction of the perseveration-consolidation hypothesis by Müller and Pilzecker (1900), it has been well accepted that memories are initially in an active, unstable state before they are fixed in long-term memory in and inactive and stable state ('consolidation'). Once consolidated (from the Latin 'consolidare' meaning 'to make firm') memories supposedly were resistant and insensitive to change (e.g., McGaugh, 2000; Dudai, 2004); a belief that still stands firm today in the general public (Simons & Chabris, 2011). Though encompassed by one term, consolidation actually refers to two types of processes. The first type is synaptic consolidation and derives its name from processes that take place in the local nodes of neuronal circuits (i.e., the synapses). Synaptic consolidation is accomplished with minutes to hours after learning. The second type, systems consolidation, is a slower type of consolidation and occurs over a period of weeks to many years. It is believed that during systems consolidation there is a reorganization of brain systems; the memory trace moves to new locations and becomes independent of memory circuits that were used for acquisition (Dudai, 2004; Dudai & Morris, 2000). Either process reflects progressive unidirectional memory stabilization that, by definition, takes place postacquisition, and that for every item in memory consolidation starts and ends only once.

Consolidation research experienced its heydays in the 20th century and iconoclastic studies that did not fit the dogma were considered anomalies (e.g., Misanin, Miller, & Lewis, 1968; Schneider & Sherman, 1968). In general, these studies showed that reactivating a consolidated memory in rats followed by electroconvulsive shock (i.e., an amnesic treatment) resulted in memory loss, which at that time was coined cue-dependent amnesia (Lewis & Bregman, 1972). These results were perpendicular to the unidirectional stabilization view of consolidation, which stated that it should be impossible to interfere with memories after their consolidation (McGaugh, 1966). The introduction of the concept of *re*consolidation provided reconciliation for the anomalous findings and for consolidation theory (Nader, Schafe, & LeDoux, 2000; Sara, 2000). The process of reconsolidation, as the name implies, is largely a repetition of the process of consolidation, only reconsolidation takes place after a consolidated memory is retrieved. Under certain conditions retrieval of consolidated memories may return these memory traces (i.e., engrams) into active and unstable states, which may then be disrupted or updated before being reconsolidated (for an overview see Ågren, 2014; Besnard, 2012; Elsey & Kindt, 2017; Schwabe, Nader, & Pruessner, 2014). It is assumed that consolidated memories can change because updating a memory to retain its relevance in changing environments would be adaptive for the individual (Lee, 2009).

Memory disruption and/or updating in humans is, in experimental studies, generally achieved by applying either a pharmacological or behavioral intervention after memory reactivation. A standard experiment usually involves a three-day design with memory acquisition on Day 1, reactivation and/or intervention on Day 2, and test on Day 3.

Crucially, only some groups reactivate the memory on Day 2, and some receive an intervention. Memory disruption or updating is only expected in the group that reactivated the Day1 memory and subsequently received an intervention on Day 2. By *pharmacological* means in humans the memory can be disrupted by, for instance, the use of the drug propranolol. It is assumed that propranolol inhibits protein synthesis that is necessary for the re-storage of the original memory, thereby creating amnesia for (parts of) the original memory (e.g., Kindt, Soeter, & Vervliet, 2009; Sevenster, Beckers, & Kindt, 2012; 2013; Soeter & Kindt, 2012). *Behavioral* manipulations primarily employ novel learning or interference interventions during reconsolidation to update or disrupt reactivated memories (e.g., Ågren, 2014; Forcato et al., 2007; Hupbach, Gomez, Hardt, & Nadel, 2007; James et al., 2015; Schwabe & Wolf, 2009; Schiller et al., 2010; Wichert, Wolf, & Schwabe, 2011, 2013a, 2013b).

The potential to change memories when they are reconsolidated is especially exciting for clinical practice. Arguably, this is relevant and interesting for different types of psychopathology in which unwanted and maladaptive memories take a central place. PTSD is just one of the potential candidates next to, for instance, anxiety disorders and addiction. How could this be applied in clinical practice? Roughly one could envision patients recalling (i.e., reactivating) their memory (hotspot) followed by an appropriate behavioral or pharmacological intervention that (gradually) modifies the maladaptive memory or memories that are at the core of their disorder (Beckers & Kindt, 2017; Kindt & van Emmerik, 2016; Lane, Ryan, Nadel, & Greenberg, 2015). Though, this idea holds potential for clinical practice, up until this point, there are only a limited number of studies describing the application of interventions during reconsolidation in patients with PTSD (e.g., Brunet et al., 2008; Kindt & van Emmerik, 2016; Wood et al., 2015; though for an application in spider phobics, see Soeter & Kindt, 2015). Currently, no standardized clinical guidelines are available or under development (for as far as the author is aware). However note that it is also possible that psychological treatments work because they already capitalize on this process; in many therapies a memory is reactivated followed by an intervention (e.g., exposure, imagery rescripting; Foa & Kozak, 1986).

Given the potential clinical application, it is important to fully understand the reconsolidation process and thus when memory change can occur (e.g., Dudai, 2006). Presently, several so-called boundary conditions have been identified which preclude memories from being changed during reconsolidation (e.g., Schwabe et al., 2014). For instance, older and stronger memories seem more resistant to modification compared to younger and weaker memories (Eisenberg, Kobilo, Berman, & Dudai, 2003; Wichert et al., 2011). Some have suggested that the memory needs to be directly (instead of indirectly) reactivated (Dębiec, Doyère, Nader, & LeDoux 2006), or that new information needs to be present (i.e., prediction error; Sevenster et al., 2012, 2013). Not only are there a number of factors that restrict initiating (memory change during) reconsolidation, there have also been problems with replicating reconsolidation research (e.g., Bos, Beckers, & Kindt, 2014; Nader & Einarsson, 2010; Schroyens, Beckers, & Kindt, 2017; Thome et al., 2016).

If behavioral interventions are to be implemented in the future in psychological treatment, it is pertinent to investigate when memory change during reconsolidation is possible and when it is not. Using laboratory designs it is possible to carefully examine these conditions. It is, however, pertinent that these designs start to mimic conditions that are associated with the psychopathology they intend to treat. One or multiple memories at the core of PTSD are likely strongly consolidated, because stress hormones are released at the time of traumatic event (and while re-experiencing the event; Pitman, 1989), and because the traumatic memory is intrusive and repetitive in nature (Hackmann, Ehlers, Speckens, & Clark, 2004). This necessitates that laboratory designs take memory strength (or age) into account, and perhaps also levels of anxiety, when studying memory change during reconsolidation. Only then can we increase our understanding of the reconsolidation process, and can we be successful in translating these findings to clinical practice of treating psychopathology.

To conclude, memory disruption and/or updating by applying a behavioral intervention after memory reactivation seems promising for changing memories in clinical practice. However, given that sometimes the effect does not replicate (in the lab) and that it is restricted by a number of factors, we need to know more about the replicability and the exact parameters of the reconsolidation process.

Retrieval Suppression of Unwanted Intrusive Memories

In daily life, people regularly encounter objects, people or situations that automatically trigger memories of earlier aversive events. Thinking of these events may unsettle us and consequently people try to limit the time they spend thinking of these memories. A way to do that is by preventing or interrupting the reflexive retrieval of unwelcome memories when confronted reminders, and thereby limiting their duration in mental awareness. This is an ability known as retrieval suppression or retrieval stopping. Being able to suppress unwanted memories may be especially important after experiencing a traumatic event particularly given that many cognitive theories of PTSD state that traumatic memories are often triggered by external cues; stimuli that have become associated with the traumatic event (e.g., Brewin, 2001; Dalgleish, 2004; Ehlers & Clark, 2000; Ehlers et al., 2002). The possibility of stopping retrieval after encountering these reminders may be a way to keeps one's mind free from unwanted memories.

The logic behind retrieval suppression is relatively straightforward and is based on the idea that, at times, a prepotent mental or physical activity needs to be stopped or overridden because it is or has become inappropriate (Anderson & Weaver, 2009). This is best exemplified by a frequently mentioned, yet elegant, example involving a falling plant and one of my co-authors (see Anderson & Levy, 2009). He knocked a potted plant off a windowsill one evening, and reflexively wanted to catch the falling plant. However, at the moment he almost caught the plant, he stopped his action realizing that the plant falling is a cactus. Without these control mechanisms some argue, "we would be slaves to habit and reflex" (Anderson & Huddleston, 2012). Stopping motor responses has typically been investigated with tasks such as the stop-signal paradigm (Logan, 1994) or the Go/No-Go paradigm. In the latter task, people press a button when they see a letter appear on screen (i.e., go stimuli), except when the X appears (i.e., no-go stimulus). When this happens they need to suppress pressing the button. How well one is able to withhold these button presses constitutes their inhibitory motor control.

Similar to stopping prepotent motor actions, people also attempt to stop prepotent memories (i.e., stopping the retrieval process). Retrieval stopping seems to build upon similar, if not the same, mechanisms of motor stopping, and it is supposed that both are part of a broader ability to override prepotent responses (Anderson & Levy, 2009; Anderson & Huddleston, 2012; Depue, Orr, Smolker, Naaz, & Banich, 2015; for a critical view see Noreen & Macleod, 2015). One paradigm to study this is the Think/No-think (TNT) paradigm (Anderson & Green, 2001, Anderson et al., 2004). This paradigm can be used to study how reflexive retrieval of unwelcome memories might be prevented or interrupted by voluntarily inhibitory control. It is modelled after the Go/No-Go paradigm and mimics situations in which we encounter reminders that make us think about our own "mental cacti" (i.e., memories we would rather not think about; e.g., Anderson & Huddleston, 2012). In the TNT paradigm, people first study cue-target word pairs till a set criterion is reached. Next, when they see a cue they are instructed to either retrieve the target memory (i.e., Think trials), or to prevent or exclude it from awareness (i.e. No-Think trials). A third set of items in this phase are neither recalled, nor suppressed (i.e. Baseline items), which controls for effects of passage of time on memory. Typically during the TNT phase each Think and No-Think cue is repeated a number of times. Afterwards, all pairs are tested on an unexpected memory test to determine how retrieval suppression influenced later retention of the excluded memories.

A large body of research shows that suppressing unwanted memories reduces recall of No-Think items on the final test compared to Baseline items. This effect is known as suppression-induced forgetting (Anderson et al., 2004; Anderson & Green, 2001; Anderson & Hanslmayr, 2014; Anderson & Huddleston, 2012). Suppression-induced forgetting arises with many types of stimuli, such word pairs (e.g., Anderson et al., 2001; 2004), face–scene pairs (Depue, Curran, & Banich, 2007), face–word pairs (Hanslmayr, Leipold, Pastötter, & Bäuml, 2009), word–object pairs (Gagnepain, Henson, & Anderson, 2014; Kim & Yi, 2013), object–scene pairs (Catarino, Küpper, Werner-Seidler, Dalgleish, & Anderson, 2015; Küpper, Benoit, Dalgleish, & Anderson, 2014) and pairs comprising words and nonsense shapes (Hart & Schooler, 2012), but also for autobiographical experiences especially for event details (Hu et al., 2015; Noreen & MacLeod, 2013; 2014; Stephens, Braid, & Hertel, 2013). There is some evidence suggesting that suppression of emotionally negative materials is possible (e.g. Depue, Banich, & Curran, 2006), yet there are also studies suggesting this is more difficult to do compared to other materials (e.g., Marx, Marshall, & Castro, 2008; Nørby, Lange, & Larsen, 2010).

After its development the TNT paradigm has been used primarily to study the effects of inhibitory control on later *voluntary* retrieval of those same memories. Recently, however, it has been adapted to study *involuntary* memory retrieval by adding a trial-by-trial judgment task during the TNT phase. After each trial participants judge whether the target entered awareness (Benoit, Hulbert, Huddleston, & Anderson, 2014; Gagnepain, Hulbert, & Anderson, 2017; Levy & Anderson, 2012). Adding this task allows for examining the assumption that a cue triggers automatic retrieval of a target unless effort is made to proactively prevent or reactively interrupt this process. Studies have shown that retrieval suppression, indeed, is able to downregulate the number of intrusions; with repeated suppression attempts the number of intrusions decreases (e.g., Benoit et al., 2014; Gagnepain et al., 2017). Moreover, steeper reductions in intrusions predicted greater suppression-induced forgetting on the final memory test. This suggests that that people can recruit inhibitory control to down-regulate their intrusive memories, and that controlling these intrusive memories leads to forgetting.

The extent to which people are able regulate their memory, as measured either by how it affects voluntary memory recall or involuntary re-experiencing, may be dependent on a number of factors. We will discuss three that are relevant for the current thesis. One of these factors it the strategy people employ to suppress retrieval. Though, there are numerous approaches a person could take not thinking about unwanted memories (Levy & Anderson, 2008), approaches are usually identified as belonging to one of two general types: thought substitution or direct suppression strategies (e.g., Bergström, de Fockert, & Richardson-Klavehn, 2009b). When participants are instructed to use thought substitution they are generally provided with an alternative target word to think of in No-Think trials (e.g., Hertel & Hayes, 2015; Hertel & McDaniel, 2010; Joormann, Hertel, LeMoult, & Gotlib, 2009). When given direct suppression instructions participants are asked to avoid thinking of the target memory without replacing it with anything else; participants are asked to block it out when the unwanted memory happens to come to mind. Some have argued that thought substitution generally increases the forgetting effect (e.g., Hertel & Calcaterra, 2005; Hotta & Kawaguchi, 2009). Interestingly, these two strategies to retrieval suppression are supported by different neurocognitive mechanisms (Benoit & Anderson, 2012). During direct suppression the prefrontal cortex is engaged, which disengages the hippocampus limiting episodic retrieval, while during thought substitution the hippocampus is actually engaged. These neurobiological studies suggests that inhibitory control is more strongly linked to a direct suppression approach to motivated forgetting, then to thought substitution (Bergström et al., 2009b; Anderson & Hanslmayr, 2014).

Another factor that may influence whether people are able to successfully forget unwanted memories is the extent to which inhibitory control can be engaged or is required (Levy & Anderson, 2008). Thus, individuals or groups of individuals that show deficits in inhibitory control should be less able to forget unwanted memories. Indeed, forgetting is affected for groups that have deficient inhibitory control: young children perform worse compared to adults (Paz-Alonso, Ghetti, Matlen, Anderson, & Bunge, 2009), and older adults are less able to forget memories compared to younger adults (Anderson, Reinholz, Kuhl, & Mayr, 2011). Furthermore, clinical populations that are hypothesized to be deficient in inhibitory control show smaller suppression induced forgetting effects compared to healthy control groups. This was evident for patients with PTSD that suppressed pictures of negative scenes (Catarino et al., 2015), patients with attention deficit hyperactivity disorder that also suppressed pictures of negative scenes (Depue, Burgess, Willcutt, Ruzic, & Banich, 2010), and patients with depression that suppressed words of neutral, negative, and positive valence (Joormann, Hertel, Brozovich, & Gotlib, 2005; Joormann et al., 2009).

Given that deficits in inhibitory control in patient groups affects forgetting, task or state-dependent strains on cognitive control might also be associated with the size of suppression induced forgetting effect. For instance, Lee, Lee, and Tsai (2007) argue that trial duration plays a role, and that for longer duration trials the suppression induced forgetting effect should be smaller than for short duration trials. On longer duration trials there is more opportunity for control to fail, and for the unwanted target to intrude into awareness. This is indeed what they observed on the final test, though this is an effect that requires further replication, because of their between-group manipulation of duration. Moreover, they did not quantify whether smaller suppression induced forgetting effects were indeed the result of more intrusions following suppression attempts on long duration trials. Nevertheless, this study foreshadows the importance of trial duration, and accordingly the need for cognitive control. The idea that cognitive control, and failures therein, might depend on the period it is required is important, because trial duration within the TNT paradigm has varied substantially: from 2 seconds (Bergström, de Fockert, & Richardson-Klavehn, 2009a) to 8 seconds (Wang, Cao, Zhu, Cai, & Wu, 2015).

A third and final factor that may influence the size of the suppression induced forgetting effect is the valence of memories people are attempting to suppress. It is possible that negative memories may be harder to forget, because of their intrusive nature. In line with this idea, studies have found smaller effects for negative materials (e.g., Chen et al., 2012; Marx et al., 2008; Nørby, Lange, & Larsen, 2010), or found no differences between materials of different valence (Gagnepain et al., 2017; Murray, Muscatell, & Kensinger, 2011). Alternatively, it could also be possible that that people are more motivated to forget these unwanted memories, because they are unpleasant by definition. In agreement with this are studies that found bigger suppression induced forgetting effects for emotionally negative compared to neutral or positive materials (e.g., Depue, Banich, & Curran, 2006; Joormann et al., 2005; Lambert, Good, & Kirk, 2010). The possibility to suppress emotional memories, specifically negative memories, is especially relevant for generalization to naturalistic situations in which people experience aversive events, of which traumatic events are an extreme example.

To summarize, retrieval suppression seems able to prevent or interrupt the reflexive retrieval of unwelcome memories when confronted reminders, which consequently limits their intrusive nature and time in mental awareness. The extent to which suppression retrieval can be engaged may be dependent on a number of factors, such as the adopted strategy (thought substitution vs. direct suppression), the time that suppression is required, and the valence of what an individual wants to suppress (i.e., whether the memory is neutral or negative).

Outline of this thesis

The current thesis comprises 9 research chapters (chapters 2 to 10), which are divided over three sections. Each section focuses on one of the abovementioned interventions to modify unwanted memories. Hence, three different interventions will be discussed in this thesis. In the first and largest section we discuss a number of studies that investigate how dual-task interventions during memory retrieval change memory vividness, emotionality, and accessibility, but also intrusions of the memory. In general, making eye movements, as a lab-based derivative from EMDR therapy, is used in these chapters, though other dual-tasks, such as counting backwards and playing Tetris, are also used.

In the first chapter of section 1, **chapter 2**, we test whether a dual-tasking intervention reduces objectively measured memory accessibility, next to frequently self-report measures of vividness and emotionality. Because traumatic memories are frequently triggered by external cues, we adopted a dual-task procedure in this study where memory recall is cued by a reminder that was previously associated with the memory. In **chapters 3** and **4** we aimed to gain further insights in the working mechanisms of dual-task intervention, and we were specifically interested in testing whether highly taxing dual-tasks are more effective for individuals with large working memory capacity (compared to those with small working memory capacity), and memories high in vividness (compared to memories low in vividness). With the adjustment to individual differences we aimed to increase the intervention's effectiveness. In the following two chapters we add an intrusion diary to the dual-task paradigm, which allows for measuring the number of intrusions people experience after performing a dual-task intervention. In **chapter 5** we test whether a dual-task intervention during memory recall truncates the memory hotspot and also reduces the number of intrusive memories. Moreover, we varied the modality of the dual-task intervention and used a visuospatial (i.e., eye movements) and verbal (i.e., counting backwards) dual-task intervention to test whether dual-tasks reduce intrusive re-experiencing. A secondary goal was to examine whether intrusion modulation is moderated by the modality of the dual-task. In **chapter 6**, we test whether deliberate recall of the memory is a necessary element of the dual-task intervention. In this study we compare a dual-task intervention (i.e., memory recall + Tetris) to a single-task intervention (i.e., Tetris only).

Section 2 (**chapters 7 and 8**) focuses on a different intervention, namely memory disruption and updating during reconsolidation. In **chapter 7**, we perform a replication of Wichert et al., (2013a) and examine whether episodic memory can be disrupted and/or updated by performing a (behavioral) novel learning intervention following reactivation of the original memory. In **chapter 8**, we further test whether the strength of such a novel learning intervention needs to be adjusted, when the original memory is strong as well.

Section 3 of this thesis addresses the third and final intervention: Retrieval suppression of unwanted intrusive memories. In **chapter 9** we test whether suppression (and hence forgetting) of unwanted memories is dependent on whether these memories are neutrally or negatively valenced, and if this depends on the strategy adopted. In **chapter 10**, we examine how sustained inhibitory control affects intrusive memories in the Think/No-Think paradigm. The need for sustained control was varied by providing participants with short and long duration trials.

The thesis will conclude with a summary and general discussion (**chapter 11**).

SECTION 1

TAXING WORKING MEMORY BY DUAL-TASK INTERVENTIONS

CHAPTER 2

TAXING WORKING MEMORY DURING RETRIEVAL OF EMOTIONAL MEMORIES DOES NOT REDUCE MEMORY ACCESSIBILITY WHEN CUED WITH REMINDERS

K. VAN SCHIE
I. M. ENGELHARD
M. A. VAN DEN HOUT

Frontiers in Psychiatry, *6*, 16 (2015)
doi:10.3389/FPSYT.2015.00016

Abstract

Earlier studies have shown that when individuals recall an emotional memory while simultaneously doing a demanding dual-task (e.g., playing Tetris, mental arithmetic, making eye movements), this reduces self-reported vividness and emotionality of the memory. These effects have been found up to one week later, but have largely been confined to self-report ratings. This study examined whether this dual-tasking intervention reduces memory performance (i.e., accessibility of emotional memories). Undergraduates (N = 60) studied word-image pairs and rated the retrieved image on vividness and emotionality when cued with the word. Then they viewed the cues and recalled the images with or without making eye movements. Finally, they re-rated the images on vividness and emotionality. Additionally, fragments from images from all conditions were presented and participants identified which fragment was paired earlier with which cue. Findings showed no effect of the dual-task manipulation on self-reported ratings and latency responses. Several possible explanations for the lack of effects are discussed, but the cued recall procedure in our experiment seems to explain the absence of effects best. The study demonstrates boundaries to the effects of the "dual-tasking" procedure.

Keywords: dual taxation, memory accessibility, reaction time task, working memory, visual imagery

Introduction

Eye Movement Desensitization and Reprocessing (EMDR) is an evidence-based treatment for posttraumatic stress disorder (American Psychiatric Association, 2004; Dutch National Steering Committee Guidelines for Mental Health Care, 2003; National Institute for Clinical Excellence, 2011). In EMDR patients are asked to recall traumatic memories while they simultaneously make eye movements (EM). Although EMDR's scientific and practical value was questioned at first (Herbert et al., 2000), EMDR has proven to be as effective as trauma-based cognitive-behavioral therapy (For meta-analyses, see e.g., Bisson et al., 2007; Seidler & Wagner, 2006).

For a long time, it has been controversial whether making EM while recalling the distressing memory added anything to the procedure (Rogers & Silver, 2002). A recent meta-analysis of clinical and laboratory studies on the role of EM has shown that EM during recall of negative memories do have additional effects (Lee & Cuijpers, 2013). For laboratory studies this additive effect is evident by larger reductions in subjective vividness and emotionality ratings of the distressing memory for recall+EM conditions, compared to a control condition in which participants merely recall the memory (i.e. recall only). Clinical studies have mainly found effects of EM on Subjective Units of Distress. There are some findings of EM effects on Subjective Units of Distress combined with symptom measures, such as the Impact of Events Scale (see Lee & Cuijpers, 2013).

The effectiveness of eye movements in EMDR can be explained by dual taxation of the limited resources of working memory (WM; e.g., Andrade et al., 1997). This dual taxation takes place when an individual recalls a distressing memory while also performing a secondary task, such as making EM, mental arithmetic, or drawing complex figures. When individuals perform the secondary task and simultaneously recall a memory, both tasks compete for scarce WM resources. During this competition, the distressing memory cannot be retrieved completely (i.e. gets blurred) and is stored as a blurred memory after this competition. As a consequence, the blurred memory will be retrieved during future recalls. Presently, a body of evidence supports the working memory hypothesis (e.g., Engelhard et al., 2011a; Gunter & Bodner, 2008; Maxfield et al., 2008; for review, see Van den Hout & Engelhard, 2012). The vast majority of studies have shown that dual taxation of working memory blurs *autobiographical* memories (see Lee & Cuijpers, 2013; Van den Hout & Engelhard, 2012); with a few exceptions, such as pictures (Andrade et al, 1997; Van den Hout, Eidhof, Verboom, Little, & Engelhard, 2014) and a film clip (Leer et al., 2013).

Most studies on the effects of recall+EM focused on changes in the subjective ratings of the experienced memory (e.g., Engelhard et al., 2011a; Smeets et al. 2012; Van den Hout et al., 2012). Therefore, the effects of eye movements in these studies could still –in part– be the effect of demand characteristics; inferences participants make based on what they think the researcher expects (Gunter & Bodner, 2008). Thus, it remains unclear whether changes in subjective ratings are the result of an experimental manipulation *per se* or of unconscious or conscious alterations in the participants' behavior to fit the hypothesis. Gunter and Bodner (2008) showed that simultaneous recall+EM reduced subjective ratings, while making EM *after* memory recall did not produce these reductions. This seems to forestall the conclusion that demand characteristics fulfil a large role in the effects of eye movements, because if they did, both conditions should yield similar results. Nevertheless, objective memory measures could preclude the demand characteristics account even further.

To date, however, little attention has been devoted to objectively measure memory accessibility. Yet WM theory predicts blurring of memories that are recalled under dual taxation conditions. This reduces its accessibility, because it is harder to access a blurred memory. Van den Hout et al. (2013a) attempted to objectively measure alterations in memory vividness with a reaction time (RT) task. In their experiment, participants studied two "neutral" pictures; though the pictures were not rated for emotional valence. One picture was followed by recall+EM and the other was followed by recall only. In the RT task participants performed old/new recognition on cut-outs taken from both pictures and cut-outs from never studied pictures. The rationale for using cut-outs was that comparing a picture cut-out with a blurred memory of that same picture takes more time, than a non-blurred memory. Van den Hout et al. (2013a) showed that in the Recall+EM condition –but not in the recall only condition– a reduction in vividness ratings was accompanied by an increase in RTs: participants were slower in deciding whether they had seen the fragments before. Therefore, RTs seem suitable as an objective behavioral measure to assess the effects of memory blurring by EM. Notably, in a different paper (Van den Hout et al., 2014) the question arose whether the allegedly neutral pictures from van den Hout (2013a) were truly neutral. As an annex, in the discussion of the later paper, the two pictures of van den Hout et al. (2013a) were rated and it was observed that the pictures were hedonically positive. It is not self-evident that the effect of reduced accessibility of emotionally positive stimuli generalizes to negative materials.

While traumatic memories may seem to occur out of the blue, they are often activated by environmental cues (Dalgleish, 2004). Therefore, the blurring of emotional memories needs to be placed in perspective. According to cognitive theories about posttraumatic stress disorder, traumatic memories are not contained in a vacuum: they can be activated by external cues. For instance, Brewin (2001) argued that certain cues reminiscent of the traumatic experience can activate memory presentations that are otherwise inaccessible. Ehlers and Clark (2000) make a comparable statement, namely that involuntary intrusive visual memories are triggered by stimuli that are temporally associated with the trauma, but are not strongly semantically related to the event. Through an associative learning process these stimuli and trauma become connected, and as a consequence the stimulus becomes a warning signal: a stimulus that signals imminent danger (Ehlers et al., 2002).

Though it is apparent that traumatic memories are not isolated memories, it is currently unclear whether the effects of EM (i.e., reductions in subjective ratings and in memory accessibility) will be found when a reminder cue is presented. On the one hand, it is easier to remember episodic information in cued recall than in non-cued recall (e.g., Tulving & Pearlstone, 1966). On the other hand, cues are frequently not uniquely encoded with one memory, which may make it difficult to recall a specific memory instantly. Furthermore, reduced subjective ratings do not necessarily have to co-occur with reduced accessibility, because declarative memory and memory for learning associations (i.e. conditioning) can be dissociated (Bechara et al., 1995). To illustrate this point, patients with posttraumatic stress disorder may display trouble with intentional recall of aspects of the trauma memory, while cue-driven triggers lead to intrusive re-experiencing of that memory (Ehlers & Clark, 2000).

Inaccessibility of episodic memories has frequently been studied with the Think/No-Think paradigm (Anderson & Green, 2001). In this task, participants first learn cue-target word pairs. They then repeatedly recall the target for "think items" when seeing the cue, or stop target retrieval for "no-think items". For one third of the items (baseline items), there is no recall or retrieval-stopping. Afterwards, memory for all pairs is assessed and generally a part of the cued no-think items –compared to baseline items– have become inaccessible (see Anderson & Huddleston, 2011 for a review). Evidently, the accessibility of cued episodic memories can be affected.

The main aim of the current study was to replicate and extend the study of van den Hout et al. (2013a) to negative memories. We tested whether recall+EM affect the accessibility of associative emotionally *negative* memory representations. Studies on the effects of dual taxation on memory typically use negative autobiographical memories. However, given that autobiographical memories are by their very nature hard to control and that positive effects have also been found with self-irrelevant memories for pictures shown in the laboratory (e.g., Van den Hout et al, 2013a), we decided to use memories of aversive pictures instead. We examined whether accessibility of emotional memories is affected when those memories are activated via *reminder cues*. We used these reminder cues as an experimental analogue of the often cued nature of trauma intrusions. Additionally, we attempted to *objectively* measure memory blurring with a response latency task. To achieve our aim, we adapted the frequently used Eye Movements Task by incorporating elements from the think/no-think paradigm and reasoned that associative accessibility, measured by RT following a reminder cue, should be reduced for recall+EM compared to a recall only or to no presentation control. Additionally, parallel to this RT reduction, we expect memory blurring in terms of reductions in self rated vividness and emotionality of the emotionally negative memory.

Ethics Statement

The research reported in this article involved healthy human participants, and did not utilize any invasive techniques, substance administration or psychological manipulations. It was conducted according to the principles expressed in the Declaration of Helsinki. The sample size was set before data collection and written informed consent of each participant was obtained. In giving consent, participants indicated to have read and to have agreed with both the rules regarding participation and proper (laboratory) behavior, and the researchers' commitments and privacy policy. They were also informed that they would be able to stop participating in the experiment whenever they wanted to do so. After consent, participants were randomly allocated to conditions and all gathered data were analyzed anonymously. Afterwards participants were debriefed.

Methods

Participants

Sixty-two undergraduates of Utrecht University (M = 22.12 years, SD = 3.16; 45 females, 17 males) participated for course credit or financial reimbursement. Two participants were excluded (one because of sudden illness, another because of EMDR knowledge prior to the experiment), resulting in a sample of 60 participants.

Materials

Words. In this paradigm, participants studied word-image association pairs that were divided over recall+EM, recall only, and control. For these pairs, 12 neutral Dutch words and two filler words (bean, caterpillar, chip, clock, gate, hawk, iron, mill, nail, plum, reptile, spleen, stamp, and wind) with moderate levels of arousal were selected from Moors et al. (2013; $M_{valence}$ = 4.08, SD = 0.15; $M_{arousal}$ = 3.59, SD = 0.41; words were rated on a scale from 1 *very negative/passive* to 7 *very positive/active*). Filler were used as buffers at the beginning and end of lists to avoid primacy and recency effects. Importantly, the three experimental conditions were matched on ratings of valence and arousal, as well as on ratings of word length, word frequency, power, and age of acquisition. The latter two referred to the extent to which a word is submissive/dominant, and to the estimated age a word was first learned.

We used neutral words as a model of cued trauma recall. Frequently, objects or situations that are associated with trauma –because of their temporal proximity– are of neutral valence (e.g., a bank in case of a bank robbery). A second reason to use neutral words instead of negative words was to avoid inter-pair associations as much as possible (i.e. associations other than those between the word and image of a pair). Because negative materials stem from a small number of categories (e.g., death, disease) they are related to each other quickly. Since, the use of negative images was crucial, use of negative words would rapidly increase inter-pair associations. Moreover, we made sure cue words were not related to other cue words in the stimulus set with the association database from the University of Leuven: www.kuleuven.be/semlab.

Images. In a pilot study, participants (N = 24) rated 52 potentially neutral and 52 potentially negatively valenced images from IAPS (Lang et al., 2005) and Google Image. Pictures were rated with the Self-Assessment Manikin (SAM; Lang et al., 2005) on a 9-point rating scale, where 9 represents a high score on each dimension, and 1 represents a low score. From the pilot data, 12 images (and two filler images) with the lowest valence and highest arousal ratings were selected ($M_{valence}$ = 2.29, SD = 0.53; $M_{arousal}$ = 6.2, SD = 0.85). Five IAPS images were selected (2053, 6313, 6821, 9433, and 9911) and seven Google Image Pictures depicting a bullfighter attacked by a bull; an anorexic woman looking in the mirror; masked soldiers carrying guns and explosives; the hanging of two men; elephants killed by poachers; and a man who set himself on fire. Images had a landscape orientation and were of the same size (500×375 pixels). For the final test, each image was divided into four equally sized cut-outs (250×187 pixels).

Visual analogue scale. Participants rated the dependent variables vividness, emotionality, difficulty to retrieve the memory of the target scene, and degree of confidence in their decision ('choice confidence'; see end of "Response Latency Task") on a visual analogue scale (VAS) from 0 (*not vivid/emotional/difficult/certain at all*) to 100 (*very vivid/emotional/difficult/certain*). Choice confidence was added as a novel outcome measure, because correct target image identification in forced choice can be independent of choice confidence. Decreases in confidence are an extremely robust finding of research on obsessive-compulsive disorder in our laboratory (e.g., Dek, van den Hout, Giele, & Engelhard, 2010).

Response latency task. Four cut-outs were taken from each picture (48 cut-outs in total). Participants were asked to identify which cut-out had been paired earlier with the cue word presented on screen. They consecutively they rated their choice confidence. The cue was displayed for 1000ms followed by four cut-outs presented in four quadrants around the cue. The correct cut-out belonging to the target image had to be selected within 4000ms by pressing the button on the numerical keypad that corresponded with the cut-out's location on screen. To avoid learning from novelty elimination, three other cut-outs displayed parts from images that had been used as targets for other cues. There was one pseudo-randomized order set, wherein the total serial position for each condition was identical, and no more than two cues from the same condition were displayed consecutively. Response latency was measured as dependent variable. Additionally, participants rated how confident they were their answer was correct (400ms intertrial interval; ITI).

Post-experimental questions. Participants rated on paper-and-pencil VAS to what extent they were compliant with and had been able to follow the instructions, which was used as a manipulation check. For recall+EM and recall only they rated to what extent they made eye movements, were able to recall the cued target memory, and how vivid and detailed that memory was.

Procedure

Learning phase. Initially, all cue-target pairs –consisting of a word and image– were presented in the middle of a black screen for 8000ms followed by 400ms ITI. The display time was taken from Depue, Banich, and Curran (2006) and was doubled, because participants were instructed not only to associate cue and target, but also to be able to recall the target image as complete and detailed as possible when seeing the cue. Based on a pilot, participants indicated that this display time was sufficient to comply with instructions. Next, participants saw the cue word for 1000ms and were instructed to select the correct target image for each cue from four image options presented in four quadrants around the cue. Three images displayed scenes were targets of other cues. Cue and image options disappeared after 4000ms or after target selection. The maximum decision time was based on a pilot in which all decisions made by participants were within 4000ms. The correct target was highlighted by presenting a green rectangle around the correct target for 2000ms followed by 400ms ITI (Levy & Anderson, 2012) regardless of the participant's answer. When participants correctly selected 11 out of 12 experimental targets, they proceeded to the learning test. They had up to six list repetitions to achieve this criterion. All participants reached this criterion within six repetitions.

Learning test. In the learning test (pre-test), participants were presented with a cue word and retrieved the accompanying target as vividly and detailed as possible, and pressed the spacebar when they did. They then rated the memory of the retrieved target on vividness, emotionality, and difficulty of retrieval. Next, the cue was presented a second time for 1000ms, and the participant had to select the correct target from four scenes within 4000ms. Contrary to the learning phase, no feedback was provided. After each decision, participants rated how confident they were their answer was correct (400ms ITI). After participants completed the learning test for all cues, they proceeded to the eye movements phase.

Eye movement phase. Participants were instructed to retrieve and visualize the target as vividly and detailed as possible after a cue was presented. For one third of the cues, participants were instructed to simultaneously follow a dot of 20 pixel that moved laterally with their eyes (1Hz frequency and 461 pixel amplitude) for 4 intervals of 24 seconds separated by 5-sec breaks ('Recall+EM'). For another third of the cues, the same procedure was used except that participants did not perform a secondary task, but simply looked at the center of the screen while thinking of the target ('Recall Only'). The final third of the cues were not presented in this phase and served as "natural decay" control condition. The duration of the experimental manipulations was identical to previous studies (see Van den Hout & Engelhard, 2012). In total, eight cues were presented in this phase and no more than two cues with the same instruction were given in a row. Word pairs were rotated through conditions over participants. After the 4 × 24s of the eighth cue, participants continued with the final test phase.

Final test phase. Memory for all experimental items was assessed in the final test (post-test). Participants were presented with the cue and were instructed to retrieve the memory of the image associated with the target. They then rated vividness and emotionality of the memory of the retrieved target, and did this for all cues before continuing to latency response task. After the latency response task, participants filled-out the post-experimental questionnaire.

Results

Data with more than three standard deviations from the mean were corrected. (Results for data with and without outlier correction were comparable.) Moreover, in order to retain sufficient power, slight violations of sphericity were corrected with Greenhouse-Geisser ($.70 \geq \varepsilon < .75$) or Huynh-Feldt corrections ($\varepsilon \geq .75$). In case of severe violations ($\varepsilon < .70$) a multivariate test statistic (Pillai-Bartlett trace; *V*) is reported. Analyses were performed only on pairs for which participants recalled the target on the final learning test. Table 1 presents means and standard deviations of the self-report ratings for the three conditions.

Manipulation checks

An analysis of variance (ANOVA) revealed that there were no differences between conditions on difficulty of retrieval, $F(2, 118) = 1.56$, $p = .21$, $\eta_p^2 = .03$, indicating comparable levels of recall before entering the eye movement phase.

On the post-experimental questions, participants indicated that they frequently made eye movements during recall+EM ($M = 83.77$, $SD = 14.58$), and rarely during recall only ($M = 13.87$, $SD = 18.68$), $t(59) = 18.79$, $p < .001$, $d = 4.2$. They were better able to recall the memory of the image during recall only ($M = 80.63$, $SD = 15.40$) compared to recall+EM ($M = 64.78$, $SD = 24.36$), $t(59) = 4.91$, $p < .001$, $d = 0.78$. Additionally, the recalled image was more vivid and detailed during recall only ($M = 72.35$, $SD = 18.50$) compared to recall+EM ($M = 52.25$, $SD = 24.15$), $t(59) = 6.56$, $p < .001$, $d = 0.96$.

Vividness

A 2 (Pre, Post) × 3 (recall+EM, recall only, control) ANOVA showed no significant main or interaction effects for vividness ratings, Fs < 1.90, ps $> .17$, $\eta_p^2 < .04$.

Emotionality

A 2 (Pre, Post) × 3 (recall+EM, recall only, control) ANOVA did not reveal significant main or interaction effects, Fs < 2.23, ps $> .13$, $\eta_p^2 < .04$.

Confidence

Because confidence ratings related to complete images before the experimental manipulation and to partial images after the manipulation, the former were entered as covariates in an ANCOVA. Pre-manipulation confidence ratings for recall+EM, $F(1, 56) = 17.38$, $p < .001$, $\eta_p^2 = .24$, and for control, $F(1, 56) = 11.87$, $p = .001$, $\eta_p^2 = .18$, related significantly to post- manipulation confidence ratings. There was no relation between pre-manipulation recall only scores and post-manipulation ratings, $F < 1$. The main analysis of condition on post-manipulation ratings showed no effect when controlling for the pre-experimental confidence ratings, $F(1.672, 93.641) = 0.17$, $p = .80$, $\eta_p^2 = .003$ (Huynh-Feldt correction).

Table 1. Means and standard deviations (in parentheses) of Difficulty, Confidence, Vividness and Emotionality Ratings for the recall+EM, recall only, and control conditions.

	Difficulty	Confidence	Vividness		Emotionality	
			Pre	Post	Pre	Post
Recall+EM	29.85	89.90	72.30	73.90	55.16	53.99
	(16.82)	(11.6)	(14.22)	(16.59)	(16.23)	(19.04)
Recall only	31.21	89.05	70.77	73.59	51.05	53.07
	(15.68)	(10.88)	(14.25)	(14.74)	(20.13)	(17.46)
Control	33.78	88.90	70.74	71.27	50.43	53.16
	(17.93)	(12.05)	(14.72)	(15.28)	(19.74)	(19.68)

Response latencies

Before analyses, response latencies were log transformed because of natural skewness in scores. Similar to confidence ratings, response latencies related to complete images before the experimental manipulation and to partial images after the manipulation. Therefore pre-manipulation response latencies were entered as covariates in an ANCOVA. Pre-manipulation response latencies for recall only, $F(1, 56) = 6.70$, $p = .012$, $\eta_p^2 = .11$, related significantly to post-manipulation response latencies. Other pre-manipulation response latencies did not, Fs < 3.10, ps $> .08$, $\eta_p^2 < .06$. The main analysis for post-manipulation response latencies, though, did not reach significance when controlling for pre response latencies, $F(2, 112) = .7$, $p = .5$, $\eta_p^2 = .01$. Table 2 presents means and standard deviations.

Accuracy

Participants needed, on average, 1.77 ($SD = 1.30$) repetitions to achieve the learning criterion. Because analyses were performed only on pairs for which participants recalled the target on the final learning test, pre-manipulation scores for accuracy reached the ceiling (100% for all conditions). Therefore, for accuracy, an ANOVA was performed on post-manipulation scores only (see Table 2). Participants did not differ in their accuracy to select the correct target images for the different conditions, $F(2, 118) = 0.45$, $p = .64$, $\eta_p^2 = .007$.

Table 2. Means and standard deviations (in parentheses) of Response Latencies and Accuracy for recall+EM, recall only, and control conditions.

	Response Latencies	Accuracy
Recall+EM	1268.7 (310)	0.97 (0.06)
Recall only	1292.4 (383.8)	0.96 (0.07)
Control	1280.9 (324.6)	0.97 (0.06)

Discussion

The aim of this study was to test whether recall+EM affect the accessibility of associative memory representations when emotionally negative materials are used. In this extended replication of Van den Hout et al. (2013a), we found no blurring of the emotional memory representations for recall+EM compared to recall only or no intervention. This was reflected in the absence of any effects on subjective ratings of vividness and emotionality, but also in objective measures of memory accessibility, specifically latency responses. Participants reported they had complied with the instruction to make EM. They also reported that mental images were less vivid, less detailed, and more difficult to retrieve *during* the intervention. However, this seems trivial, because these effects did not persist at the post-test.

The data show no reductions in memory accessibility, but they do not necessarily falsify working memory theory as an explanation for the effects of dual taxation. The absence of effects may be a consequence of our materials. Could the data have been different if self-relevant or less negative pictures would have been used? The vast majority of earlier studies on the effects of demanding secondary tasks have used personally relevant, autobiographical memories (e.g., Engelhard et al., 2011a; Engelhard et al., 2011b; Gunter & Bodner, 2008), while our study used novel emotional images. These images could have lacked the potential to elicit sufficient levels of arousal, which may be necessary in the dual-tasking procedure to reduce vividness (Little, Remijn, Tinga, Engelhard, & van den Hout, 2017). Van den Hout et al. (2014) showed that only negative autobiographical memories –which are associated with transient levels of arousal– were reduced in their vividness ratings, while neutral memories were not. This suggests arousal is a prerequisite for (re-)encoding of memories after dual taxation. Although our materials were not negative autobiographical memories, they were, however, thoroughly piloted and showed sufficient arousal during the pilot and on pre-test measures. Moreover, other studies have used non-idiosyncratic materials and showed effects for recall+EM compared to recall only (e.g., Leer et al., 2013) and specifically showed that effects can be found for materials that are neither autobiographical, nor self-relevant (Andrade et al., 1997; Van den Hout et al., 2013a). It therefore seems unlikely that intrinsic qualities of our novel images per se explain the absence of effects.

The tasks we used may also have limited the effects of dual taxation. This study used a latency response task that was based on a similar task in Van den Hout et al. (2013a). In their task, participants were instructed to react as fast and accurately as possible, and had to decide whether a cut-out was old or new. In our experiment, participants received comparable instructions. Yet they did not make an old-new judgment, but a source judgment: they had to indicate which of the four displayed images was the cue's target. Though these tasks look similar, they probably draw on different types of recognition: old-new recognition and source recognition. Old-new recognition can generally be performed at lower levels of item differentiation than source decisions (Johnson et al., 1993). As a consequence, slightly blurred recall+EM pairs may show effects for old-new decisions compared to recall only pairs, but not for source decisions. It is possible that successful source memory differentiation does occur when pairs from the recall+EM condition differ more in the level of memory blurring from the recall only or no-presentation control conditions. Theoretically, it may be possible to find differences between conditions when an old-new recognition task is used instead of a source recognition task.

A different explanation may also be found in the response latency task, specifically in the distractors that were used as targets in previous trials. During the final part of the experiment, participants had to select the correct target out of four cut-out images. Here, the use of distractors that have been used as targets in previous trials may cause response-inhibition, and related response delays. Alternatively, it is also possible that the blurring effects were abolished, when participants specifically saw recall+EM images as distractors in earlier trials, and in later trials saw these same images as targets. As a consequence the image could have been reinstated in full and any condition effects were abolished as well. This does not, however, explain why there were no effects on any of the subjective ratings, which preceded the response latency task. Though, it might be possible that the effects of recall+EM in our design were subtle and only detectable with RT, but that this effect was abolished by how our response latency task was designed. Furthermore, emotional interference may have played a role, because participants had to select one out of four highly unpleasant images, which may have caused response delays that are not due to simple blurring or retrieval delay. Indeed, it cannot be ruled out that there might be a floor effect in that the delay is beyond the critical point where retrieval differences can be found.

Though several plausible explanations for the lack of effects can be found in the response latency task, it still leaves unanswered why subjective ratings of the images did not change, because these ratings preceded the latency response task, and were recalled without seeing any of the targets from other pairs. Perhaps cued recall of the to-be-recalled material influenced the effectiveness of dual taxation. On the one hand, cued recall could facilitate episodic memory retrieval compared to non-cued recall (Tulving & Pearlstone, 1966), and thus should allow the participant to vividly retrieve the associated target image. This is reflected in relatively low pretest scores for self-assessed difficulty of retrieval and high scores for vividness. Subsequent retrieval of a vividly cued image should therefore be blurred as a consequence of dual taxation. This, however, did not happen. On the other hand, even though participants quickly learned the associations –which hints at strong relationships between cue and target– these may simply not have been strong enough. We used neutral words referring to objects that people may encounter frequently in daily life and thus could be linked to various situations. Although the association between the cue and target is novel and recent, it is unlikely that the cue's path exclusively leads to the target. Therefore, after images were cued in a recall+EM trial, participants may not have thought of the target image all the time, but of other images, objects, or words. As a consequence, the target may not have been sufficiently blurred, and reductions in self-reported vividness and emotionality may not have been experienced for recall+EM. Hence, generally cued recall may ameliorate memory, but not when a multipath cue needs to prompt one specific target for prolonged periods of time. Additionally, this might also explain why there were no latency response effects. If cues did not elicit specific and continuous target retrieval, then differential item blurring and successful source recognition could not have occurred.

Provided that cued recall was primarily responsible for the lack of effects, the question ensues whether dual taxation is able to affect associative memory networks. A review of earlier studies showed that WM taxation, specifically eye movements, is able to reduce subjective ratings of emotional memories when those memories were recalled immediately after the intervention (Lee & Cuijpers, 2013). The current study showed that these memories were not changed subjectively or objectively when cued with a memory reminder. Perhaps this limitation signals a boundary condition for this paradigm and limits the robustness of the dual taxation paradigm. Until now, Van den Hout et al. (2013a) conducted the only study that found effects of the eye movement intervention on objective measures of memory accessibility, and it cannot be ruled out that this represents a chance finding. It should be noted that other studies using objective measures of memory valence or emotionality have shown effects after dual taxation, such as eye blink startle reflex diminution (Engelhard et al., 2010), reduced heart rate variability (Schubert et al., 2011), and decreased electrodermal responses (Barrowcliff et al., 2004), but these objective measures primarily related to arousal levels and not memory accessibility per se.

If it does, however, signal a limitation of the dual taxation paradigm, then memory change under dual taxation conditions may require a specific form of recall. This would imply that a memory must be directly recalled as opposed to memory recall that is initially cued by reminders. The former has been frequently used in previous work (e.g., Van den Hout et al, 2014). Interestingly, this does not preclude the possibility of finding effects with cued recall. Cued recall *after* dual taxation may still lead to reductions in subjective ratings, but only if the memory is recalled directly *during* dual taxation. This does; however, seems to contradict predictions from current trauma theories (Brewin, 2001; Ehlers & Clark, 2000), which state that encountering cues associated with an emotional memory could instantly trigger vivid intrusions of that memory. In the current study, cues probably elicited retrieval of the target memory, but most likely only briefly.

The lack of effect of cued recall also has implications for associative network theories, in which connectivity between different representations is paramount. A central tenet of network theories is modifiability of the network's structure after activation (e.g., Foa et al., 1989). It is possible that small networks (e.g., two nodes representing only stimulus characteristics, such as gate – car accident) are difficult to modify. Larger, more ecologically valid networks typically also contain elements regarding responses ('panic') or meanings ('I am helpless'), next to mere stimulus relations. Perhaps these former relations are an inherent changeable part of the associative network, but also a part that does not need to be targeted directly. It is possible that stimulus characteristics change during dual taxation, and affect response and meaning elements, which changes how a person feels or thinks about an event. As a consequence, change in subjective experiences may be difficult to accomplish in smaller, laboratory created networks because these lack elements of meaning.

In sum, we found that eye movements during recall did not blur emotional memory representations measured by subjective or objective measures of memory accessibility. Response inhibition and emotional interference do not seem able to explain all the effects. Cued recall, on the other hand may; it may not have been potent enough to elicit specific and continuous target retrieval for differential item blurring to occur. Although memory effects following eye movements were not observed, it is unlikely –given the substantial body of evidence– that reductions in self-reported ratings are a chance discovery. Changes in objective measures of memory accessibility therefore still need to pass the critical test of replication.

Acknowledgements

This research was supported by a TOP grant (dossier number: 40-00812-98-12030) from the Netherlands Organization for Health Research and Development (ZonMw) awarded to Marcel A. van den Hout. We are grateful to Anne Klein Schiphorst for her assistance in piloting the materials and to Sheanen Danoe and Nikki van de Loo for their assistance in testing. We are especially thankful to the reviewers who suggested very knowledgeable and thoughtful explanations for our effects.

Author contributions

KvS and MAvdH developed the study concept. All authors contributed to the study design. KvS was responsible for data collection. KvS performed the analyses and drafted the manuscript. MAvdH and IME provided critical revisions.

CHAPTER 3

SPEED MATTERS: RELATIONSHIP BETWEEN SPEED OF EYE MOVEMENTS AND MODIFICATION OF AVERSIVE AUTOBIOGRAPHICAL MEMORIES

S. C. VAN VEEN
K. VAN SCHIE
L. D. N. V. WIJNGAARDS-DE MEIJ
M. LITTEL
I. M. ENGELHARD
M. A. VAN DEN HOUT

Frontiers in Psychiatry, 6, 45 (2015)
doi:10.3389/fpsyt.2015.00045

Abstract

Eye Movement Desensitization and Reprocessing (EMDR) is an efficacious treatment for posttraumatic stress disorder. In EMDR, patients recall a distressing memory and simultaneously make eye movements (EM). Both tasks are considered to require limited working memory (WM) resources. Because this leaves fewer resources available for memory retrieval, the memory should become less vivid and less emotional during future recall. In EMDR analogue studies, a standardized procedure has been used, in which participants receive the same dual task manipulation of 1 EM cycle per second (1 Hz). From a WM perspective, the WM taxation of the dual task might be titrated to the WM taxation of the memory image. We hypothesized that highly vivid images are more affected by high WM taxation and less vivid images are more affected by low WM taxation. In study 1, 34 participants performed a reaction time task, and rated image vividness and difficulty of retrieving an image, during five speeds of EM and no EM. Both a high WM taxing frequency (fast EM; 1.2 Hz) and a low WM taxing frequency (slow EM; 0.8 Hz) were selected. In study 2, 72 participants recalled three highly vivid aversive autobiographical memory images (n=36) or three less vivid images (n=36) under each of three conditions: recall + fast EM, recall + slow EM or recall only. Multi-level modelling revealed a consistent pattern for all outcome measures: recall + fast EM led to less emotional, less vivid and more difficult to retrieve images than recall + slow EM and recall only, and the effects of recall + slow EM felt consistently in between the effects of recall + fast EM and recall only, but only differed significantly from recall + fast EM. Crucially, image vividness did not interact with condition on the decrease of emotionality over time, which was inconsistent with the prediction. Implications for understanding the mechanisms of action in memory modification and directions for future research are discussed.

Keywords: EMDR, Eye movements, Autobiographical memory, Working memory, Vividness, Emotionality

Introduction

Trauma-exposed individuals may suffer from distressing and intrusive memories of their traumatic experience and some even develop posttraumatic stress disorder (PTSD; Kessler, Tat Chiu, Demler, & Walters, 2005). Eye Movement Desensitization and Reprocessing (EMDR) is a psychological treatment for PTSD, and its efficacy is comparable to Cognitive Behavioral Therapy (Bisson et al., 2007; Seidler & Wagner, 2006). A key aspect of EMDR is that the patient makes bilateral Eye Movements (EM) during the retrieval of traumatic memory images. Empirical research has confirmed that this dual-task approach reduces the image vividness and emotional intensity of an aversive memory, both in healthy persons and in patients with PTSD (for a meta-analysis, see Lee & Cuijpers, 2013). Note that in EMDR analogue studies, a standard "dose" is typically used: EM with a speed of 1 cycle per second (1Hz), in sets of 24 seconds (e.g., van den Hout, Muris, Salemink, & Kindt, 2001). This presumes that patients and aversive memories respond equally well to the same dual-task manipulation. Recent insights from experimental studies challenge the efficacy of this standardized procedure (e.g., Gunter & Bodner, 2008; Maxfield, Melnyk, & Hayman, 2008; van den Hout & Engelhard, 2012). Therefore, the aim of the current research was to test whether titration based on image vividness enhances the effects of dual-task manipulation on aversive memories.

A range of experimental studies provides support for a working memory (WM) account to explain how EM decrease the image vividness and emotional intensity of negative memories (for an overview, see van den Hout & Engelhard, 2012). More specifically, holding an emotional memory image in mind and performing EM will both tax the limited resources of WM (Andrade, Kavanagh & Baddeley, 1997; Gunter & Bodner, 2008). Consequently, competition between these tasks should impair retrieval of the image with its accompanied details and emotions, and result in immediate decreased image vividness and emotional intensity of the memory before its return to long-term store. A laboratory model has been used to critically test this WM account. In this model, participants recall a negative memory image with or without simultaneously making EM. Image vividness and emotional intensity are measured before and after this intervention. Studies with healthy participants have shown that recall + EM decreases the vividness and/or emotionality of the recalled memory image, while recall without EM (recall only; RO) does not (van den Hout & Engelhard, 2012). This effect has been replicated with other cognitively demanding tasks, such as counting backwards (Engelhard, van den Hout, & Smeets, 2011), attentional breathing (van den Hout et al., 2011), drawing a complex figure (Gunter & Bodner, 2008), and playing the computer game Tetris

(Engelhard, van Uijen, & van den Hout, 2010). Furthermore, it has not only been found for mental images of adverse past events, but also for mental images of imagined, aversive future events (e.g., Engelhard et al., 2011). As predicted, tasks that barely tax WM, such as passively listening to sounds, are less effective than more cognitively demanding tasks (e.g., van den Hout et al., 2012). These studies suggest that any dual-task that sufficiently taxes WM may decrease the vividness and/or emotionality of the recalled memory image.

Although many studies have shown that various dual-tasks affect emotional memory images, less is known about boundary conditions and optimization of the dual-task manipulation. The degree to which competition will occur between the WM load of the memory image and the WM load of the dual task partly depends on a person's WM capacity. Individuals with a large WM capacity are expected to be relatively proficient in performing tasks simultaneously (multitasking). Because there will be less competition between the two tasks (memory image recall and dual task) for them, compared to individuals with a low WM capacity, the effects on memory image should be smaller. Evidence for a correlation between WM capacity and memory effects comes from a study by Gunter and Bodner (2008) who found medium negative correlations between automated reading span scores –an indicator of WM span– and decreases of vividness and emotionality within the recall + EM condition. This finding was replicated by two other studies that showed that individual differences in WM span are negative related to beneficial effects of dual taxation of memory image recall + WM taxing: the larger the WM span, the smaller the benefits of recall + WM taxing (van den Hout et al., 2010, 2011). To test the feasibility of the WM theory, Maxfield, Melnyk, and Hayman (2008) manipulated the speed of EM. As predicted, they found that fast EM (1.25Hz) resulted in larger decreases in image vividness and emotional intensity than slow EM (1Hz), and both EM conditions led to larger decreases than a control condition. The authors argue that fast EM are more difficult to perform (i.e., they are more taxing), which leads to larger effects on memory images. Although this is plausible, the actual WM load of the two speeds of EM was not measured. Also, the stimulus presentation was a repetition of short intervals of dual-task manipulation (left-right-left appearance of a stimulus). One could argue that this procedure tested the capability of task switching, rather than ongoing dual-task performance.

Contrary to the prediction that the higher the WM load of the dual task, the larger the dual-task manipulation effects, Gunter and Bodner (2008) hypothesized that this relationship may not be linear. A task that is slightly taxing may not disrupt the memory image enough, and a task that is overly taxing might preclude holding the memory image in mind, thereby preventing competition effects. Therefore, they proposed an inverted U-shape function. In other words, too little or too much WM taxing may lead to smaller effects than WM taxing that is intermediate. This was tested and partially confirmed by Engelhard, van den Hout and Smeets (2011), who found an inverted U-shape function for emotionality, but not for vividness. Participants recalled a negative memory image and performed one of four arithmetic tasks: exposure alone, or exposure with "simple" subtraction, "intermediate" subtraction, or "complex" subtraction. Prior to the memory experiment, the WM taxation of the four tasks was assessed using a discriminative reaction time (RT) task and the results indicated that the subtraction tasks indeed increasingly taxed the WM, with simple subtraction taxing WM the least and complex subtraction taxing the most. In line with the inverted U-shape hypothesis, emotional intensity of the memory decreased more after recall during simple or intermediate subtraction than when after recall during complex subtraction or no subtraction. Results for vividness were in the expected direction, but were not significant. Variation was larger for vividness ratings than for emotionality ratings, and this latter may have caused the difference between the dependent variables. To sum up, research indicated that the WM load of that dual task is related to the effectiveness of the intervention, and that this relation presumably follows an inverted U-shape function. It is unclear, however, whether these effects are translated to various speeds of EM.

From a theoretical perspective, the effectiveness of the dual-task manipulation depends not only on the WM load of the dual task, but also on its interaction with the WM load of the memory. The WM load of the memory may be affected by variation in memory image vividness: highly vivid images are presumed to tax the WM more than less vivid images (Baddeley & Andrade, 2000). Obviously, the degree of image vividness of aversive memories varies between individuals who have experienced the same situation and within one individual over time. These variations in image vividness may therefore influence the variation in WM load. According to the inverted U-shape hypothesis, if a memory image is highly vivid, a relative low degree of taxing WM by the dual task may produce insufficient blurring. Conversely, if the memory image is less vivid, strong WM taxing may preclude memory recall. Therefore, in order to maximize memory effects, the WM theory implies that there is a need for titration: highly vivid memories require a relatively high WM load and less vivid memories a lower load.

The current study used the WM framework to investigate the interaction between the WM load of the memory image and the WM load of the dual task. In study 1, we examined the WM load of five different speeds of EM. We hypothesized that faster EM are more taxing. This study resulted in the selection of two conditions: fast EM and slow EM. In study 2, participants recalled three highly vivid distressing memories or three distressing memories that were less vivid. These memories were randomized to each of three conditions: recall + fast EM, recall + slow EM or RO. We predicted that (1) relative to RO, both EM conditions result in memory images that are less emotional, less vivid and more difficult to retrieve, and more importantly (2) highly vivid memories benefit more from fast than slow EM during recall, while less vivid memories benefit more from slow than fast EM during recall.

Study 1: WM taxation of different speeds of EM

In order to select two speeds of EM that significantly differ in WM taxation, we tested the WM load of different speeds of EM in a within-subjects design. Participants performed a discrimination RT task during the performance of six tasks: five different speeds of EM and no EM. Slower RTs indicate the degree of taxation (Bower & Clapper, 1989). In addition, participants were asked to hold six well-known images in mind (e.g., "your own kitchen"), while carrying out the same six tasks, and rated the vividness and difficulty to hold an image in mind during each task. We included vividness and difficulty ratings to test whether participants were still able to recall an image while simultaneously making the EM. We hypothesized that EM are more taxing than no EM, and that faster EM are more taxing than slower EM, resulting in larger RTs.

Methods

Participants

Participants were recruited through advertisements at Utrecht University and the University of Applied Sciences (Hogeschool Utrecht), located at the same campus. Thirty-six participants (8 men, 28 women, M_{age}= 21.89, *SD* = 2.08) were tested, using no exclusion criteria. Two participants were removed from analyses due to technical problems. Participants received course credit or financial compensation for participation.

Materials and procedure

Participants were seated in front of a computer screen with a screen resolution of 1280 x 1024 at a distance of approximately 45cm. OpenSesame 2.8.3 (Mathôt, Schreij, & Theeuwes, 2012) was used to present stimuli. First, the low tone and high tone 1-s beeps (44.1 kHz) of the discrimination RT task were introduced. Beeps were administered to both ears through headphones using a constant volume. Participants pressed the *z*-key with their left index finger for low beeps and the */*-key with their right index finger for high beeps. Beeps were presented randomly with a mean stimulus-onset asynchrony of 2.6-s (*SD* = 0.4). After a practice trial of 10 beeps, the experiment started. Participants were asked to categorize 20 low and 20 high beeps with or without making EM. In the EM conditions, a white 20 pixel dot appeared in the middle of a black screen and moved horizontally from side-to-side, with a movement amplitude of 461 pixel. The EM conditions had speeds of 0.4, 0.6, 0.8, 1.0 and 1.2Hz (number of left-right-left cycles per second). Participants in the EM conditions were instructed to keep their head still and follow the dot with their eyes, and participants in the no EM were instructed to look at the middle of the screen (no dot was shown). The experimenter sat next to the participant and checked whether the eye movements were in accordance to the manipulation. If needed, the experimenter shortly repeated the instruction. In all conditions, the task was presented for a period of 106.6-s, adjusted to the average total time of beeps plus one (41x2.6-s). The order of the speed of EM was randomly assigned, but each participant completed all six conditions.

To test whether participants were still able to recall a mental image while they simultaneously made EM, participants received the same condition again immediately after the RT trial, but instead of responding to beeps, they were instructed to simultaneously hold a well-known image in mind as vividly as possible. After 24-s, participants rated the vividness and difficulty of that image during manipulation on a Visual Analogue Scale (VAS), ranging from 0 (*not vivid/difficult at all*) to 100 (*very vivid/difficult*). The well-known mental images were the participant's kitchen, bathroom, bed, wardrobe, front door, and bicycle. Latin-square counterbalancing was used to order the sequence of these six images.

Design and data analyses

To test the WM taxation of the various speeds of EM, relative to no EM, a repeated measures analysis of variance (ANOVA) was performed with speed of EM as within-subjects factor and average RT as outcome measure. The first and last beeps were excluded from the calculation of the average RT, to exclude potential transition delays. Differences in vividness and difficulty of holding an image in mind between conditions were analyzed by two repeated measures ANOVAs with speed of EM as within-subjects factor and vividness or difficulty as outcome measure. Alpha levels of .05 were used; they were one-tailed for tests crucial to the hypothesis. For small violations of sphericity, the degrees of freedom of the F-distribution were corrected with either Green-Geisser ($.70 \geq \varepsilon < .75$) or Huynh-Feldt corrections ($\varepsilon \geq .75$). More severe violations ($.70 < \varepsilon$) were corrected using a multivariate test statistic (Pillai-Bartlett trace; V).

Results

The average RT varied significantly across conditions, $V = 0.48$, $F(5, 29) = 5.40$, $p = .001$, $\eta_p^2 = .48$ (see Figure 1). Pairwise comparisons showed that all EM conditions during the RTT yielded increased RTs compared to no EM, range $M_{dif} = 60\text{-}100$, ps $< .001$. Furthermore, simple contrasts indicated that RTs in the fastest condition (1.2Hz) were significantly greater than in the 0.8Hz EM condition, $F(1, 33) = 4.50$, $p = .02$, $\eta_p^2 = .12$, or 1.0Hz EM condition, $F(1, 33) = 4.22$, $p = .03$, $\eta_p^2 = .11$. The average number of correct items was high ($M_{range} = 36–37$ out of 39) and did not differ between the conditions.

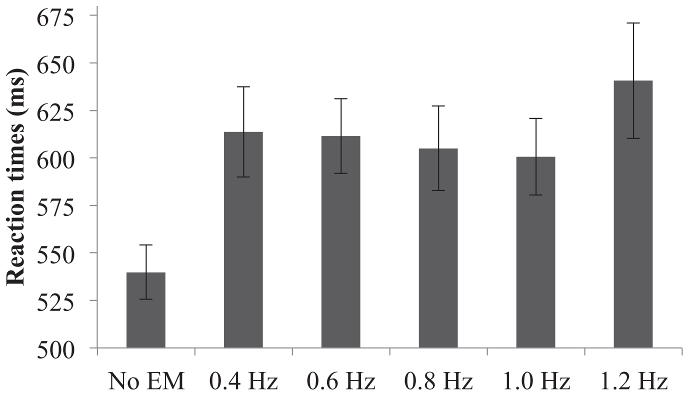


Figure 1. Mean reaction times (ms) and SEs for the different speeds of EM and no EM

Average vividness scores differed between conditions, $F(5, 29) = 11.42$, $p < .001$, $\eta_p^2 = .26$ (see Figure 2). Pairwise comparisons showed that all EM conditions decreased the vividness of the image compared with no EM, $M_{dif} = 15\text{-}31$, ps $< .002$. The relation between WM taxation and vividness indicated a clear negative linear relationship: vividness decreased as WM taxation increased. Difficulty retrieving the image while performing the dual-task differed between the conditions, $F(3.58, 118.15) = 9.25$, $p < .001$, $\eta_p^2 = .22$ (see Figure 2). Pairwise comparisons showed that it increased for all EM conditions compared with no EM, $M_{dif} = 16\text{-}29$, ps $< .002$. Simple contrasts indicated that for both vividness and difficulty ratings, the 1.2Hz EM condition differed significantly from the 0.8Hz and 1.0Hz condition, ps $< .05$. The highest speed (1.2Hz) resulted in a mean vividness of 45.20 ($SD = 26.46$) and mean difficulty of 51.90 ($SD = 27.91$).

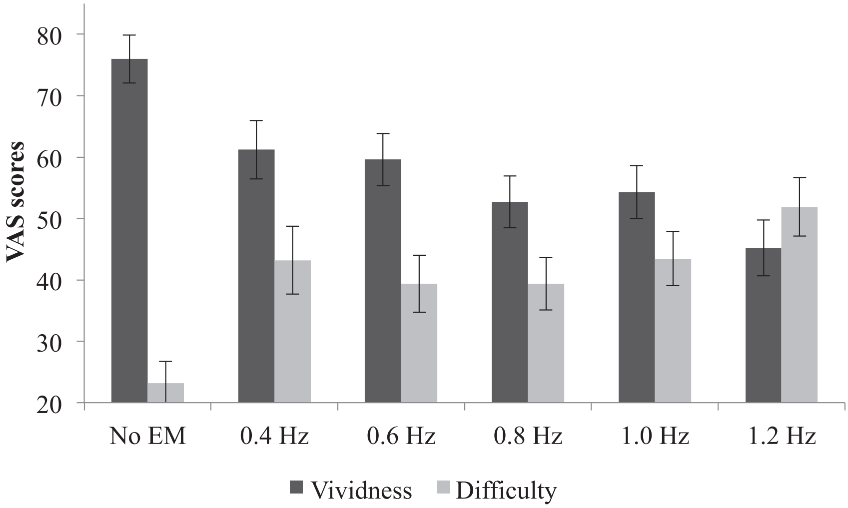


Figure 2. Mean VAS scores and SEs for vividness and difficulty for the different speeds of EM and no EM

Discussion study 1 and introduction study 2

In line with previous research (e.g., van den Hout et al., 2011), all EM conditions resulted in slower RTs compared to no EM, indicating that performing EM indeed taxes WM. Between the EM conditions, EM of 1.2Hz produced more WM taxation, lower image vividness and higher difficulty to retrieve the image during manipulation compared to EM of 0.8Hz and 1.0Hz. The 0.8Hz and 1.0Hz conditions did not differ from each other on any of the outcome measurements. Since 1.0Hz is the standard EM speed in EMDR, this could be considered an "intermediate" speed of EM. To keep the amount of variation equal on both ends, we selected 0.8Hz for the slow EM condition and 1.2Hz for the fast EM condition. Study 2 tested whether the WM load of EM interacts with the image vividness of a negative memory.

Study 2: interaction between speed of EM and image vividness

Methods

Participants

We recruited 92 undergraduate students through advertisements at the Utrecht University and the University of Applied Sciences (Hogeschool Utrecht). Exclusion criteria were knowledge about EMDR, prior participation in an experiment from our laboratory that required participants to recall memories, or medication use that may affect concentration, such as benzodiazepines. We excluded 20 students based on these exclusion criteria. The final sample consisted of 72 participants (22 male, 50 female, M_{age}= 22.40, *SD* = 3.81). They were randomly assigned to one of two groups: 'highly vividness memories', *n* = 36; 'less vividness memories', *n* = 36. Participants received course credit or financial compensation.

Materials and general procedure

Participants were tested individually in a quiet room. After providing written informed consent, participants were interviewed by the experimenter (see below). Participants selected three negative memories following the procedure used by van den Hout et al. (2001). Next, in line with the Dutch EMDR standard protocol (de Jongh, & ten Broeke, 2012), they selected a target image of each memory. During the second half of the experiment, participants were seated behind a computer with a screen resolution of 1280 x 1024 at a distance of about 45cm. OpenSesame 2.8.3 (Mathôt et al., 2012) was used to present stimuli.

Memory selection

During the first half of the experiment, participants selected three negative memories that were at least one week old and still evoked relevant feelings (i.e. fear/anxiety/sadness). Participants in the highly vivid memories group were instructed to select three negative 'memories that are very clear and detailed', and participants in the less vivid memories group were instructed to select three negative 'memories that are relatively vague and low on details'. If participants found it difficult to select memories, the experimenter presented a list of examples (e.g., eye-witness of a traffic-accident, a job rejection, an argument with a family member), and stressed that vividness of memories is subjective, so the example memories were merely given to stimulate the selecting process. Participants wrote down the content of each memory on a card and indicated the vividness (with 0 *not at all vivid* to 100 *very vivid*) and emotionality (0 *not at all unpleasant* to 100 *very unpleasant*) of each memory. The experimenter checked if these ratings were within the intended range, which was 70-100 for vividness in the highly vivid memories group, 30-60 for vividness in the less vivid memories group, and 50-90 for emotionality in both groups. If it was not, the experimenter asked the participants to select another memory. Memories were ranked based on vividness ratings (1 = *most vivid*, 3 = *least vivid*, 2 = *in between*). The order of the target image selection, as well as the order of the conditions, was counterbalanced based on this ranking.

Target image selection

Next, the experimenter asked the participants to describe the memory in global story lines. Then the experimenter asked the participant to identify the worst moment of this memory and describe this moment as a still image (i.e., "target image"). The participants assigned a descriptive, relatively neutral label to each target image, to act as a cue during the experiment.

Experiment

Then the participants performed a pre-test, an intervention phase, and a post-test for each condition. In the pre-test, participants recalled their target image for 10-s and gave ratings of emotional valence, vividness and difficulty of retrieving the target image on the VAS (ranging from 0 *not at all unpleasant/vivid/difficult* to 100 *very unpleasant/vivid/difficult*). In the intervention phase, they recalled their target image six times for 24-s, with 10-s rest periods in between. Each rest period ended with a 2-s instruction to recall the target image again. In each EM condition, participants held their head still and looked at a horizontally moving white dot (20 pixel) on a black screen. The dot had a movement amplitude of 461 pixel, and a speed of 0.8Hz in the slow EM condition and 1.2Hz in the fast EM condition. In the RO condition, participants recalled the target image and looked at the black screen. If participants moved their head or eyes incorrectly, the experimenter briefly repeated the instructions. The post-test was immediately after the intervention. In the post-test, participants again brought the target image to mind for a 10-s period and rated the same VAS.

Results

Manipulation check

During memory selection, all participants managed to select three memories that matched the vividness criteria. However, a manipulation check based on the vividness ratings in the pre-test indicated that only 33 participants (45.8%) had three target images within the vividness range of their condition. For the less vivid memories group, vividness scores during the memory selection were significantly lower (M = 50.31, SD = 6.07) compared to the pre-test ratings of the target image (M = 63.64, SD = 13.05; $t(35) = -5.82$, $p < .001$). For the highly vivid memories group, vividness scores during memory selection and the pre-test did not differ from each other ($M_{selection}$ = 79.85, SD = 5.24; $M_{pre\text{-}test}$ = 79.67, SD = 8.13, p = .91). Because our manipulation check indicated that target image vividness did not match the intended group criteria (highly vivid memories vs. less vivid memories), we analyzed the data on the memory level instead of on the participant (group) level.

Analysis strategy

Memories were nested within participants. Therefore, we analyzed the data with multilevel modelling using three levels: 432 repeated measures (level 1) of 216 memories (level

2), nested within 72 participants (level 3). We conducted the analyses with Hierarchical Linear and Nonlinear Modeling, version 6 (HLM6, Raudenbush, Bryk, & Congdon, 2004). For our first hypothesis that EM would decrease emotionality and vividness, and increase the difficulty of retrieving the memory image more than RO, we analyzed emotionality, vividness, and difficulty over time between the conditions. Figure 3 shows the mean difference scores (post-test minus pre-test) and standard errors of all three conditions on emotionality, vividness, and difficulty. Table 1 shows the fixed and random parts of the same multilevel model applied to each outcome measure. Condition was coded as dummy variable, with RO as reference condition. Therefore, the variable *RO_slowEM* indicated the difference between RO and the slow EM condition, and *RO_fastEM* indicated the difference between RO and the fast EM condition. The mixed equation for each model was: Outcome variable$_{ijk}$ = $\beta_{00} + \beta_{10}$ $(time)_{ijk}$ + β_{01} $(RO\_slowEM)_{jk}$ + β_{02} $(RO\_fastEM)_{jk}$, + β_{11} $(RO\_slowEM)_{ijk} * (time)_{ijk}$ + β_{12} $(RO\_fastEM)_{ijk} * (time)_{ijk} + v_{0k} + u_{0jk} + u_{1jk}$ (i = time, j = memory, k = person).

The second hypothesis was that highly vivid memory images benefit more from fast EM than slow EM during recall, and less vivid memory images benefit more from slow EM than fast EM during recall. To test the difference between slow and fast EM, we used slow EM as reference condition. Accordingly, the dummy *slowEM_RO* indicated the difference between slow EM and RO, and the dummy *slowEM_fastEM* indicated the difference between slow EM and fast EM. To observe the three-way-interaction between pre-test vividness, condition and time, interaction variables between the centered *pre-test vividness* variable and the dummy condition variables were added as predictor for the intercept at the second level and as predictor for the slope of time at the first level (Model 4). Support for the hypothesis should materialize as a significant negative coefficient in predicting the slope of time for the variable *pre-test vividness* x *slowEM_fastEM*: the higher the vividness of the target image at pre-test, the more decrease in emotionality for the fast EM condition compared to the slow EM condition. Likewise, the lower the vividness of the target image at pre-test, the less decrease in emotionality for the fast EM condition when compared to the slow EM condition.

The mean pre-score vividness was 71.66 with a pile-up of scores on the right of the distribution (range 30.75-98.63, SD = 16.82, N = 216, $z_{skewness}$ = -3.68, $z_{kurtosis}$ = -1.39). To establish that there was no detrimental effect of the skewed distribution on the analyses, the distribution of errors of the second and third level were inspected for the final models. No abnormalities were detected.

3

Emotionality over time between conditions

Memories in the RO condition were stable in emotionality over time, $\beta_{10} = -1.57$, $p = .340$. Contrary to expectations, memories in slow EM did not decrease emotionality when compared to RO, $\beta_{11} = -3.28$, $p = .158$. However, fast EM did result in a larger decrease of emotionality over time than RO, $\beta_{12} = -8.16$; $t(213) = -3.52$, $p = .001$: post-test scores were lower (predicted mean = 65.89) than pre-test scores (predicted mean = 74.04). Next, to test whether the fast EM condition differed from the slow EM condition, we analyzed the same model with slow EM as reference condition. This revealed that fast EM led to larger decreases in emotionality than did slow EM, $\beta_{12} = -4.87$; $t(213) = -2.10$, $p = .04$. This means that fast EM were superior to both RO and slow EM in decreasing the emotional intensity of target images. Finally, Table 1 (Model 1) summarizes the random components of the model. Emotionality ratings of the memories varied significantly across participants (σ^2_{v0k}), across memories within participants (σ^2_{u0jk}), and across time within memories within participants (σ^2_{u1jk}), ps <.001.

Vividness over time between conditions

Similar to the differences between conditions on emotionality, memories in RO showed stable vividness ratings over time, $\beta_{10} = -0.28$, $p = .885$, memories in slow EM did not decrease vividness compared to RO, $\beta_{11} = -3.91$, $p = .151$, while memories in fast EM yielded a significant difference compared to RO, $\beta_{12} = -8.63$; $t(213) = -3.18$, $p = .002$: post-test scores were lower (predicted mean = 62.89) than pre-test scores (predicted mean = 71.52; Model 2, Table 1). A re-run of the model with slow EM as reference condition revealed that fast EM showed a non-significant trend towards larger decreases in vividness ratings, $\beta_{12} = -4.72$; $t(213) = -1.74$, $p = .083$. So, it seems that images that were recalled while making fast EM decreased more in vividness than images that were recalled while making slow EM or were only recalled without dual task. Vividness ratings of the memories varied significantly across participants (σ^2_{v0k}), across memories within participants (σ^2_{u0jk}), and across time within memories within participants (σ^2_{u1jk}), ps <.001.

Difficulty over time between conditions

Likewise, the same pattern between the conditions was found for the difficulty of retrieving the target image. Memories in RO showed a stable score over time, $\beta_{10} = -1.57$, $p = .539$, slow EM did not increase difficulty more than RO, $\beta_{11} = 4.83$, $p = 0.182$, but fast EM did increase difficulty recalling the memory compared to RO, $\beta_{12} = 12.04$; $t(213) = 3.34$, $p = .001$: post-test scores were higher (predicted mean = 49.74) than pre-test scores (predicted mean = 37.70; Model 3, Table 1). A re-run of the model with slow EM as reference condition revealed that fast EM led to larger increases in difficulty than did slow EM, $\beta_{12} = -4.83$; $t(213) = 2.00$, $p = .046$. So, fast EM caused more difficulty in retrieving the image after intervention than both RO and slow EM. Again, difficulty ratings varied significantly across participants (σ^2_{v0k}), across memories within participants (σ^2_{u0jk}), and across time within memories within participants (σ^2_{u1jk}), ps < .001.

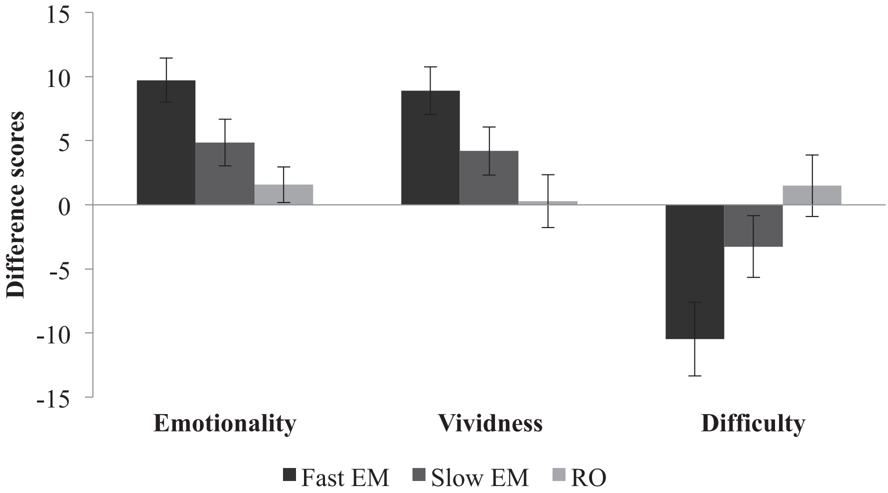


Figure 3. Mean difference scores (post-test minus pre-test) and SEs of fast EM, slow EM, and RO on emotionality, vividness and difficulty

Table 1. Fixed and random parts of Model 1 (emotionality over time between conditions), Model 2 (vividness over time between conditions), and Model 3 (difficulty over time between conditions)

		Model 1		Model 2		Model 3	
		Emotionality		Vividness		Difficulty	
Fixed part		β	S.E.	β	S.E.	β	S.E.
Intercept	β_{00}	72.59*	1.54	72.69*	1.97	40.17*	2.88
RO_slowEM	β_{01}	-2.31	1.87	-1.94	2.18	0.28	3.02
RO_fastEM	β_{02}	1.46	1.87	-1.17	2.18	-2.47	3.02
Time	β_{10}	-1.57	1.63	-0.28	1.92	-1.57	2.54
RO_slowEM	β_{11}	-3.28	2.32	-3.91	2.72	4.83	3.60
RO_fastEM	β_{12}	-8.16*	2.32	-8.63*	2.72	12.04*	3.60
Random part							
σ^2_{v0k}		45.99*		107.67*		267.98*	
σ^2_{u0jk}		125.74*		171.10*		328.41*	
σ^2_{u1jk}		193.09*		265.42*		467.83*	
Deviance		3457.23		3567.08		3819.20	

Note: In all models, RO was the reference condition, * $p < .05$

Interaction between pre-test vividness, condition and emotionality over time

Model 4 revealed that the coefficient of *pre-test vividness* x *slowEM_fastEM* in predicting the slope for time was -0.14 (SE = 0.14) and not significant, $t(210) = -1.05$, $p = .297$. So, inconsistent with our predictions, pre-test vividness did not interact with condition on changes in emotionality ratings.

Discussion

This study aimed to examine whether WM load of a dual-task carried out during memory image recall interacts with WM load of that memory on reducing its emotional intensity. We found a consistent pattern for all outcome measures: high WM taxation (recall + fast EM) was superior to low WM taxation (recall + slow EM) and no WM taxation (recall only; RO), and the effects of low taxation felt consistently in between the effects of high taxation and RO, but only differed significantly from high taxation. High WM taxation during recall produced memory images that were less vivid, less emotional and were more difficult to retrieve after the intervention. This is in line with WM theory: the more taxing a dual-task is, the more a memory image degrades. Crucially, image vividness did not interact with condition (high taxation vs. low taxation) with regard to the decrease of emotionality over time. Thus highly vivid and less vivid images showed the same responsiveness to dual-task manipulation: both memories benefitted the most from high WM taxation during recall.

The finding that recall + dual WM taxing reduced memory vividness and emotionality, compared to RO, is in line with a large body of experiments (see van den Hout and Engelhard, 2012, for an overview). More specifically, this study replicated the findings of Maxfield et al. (2008; experiment 2), who also found that fast EM (1.25Hz) yielded stronger reductions in memory image vividness and emotional intensity than slow EM (1.0Hz) and no EM. They contributed this difference in effects to presumed variation in WM taxation, but did not experimentally assess the WM taxation of both EM tasks. We extended their design and used RT methods to select two speeds of EM that significantly differed in WM taxation (study 1). We found the same superiority effects of fast EM compared to slow EM on image vividness and emotional intensity. Furthermore, we measured the difficulty of retrieving the memory image before and after the intervention and found that the higher the WM taxation, the more difficult it was to retrieve the memory image after intervention. Together these studies provide strong evidence for the WM theory in explaining the effectiveness of dual-task manipulation on memory modification. Low WM taxation produced memory effects in the same direction as high WM taxation; however, only high WM taxation was effective enough to produce memory effects that differed significantly from a control condition after a short intervention (6 x 24s).

The superiority of the 1.25Hz condition over the 1.0Hz condition in the study of Maxfield et al. (2008) suggests a linear relationship: higher WM taxing results in larger memory effects than lower WM taxing. However, according to the inverted U-curve hypothesis (Gunter & Bodner, 2008), strongest effects are found when competition between memory recall and the dual-task use approximately the same amount of WM resources. Too little taxation of the dual task will leave too many resources available for vivid memory recall and its accompanying emotions, while too much taxation of the dual task prevents the memory from being recalled. In a recent study, an inverted U-curve pattern was observed for emotionality, but not for vividness (Engelhard et al., 2011). In the current study, we examined whether the EM intervention would be more effective if the load of the dual task is matched with the load of the memory. We hypothesized that highly vivid memory images would benefit more from fast EM than from slow EM, and less vivid memory images would benefit more from slow EM than from fast EM. Contrary to these hypotheses, there were no interactions between image vividness and dual task WM taxation. Several explanations will be discussed.

First, it could be argued that slow EM were not sufficiently demanding and did not trigger the hypothetical threshold of the inverted U-curve. However, results of study 1 showed that slow EM tax WM more than no EM. Furthermore, EM with a speed of 0.8Hz had comparable WM taxation as EM with a speed of 1.0Hz. Because of these results, and because many laboratory studies have found memory effects with 1.0 Hz, which can be considered the “standard speed” (Lee & Cuijpers, 2013), the argument that slow EM were not taxing enough seems not plausible. The fact that in study 2 slow EM was attended by effects on memory that were in the same direction as fast EM, could indicate a dosage effect: the more cognitive demanding a dual task, the larger the memory effects. It could be hypothesized that an extended duration (e.g. more sets of recall+EM) would lead to a difference between slow EM and RO. For example, Leer, Engelhard, and van den Hout (2014) found that eight sets of recall with EM, compared to RO, caused a decrease in emotionality at a 24h follow-up test, while four sets did not.

There may be a second explanation for the absence of an interaction effect between dual task load condition and image vividness. Possibly, image vividness does not influence the amount of WM load. In the present study, WM load of the memory itself was not measured. However, the relation between WM and vividness of imagery was examined in series of experiments with dual task manipulations by Baddeley and Andrade (2000). It was concluded that vividness of imagery reflects the richness of representation in WM. Moreover, more recent evidence indicates that emotional memories tax WM to a greater extent than neutral memories (Van den Hout, Eidhof, Verboom, Littel, & Engelhard, 2014; discussion). Based on these previous studies, it seems justified to presume that image vividness affects the degree of WM taxation. In order to fully clarify this issue, it would be interesting to have participants recall images with a wide variation of vividness while performing a simple RT task. This would enable us to measure the cognitive demanding qualities of the target memory.

Alternatively, because WM load of the dual task did not interact with WM load of the memory image, one may question whether individuals are actually able to hold a memory image in mind while performing a dual task. The WM account is derived from the WM theory by Baddeley and Hitch (1974) in which three memory components are described: an attentional control system (central executive) and two slave storage systems (visuospatial sketch path and phonological loop). Later, Baddeley (2000) added a fourth component: the episodic buffer, which is a limited-capacity temporary storage system that allows integration from both the slave systems with material from long-term memory. The central executive is thought to control the retrieval and modification of information that is temporally stored in the episodic buffer. The central executive may therefore influence the content of information, by directing attention to a specific source: the slave systems or long-term memory. Based on this model, it seems likely that during a dual-task manipulation, the central executive is involved in attending to both tasks, while the temporal storage and integration of information takes place in the episodic buffer. During dual-task manipulation in our study, information is retrieved from long-term memory and maintains active in the episodic buffer. This process of constant reactivation to maintain an image active requires much effort (see Smeets, Dijs, Pervan, Engelhard, & van den Hout, 2012). A crucial question is whether performing EM *interferes* with the memory material due to integration of both tasks in the episodic buffer or whether division of attention between the two tasks by the central executive *inhibits* the memory material to be fully activated. If the former is true, then maximizing the complexity of a cognitive demanding task may leave almost no resources available for active recall of material from long-term memory and therefore there will be little interference. If the latter is true, then maximizing the complexity of a cognitive demanding task may lead to memory retrieval strategies, such as rapid shifting between tasks, which could lead to partial exposure to the memory and result in devaluation of the memory. More fundamental studies are needed to investigate these hypotheses about the cognitive processes that underlie the effects of dual task procedures on emotional memories.

Finally, there were some short-comings of the current study. First, we selected memories high or low in vividness, but the image vividness changed during the experiment prior to the intervention. This may have resulted in unreliable conditions. That is, selecting the target image seemed to inflate its vividness. We therefore analysed the data on the memory level instead of on the person level and used multilevel modelling to correct for the assumption violation of independent data. A strength of multilevel modelling is that it allowed the use of vividness as a continuous predictor, and therefore provides more detailed information than a dichotomous division in target image vividness. Second, the general ability to use mental imagery was not measured. Individual differences in imagery may influence the effectivity of dual task manipulation. Future studies could test this influence through assessment of the ability to use mental imagery with the Spontaneous Use of Imagery Scale

(SUIS; Reisberg, Pearson, & Kosslyn, 2003) or, more specified to visual imagery, the revised version of the Vividness of Visual Imagery Questionnaire (VVIQ-2; Marks, 1995). Third, vividness and emotionality ratings were based on subjective ratings. Psychophysiological measures could be used as objective indicator of memory emotionality (e.g. Barrowcliff, Gray, Freeman, & MacCulloch, 2004; Engelhard et al., 2010; Kearns & Engelhard, 2014, this issue). Fourth, the current study did not use standardised compliance measures: we manipulated the speed of the dot moving from left to right and corrected the participant if they did not follow the dot properly, but we did not test the actual speed of participants' eye movements. Using electrooculogram analysis might help here. Finally, we only analysed the immediate influence of dual-task manipulation on memory modification. Research has yet to determine whether memory modification effects are maintained over time (Leer et al., 2014).

In sum, we found consistent effect patterns that are in line with WM theory: the more cognitively demanding the dual task, the more an aversive memory can be modified, in that these images become less emotional, less vivid and more difficult to retrieve. In our study, WM load of the memory - operationalized by image vividness - did not interact with the WM load of the dual task. Therefore, we found no evidence for the inverted U-curve hypothesis proposed by Gunter and Bodner (2008). Further research is needed to critically test whether the inverted U-curve hypothesis does occur for other intra-individual variables, such as differences in WM capacity. Unraveling the complexities of WM theory may provide a better idea of how titration between the recalled memory image and the WM load of the dual task may be optimized.

Acknowledgements

This study was supported with a TOP-grant (40-00812-98-12030) from the Netherlands Organization for Health Research and Development (ZonMw) awarded to Marcel van den Hout. Iris Engelhard is supported with a Vidi grant (452-08-015) from the Netherlands Organization for Scientific Research (NWO). We thank Khatera Zamani and Rianne Westhof for their assistance with data collection.

Author contributions

MAvdH, IME, SCvV, KvS, and ML designed the study; SCvV collected the data, SCvV and LDNVWdM analyzed the data; SCvV drafted the manuscript and MAvdH, IME, KvS and ML provided critical revisions.

CHAPTER 4

BLURRING EMOTIONAL MEMORIES USING EYE MOVEMENTS: INDIVIDUAL DIFFERENCES AND SPEED OF EYE MOVEMENTS

K. VAN SCHIE

S. C. VAN VEEN

I. M. ENGELHARD

I. KLUGKIST

M. A. VAN DEN HOUT

European Journal of Psychotraumatology, 7(1), 29476 (2016)

doi:10.3402/ejpt.v7.29476

Abstract

Background: In eye movement desensitization and reprocessing (EMDR) patients make eye movements (EM) while recalling traumatic memories. Making EM taxes working memory (WM), which leaves fewer resources available for imagery of the memory. This reduces memory vividness and emotionality during future recalls. WM theory predicts that individuals with small working memory capacities (WMC) benefit more from low levels of taxing (i.e., slow EM) whereas individuals with large WMC benefit more from high levels of taxing (i.e., fast EM). **Objective**: We experimentally examined and tested four pre-specified hypotheses regarding the role of WMC and EM speed in reducing emotionality and vividness ratings: 1) EM − regardless of WMC and EM speed − are more effective compared to no dual task, 2) increasing EM speed only affects the decrease in memory ratings irrespective of WMC, 3) low WMC individuals − compared to high WMC individuals − benefit more from making either type of EM, 4) the EM intervention is most effective when − as predicted by WM theory − EM are adjusted to WMC. **Method**: Undergraduates with low (n = 31) or high (n = 35) WMC recalled three emotional memories and rated vividness and emotionality before and after each condition (recall only, recall + slow EM, and recall + fast EM). **Results**: Contrary to theory, the data do not support that EM speed should be adjusted to WMC (hypothesis 4). However, the data show that a dual task in general is more effective in reducing memory ratings than no dual task (hypothesis 1), and that a more cognitively demanding dual task increased the interventions' effectiveness (hypothesis 2). **Conclusions**: Although adjusting EM speed to an individual's WMC seems a straightforward clinical implication, the data do not show any indication that such a titration is helpful.

Keywords: visual imagery, dual taxation, working memory capacity, EMDR

Introduction

Mental imagery allows us to think about past or anticipated events and is a powerful process with which we can re-experience and recombine perceptual information from memory (Kosslyn, Ganis, & Thompson, 2001). However, at times, it becomes maladaptive, for instance, when unwanted memories of upsetting life events come to mind. Intrusive, recurrent memories are core symptoms of post-traumatic stress disorder, but can also occur in other psychiatric disorders, including obsessive-compulsive disorder, depression, body dysmorphic disorder, and several phobias (Hackmann & Holmes, 2004). These images can be past or future-oriented (e.g., Engelhard, van den Hout, et al., 2011).

In clinical practice, cognitive behavioral therapy is often used to reduce intrusive imagery in post-traumatic stress disorder (Deacon & Abramowitz, 2004). It includes techniques that encourage patients to repeatedly relive these images (i.e., imaginal exposure) or to confront the feared object or situation in real life for prolonged periods of time (i.e., in vivo exposure; Rothbaum, Meadows, Resick, & Foy, 2000). However, patients with PTSD may be reluctant to be exposed to their feared images for longer periods of time (Arntz, Tiesema, & Kindt, 2007).

A different technique for manipulating image vividness and emotionality is used in eye movement desensitization and reprocessing (EMDR). In EMDR, patients make eye movements (EM) while they simultaneously recall traumatic memories. Research has shown that making EM reduces self-reported ratings of vividness and/or emotionality of unpleasant autobiographical memories (Engelhard, van Uijen, & van den Hout, 2010; Kavanagh, Freese, Andrade, & May, 2001; Maxfield, Melnyk, & Hayman, 2008; van den Hout, Eidhof, Verboom, Littel, & Engelhard, 2014; for a meta-analysis of patient studies and analogue studies, see Lee & Cuijpers, 2013), but also reduces the vividness and emotionality of imagined feared future events (i.e., flash-forwards; Engelhard, van den Hout, Janssen, & Van der Beek, 2010; Engelhard, van den Hout, et al., 2011). Moreover, other secondary tasks besides EM also reduce image vividness and/or emotionality, including drawing a complex figure (Gunter & Bodner, 2008), playing Tetris (Engelhard, van Uijen, & van den Hout, 2010), arithmetic (van den Hout et al., 2010; Engelhard, van den Hout, & Smeets, 2011), and complex tapping (Andrade et al., 1997).

How secondary tasks reduce vividness and emotionality of mental images can be conveniently explained by the interplay of dual taxation of working memory (WM; e.g., Andrade et al., 1997; Gunter & Bodner, 2008) and destabilization induced by memory reactivation (Lewis, 1979). A previously consolidated memory that is recalled (i.e., reactivated) can become labile and sensitive to disruption. When at the same time an individual performs a secondary task (e.g., making EM) dual taxation of WM takes place. Both tasks compete for limited WM resources and therefore the distressing memory cannot be retrieved completely (i.e., gets blurred). It is suggested that as a consequence of its temporary labile state the blurred memory reconsolidates after competition, and the reconsolidated blurred memory will be retrieved during future recalls (see van den Hout & Engelhard, 2012).

WM theory predicts that the effectiveness of dual taxation depends on an individual's working memory capacity (WMC). For competition to occur, it is necessary that both tasks (mental image activation and secondary task) are sufficiently taxing. Individuals with a large WMC may be able to perform memory recall and a secondary task simultaneously without much competition between these tasks. As a consequence, they may experience fewer benefits (i.e., less blurring) from performing the dual task, compared to individuals with a relatively small WMC. Findings from Gunter and Bodner (2008, experiment 3) suggest that this is indeed the case: WMC – as measured with the automated reading span – correlated negatively with decreases in self-reported memory vividness/emotionality as a consequence of dual taxation. Other similar correlational findings support this (van den Hout et al., 2010; van den Hout, Engelhard, Beetsma, et al., 2011; Engelhard, van Uijen, & van den Hout, 2010). Moreover, the theory predicts that the dual task should not be too easy (this leaves too much capacity for the memory) or too hard (then the memory can hardly be recalled); the optimal load lies in between. There is preliminary evidence for this inverted U-curve shape: mild and moderately taxing of WM resulted in larger drops in emotionality ratings compared to little or extreme taxing (Engelhard, van den Hout, & Smeets, 2011). This suggests that dual taxation of WM would be more effective if the degree of taxing is adjusted to an individual's WMC. More specifically, the theory predicts the presence of inter-individual differences; individuals with a relatively small WMC benefit more from relatively low levels of taxing and individuals with a relatively large WMC benefit more from relatively high levels of taxing. An interaction between WMC and WM taxation has obvious clinical implications. It would suggest that, in clinical practice, the degree of WM taxation (i.e., speed of EM) should be adjusted to the WMC of the patient treated. This is especially relevant because PTSD has been linked with poor performance on WMC measures (e.g., Samuelson et al., 2006). Since tasks measuring WMC are widely used and validated (Conway et al., 2005), it would be possible to determine a patient's individual WMC before the dual task intervention in EMDR is started.

Though WM theory predicts increased effectiveness when dual taxation is adjusted to individual differences, this might not necessarily be the case as suggested by an analogous study investigating intra-individual differences in memory vividness. Van Veen et al. (2015) inferred from WM theory that the effectiveness of dual taxation depends on intra-individual differences in memory vividness. They tested whether emotionality ratings of highly vivid memories (i.e., memories that are more taxing on WM) are more effectively reduced by a highly taxing secondary task (i.e., fast EM), and emotionality ratings of memories low in image vividness are more effectively reduced by an easy dual task (i.e., slow EM). Inconsistent with this prediction, there was no interaction with memory vividness. Emotionality ratings were reduced only as a result of increased dual taxation of WM; a high load dual task led to higher reductions compared to a low load dual task. This is in accordance with an earlier study by Maxfield, Melnyk, and Hayman (2008), that also showed fast EM were more effective than slow EM, and that those were more effective than a control condition in decreasing image vividness and emotionality ratings. Therefore, there may only be a gradual effect of increasing dual taxation without an effect of WMC or load of the dual task.

All in all, it is possible that individuals with a small WMC benefit more from dual taxation in general compared to individuals with a large WMC as frequently evidenced in correlational studies. Alternatively, WM theory predicts that taxation is more effective if it is adjusted to an individual's WMC or to memory vividness. Previous research, however, showed that increasing dual taxation of WM increases effectiveness, regardless of intra-individual differences in memory vividness. It is therefore possible that only the dual taxation of WM determines its effectiveness irrespective of inter-individual differences in WMC.

Therefore, we tested four pre-specified hypotheses regarding the role of individual WMC and dual taxation in reducing emotionality and vividness ratings. Within each hypothesis, we expect emotionality and vividness to display similar patterns. Our first hypothesis tested if making EM –regardless of EM speed – is more effective in reducing memory ratings compared to a control condition. The second examined if the speed of making EM solely affects the decrease in memory ratings irrespective of WMC. The third tested if low WMC individuals – compared to high WMC individual – benefit more from making EM in general, and our final hypothesis examined if reductions in memory ratings are highest when EM are adjusted to an individual's specific WMC. All hypotheses are visually presented in figure 1 and exact hypothesis constraints can be found in the *Appendix*. We used an experimental design – in which we manipulated speed of EM (i.e., load of the dual task) and tested individuals with low and high WMC – and we used a Bayesian approach to critically test which of the hypotheses is most likely.

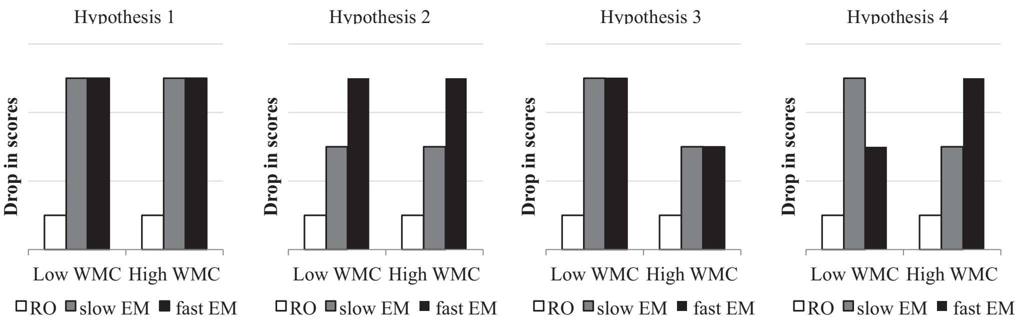


Figure 1. Visual representations of all four hypotheses for the Low and High WMC group after Recall Only (RO), slow EM, and fast EM

Method

Participants

Prior to participation undergraduate and graduate students from Utrecht University and the University of Applied Sciences (Hogeschool Utrecht) were screened for knowledge about the working mechanism of EMDR, participation in previous EMDR experiments, and medication intake that affected memory or concentration (N = 166). Based on these exclusion criteria, 18 individuals were excluded from further participation. One hundred forty-eight participants (M_{age} = 20.28, SD_{age} = 2.36; 64 males, 84 females) completed a WMC test for course credit or financial reimbursement. Thirteen of them were excluded because they failed to adequately perform the WM task (see below), or did not want to participate in the second part of the study, leaving a final sample of 135 participants. Via a tertiary split three groups were created. Individuals with the lowest and highest WMC were invited for the second part of our study and performed a dual-taxation task. Of the 90 participants invited, 37 low WMC individuals and 38 high WMC individuals were willing to participate further. After the dual-taxation task, six low WMC individuals and three high WMC individuals were excluded from final data-analysis, because they refused to make EM (n = 1), failed to recall three emotional memories (n = 2), or displayed emotionality and/or vividness scores that were extremely improbable and are indicative of task non-compliance (e.g., identical pre and post scores to two decimal places; n = 6). Our final sample consisted of 31 and 35 participants, in the low and high WMC groups respectively (M_{age} = 20.42, SD_{age} = 2.38; 25 males, 41 females). The Ethical Committee of the Faculty of Social and Behavioural Sciences of Utrecht University

[FETC14-008 (Hout)] approved this study. Written informed consent was obtained from all participants.

Materials & Procedure

Automated Reading Span. Participants first completed the automated reading span (Conway et al., 2005; Daneman & Carpenter, 1980), which was used to assess WMC. Before the experimental trials, participants performed three practice sessions. In the first session, participants viewed two or three letters presented individually on screen for 1000ms. These letters were a subset of letters taken from 12 available letters: F, H, J, K, L, N, P, Q, R, S, T, and Y. Participants were instructed to remember the letters in the order in which they were presented. After a set of presented letters, participants were instructed to recall and indicate the letters they had just seen, in the same order, by clicking on letters presented in a 4 × 3 letter matrix. After each trial, the number of correctly selected letters was presented as feedback.

In the second session, participants read 15 sentences (e.g., "The young pencil kept his eyes closed until he was told to look.") and decided after each sentence whether that sentence made sense by clicking a TRUE or FALSE box. After each decision, they received feedback on their accuracy. The program calculated each individual's mean decision time. This time plus 2.5 standard deviations served as maximum response latency for sentence evaluations in the experimental trials.

In the third session, participants practiced with both making sentence evaluations and remembering letters. Participants read a sentence and made an evaluation. This was immediately followed by one letter and the instruction to remember it. In this phase, trials comprised of two sentence-letter sequences (i.e., a set size of two). Thus, set size corresponded with the number of letters that had to be recalled at the end of a trial. After each trial, the letter matrix (cf. practice session 1) was shown, and participants recalled the letters they had just seen, in the same order. There were three practice trials with a set size of two.

Experimental trials were similar to trials in the third practice session except that the set size varied from three to seven, and set sizes were randomly presented. In total, 75 sentences and letters were presented in the experimental phase (three trials of each set size). Half of the sentences made sense, and half did not.

For scoring the automated reading span, we used partial-credit unit (PCU) scoring, in which credit is given to partly correct items as opposed to all-or-nothing unit scoring, where credit is only given to completely correct items (i.e., items where *all* letters were recalled in the correct order). That is, for a trial with a set size of four, two omissions followed by two correctly selected letters still constituted a score of 0.5 with PCU scoring. The PCU score expresses the proportion of correctly recalled letters within a trial averaged over all trials (see Conway et al., 2005, for a discussion on scoring in complex span tasks). Partial scores in the automated reading span show good test-retest reliability (r = .82) and internal consistency (α = .86-.88), and the automated reading span correlates strongly with other complex span tasks measuring WMC (Conway et al., 2005; Redick et al., 2012).

Dual-Taxation Task. After a tertiary split on WMC, the highest and lowest groups performed a dual-taxation task. A participant was instructed to recall three negative autobiographical memories. These memories were rated on emotionality (0 *not unpleasant* to 100 *very unpleasant*). The emotionality score had to be in the range 50-90; if this was not the case, the participant was asked to recall a different memory. Next, the three memories were ranked from highest to lowest based on the participant's emotionality ratings. Then, for each memory, the participant selected the worst mental image, which served as that memory's "hotspot". During this selection procedure, we counterbalanced which mental image was selected first, second, and third. For instance, some participants started with the memory ranked first, followed by the memory ranked second, and finally the memory ranked third (i.e., a 1-2-3 sequence). To avoid order effects, other counterbalancing sequences were used equally often (i.e., 1-3-2, 2-1-3, 2-3-1, 3-2-1, and 3-1-2). Next, the participant wrote down a label for each memory's "hotspot", which served to refer to that specific mental image. After mental image selection, each of the participant's three mental images was assigned to one of the three conditions: recall + fast EM, recall + slow EM (henceforth called "fast EM" and "slow EM"), and recall only. So, a participant performed all conditions. Again, assignment to conditions depended on counterbalancing in a similar way as described for the selection phase. Counterbalancing assured that for all participants all three memories were equally often assigned to each of the conditions. The conditions (fast EM, slow EM, and recall only) were presented randomly to a participant. Before each condition, participants were presented with the label of their mental image and were instructed to recall the mental image. They then rated their memory on a visual analogue scale (VAS) that ranged from 0 (*not vivid/unpleasant*) to 100 (*very vivid/unpleasant*). In each condition, participants recalled the memory for 6 intervals of 24 seconds separated by 10-second breaks (Engelhard et al, 2012; van Veen et al., 2015). In both EM conditions participants were seated approximately 45 cm from the computer screen and were instructed to recall the memory and to simultaneously follow a 20 pixel dot that moved horizontally with their eyes (600 pixel amplitude on a 1280×1024 pixel screen). For the slow EM condition, a dot moved with 0.8Hz across the screen, and for the fast EM condition this was 1.2Hz. A 1Hz cycle corresponds with one left-right-left movement within 1 second. Van Veen et al. (2015) showed that these two speeds differed significantly in WM taxation. After each condition, participants again recalled the memory and rated it on vividness and emotionality using VASs. Groups were tested double-blind.

Data analysis

The hypotheses were evaluated using a Bayesian model selection criterion based on the Bayes factor (BF; Kass & Raftery, 1995) that was analyzed with the software BIEMS (see Mulder, Hoijtink, & Klugkist, 2010; Mulder, Hoijtink, & de Leeuw; 2012; Mulder et al., 2009). Contrary to null hypothesis significance testing, the Bayesian framework is not based on *p*-values, dichotomous decisions (i.e., the result is significant or not), or the assumption that a null hypothesis is true. The Bayesian approach uses the observed data and computes support for each hypothesis given all constraints specified in each hypothesis. This approach can also be used to evaluate competing hypotheses. Thus, the calculated BF states the likelihood of a specified hypothesis. The program BIEMS specifically computes BFs for constrained hypotheses against the unconstrained hypothesis. A BF of 1 means that compared to an unconstrained model, the hypothesis receives equal support. BF > 1 indicates that the hypothesis outperforms the unconstrained model, and BF < 1 means the opposite.

Results

Automated Reading Span: Working Memory Capacity Scores

After the tertiary split, the group with the low WMC group had a mean PCU score of .68, 95% CI [.65, .71] and the high WMC group had a mean score of .91, 95% CI [.90, .92]. These mean scores correspond with the 30th and 75th percentile for the low and high group respectively (Redick et al., 2012).

Bayesian Analysis on Reductions in Vividness and Emotionality Ratings

For each participant, a pre-post change score was calculated per condition for vividness and emotionality ratings; with higher scores indicating a greater pre-post drop.[2] Bayesian analyses showed BFs of 3.02 and 4.02 for vividness and emotionality for hypothesis 1, 3.28 and 4.00 for hypothesis 2, 0.99 and 2.58 for hypothesis 3, and 0.11 and 1.08 for hypothesis 4. Overall, this shows that given the data, models 1 and 2 appear more likely than models 3 and 4 (see Table 1).

[2] Change scores with more than three SD from the group mean were corrected to a score with three SD from the group mean. In total there were 396 difference scores of which 6 were corrected.

Table 1. Bayes Factors for Vividness and Emotionality for all four Hypotheses

	Hypothesis 1	Hypothesis 2	Hypothesis 3	Hypothesis 4
Vividness	3.02	3.28	0.99	0.11
Emotionality	4.02	4.00	2.58	1.08

Observed Reductions in Vividness and Emotionality Ratings

Figure 2 and Table 2 show that the observed data patterns are indeed moderately in line with hypothesis 1 and 2, but not with hypothesis 3 and 4 [note that in the hypotheses with between group comparisons, differences in recall only (i.e., control condition) are taken into account when comparing slow and fast EM conditions]. However, the observed scores seem to display an unexpected pattern that was not hypothesized. In the low WMC group, fast EM achieved the largest change scores, and in the high WMC group, slow EM and fast EM seem to be equally effective in reducing memory ratings. It is also worth noting that the low and high WMC group recall only conditions differ substantially: the low WMC group showed a decrease in scores, while the high WMC group showed an increase.

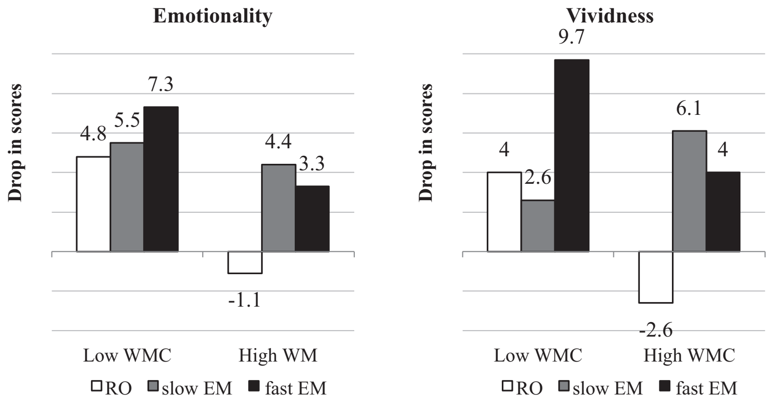


Figure 2. Observed reductions (pre-post difference scores) in Emotionality and Vividness for Low and High WMC group after Recall Only (RO), slow EM, and fast EM.

Table 2. Mean Raw Scores and Standard Deviations on Emotionality and Vividness before and Recall Only (RO), slow EM, and fast EM for the Low and High WMC group

	Emotionality			
	low WMC		high WMC	
	pre	post	pre	post
fast EM	68.87 (*14.40*)	61.59 (*22.05*)	68.16 (*16.93*)	64.48 (*17.90*)
slow EM	71.24 (*16.37*)	65.62 (*21.93*)	70.74 (*17.47*)	66.31 (*18.09*)
RO	76.32 (*16.03*)	71.52 (*16.60*)	71.19 (*16.44*)	72.17 (*16.72*)
	Vividness			
	low WMC		high WMC	
	pre	post	pre	post
fast EM	73.25 (*22.42*)	63.51 (*26.03*)	71.37 (*20.51*)	67.38 (*20.05*)
slow EM	71.22 (*23.42*)	68.66 (*15.09*)	73.78 (*20.26*)	68.03 (*20.57*)
RO	80.22 (*14.95*)	76.11 (*20.27*)	75.96 (*17.55*)	78.48 (*16.03*)

Note. These scores are not corrected for outliers. As a result, there are small deviations with the difference scores that were used for the analyses; WMC = Working Memory Capacity

Discussion

The aim of this study was to test if and how inter-individual differences in WMC affect self-reported memory ratings after recall of an emotional memory under different dual taxation conditions. Using a Bayesian approach, we compared four hypotheses, which specified the relation between WMC and WM taxation. Two of these hypotheses (hypothesis 1 and 2) were equally supported by the data. The data show that a dual task in general is more effective than no dual task, and that effectiveness increases with more cognitively demanding dual taxation. Moreover, it seems that individuals with smaller WMC do not benefit consistently from dual taxation compared to individuals with larger WMC (hypothesis 3), and that the hypothesis about adjusting WM taxation to WMC – as predicted by WM theory – (hypothesis 4) received very little support from the data and was therefore the least likely of all hypotheses. Unexpectedly, as was evident from the observed reductions, the low WMC group showed the largest decrease after a high load dual task, while the high WMC group showed similar decreases in memory ratings after either type of dual task.

The finding that an EM dual task in general (i.e., slow or fast EM) works better than no dual task (i.e., recall only) – as stated in hypothesis 1 – joins a corpus of data showing that EM can be used as an effective dual task (see Lee & Cuijpers, 2013), and that any dual task that taxes WM is effective in reducing memory ratings of vividness and/or emotionality (e.g., Andrade et al., 1997; Engelhard, van Uijen, van den Hout, 2010; Engelhard, van den Hout, & Smeets, 2011; van den Hout et al., 2010). Evidently, competition for WM resources between a dual task and recall makes memories less vivid and less emotional during future recalls. Moreover, our study nicely fits with two other studies that have shown that a more cognitively demanding dual task results in larger decreases in memory ratings, and that different speeds of EM can be used to tax WM differentially (Maxfield et al., 2008; van Veen et al., 2015). Our study replicates these findings and shows that a higher speed of EM leads to higher drops in memory ratings, as was evidenced by the support from the Bayes factor for hypothesis 2. The evidence in favor of hypothesis 3 was not only considerably smaller compared to hypothesis 1 or 2, but it was also mixed. The discrepancy between Bayes factors for emotionality and vividness makes interpretation of these results difficult. However, a change in emotionality, but not in vividness or vice versa is not unique (e.g., Andrade et al., 1997; Engelhard, van Uijen, & van den Hout, 2010; Maxfield et al., 2008). It is currently still unclear why the evidence is not consistent for both measures.

Though we found larger drops in vividness and emotionality after more cognitively taxing dual-tasks using *digitalized* EM, these effects most likely translate well to clinical practice. Meta-analysis has shown that comparable effects have been found for digitalized EM and therapist-driven EM (Lee & Cuijpers, 2013). Moreover, from a theoretical point of view, the method of application is irrelevant as long as it sufficiently taxes WM. Indeed, other secondary tasks besides EM also reduce image vividness and/or emotionality, such as complex spatial tapping, playing Tetris, and mental arithmetic (Andrade et al., 1997; Engelhard, van Uijen, & van den Hout, 2010; van den Hout et al., 2010). Therefore, any intervention that sufficiently taxes WM should be effective.

The question remains why we did not find an interaction of WMC and WM taxation that was hypothesized based on WM theory (hypothesis 4). Possibly the range of WMC of the university students recruited here was low and the tertiary split may not have resulted in between group differences that were sufficiently large. This nevertheless seems unlikely, because even in comparable student samples inverse correlations have been found between WMC and memory ratings after dual taxation (Gunter and Bodner, 2008; van den Hout et al., 2010; van den Hout, Engelhard, Beetsma, et al., 2011; Engelhard, van Uijen, & van den Hout., 2010). Moreover, the average WMC scores for the low and high group corresponded with the 30th and 75th percentile, respectively (Redick et al., 2012), which shows that our groups differed meaningfully.

WM theory clearly predicted that inter-individual differences in WMC and different conditions of WM taxation should interact (Gunter & Bodner, 2008; Engelhard, van den Hout, & Smeets, 2011). However, this study shows that there is little evidence that WM taxation is more effective when the degree of taxing is adjusted to inter-individual differences in WMC. This is contradictory to studies suggesting an inverted U-curve in terms of WM taxation (Gunter and Bodner, 2008; Engelhard, van den Hout, & Smeets, 2011) or to correlation studies (e.g., Gunter and Bodner, 2008; van den Hout et al., 2010; van den Hout, Engelhard, Beetsma, et al., 2011) that show that a standard speed of EM is most beneficial for individuals with smaller WMC, and that by extension higher speeds should be used for individual with large WMC. The inconsistency in our study is not isolated, but joins another found in van Veen et al. (2015). They showed that intra-individual differences in memory vividness do not interact with WM taxation, though this was also hypothesized based on WM theory. The fact that in our study there was no trace of evidence suggesting an interaction between WMC and WM taxation serves as an anomaly for WM theory.

Still, the observed data shows patterns that may be reconciled with WM theory. We found that for the low WMC group, fast EM outperformed the other conditions, while for the high WMC group either EM condition outperformed a no dual task control. Small reductions after slow and fast EM for the high WMC group may be the result of a limitation of the amount of WM taxation from EM as a dual task for this group specifically. It is imaginable that fast EM in the study by Maxfield et al. (2008) and van Veen et al. (2015) – and in our experiment – were the most taxing given the constraints determined by our muscular system, but that for the high WMC group it may not have been cognitively taxing enough to achieve substantial drops in memory ratings. This problem of too little taxation for high WMC individuals could be circumvented by increasing taxation by adding another task to EM that increases WM taxation as a whole, such as arithmetic (Engelhard, van den Hout, & Smeets, 2011; van den Hout et al., 2010; van den Hout, Engelhard, Rijkeboer, et al., 2011). Alternatively, EM could be substituted by this task altogether to avoid constraints associated with making EM for prolonged periods of time. Arithmetic would be ideal, because several factors of calculations can be manipulated to vary WM taxation (e.g., required operations or problem complexity; DeStefano & LeFevre, 2004).

In sum, we found that the data are in line with some predictions from WM theory: a dual task in general was effective in reducing memory ratings than no dual task, and a more cognitively demanding dual task increased the interventions' effectiveness. However, differently taxing dual tasks did not interact with differences in WMC as hypothesized. This anomaly suggests that titration based on WMC does not increase the effectiveness of the dual task intervention. Adjusting WM load (i.e., EM speed) to the WMC of individual patients appears a straightforward clinical implication from WM theory for treatment with EMDR, but the data do not show any indication that such a titration is helpful. Based on our results and a study by van Veen et al. (2015), the only clinical implication that follows is that increasing speed of EM, increases the intervention's effectiveness.

Acknowledgements

This research was supported by a TOP grant (dossier number: 40-00812-98-12030) from the Netherlands Organization for Health Research and Development (ZonMw) awarded to MAvdH. IME is supported with a Vidi grant (452-08-015), and IK is supported with a Vidi grant (452-12-010) from the Netherlands Organization for Scientific Research (NWO). We are grateful to Annalyne van Oordt-Florentinus and Anne Tiemensma for their assistance in data collection. We thank Arne Leer for helpful suggestions and manuscript feedback.

Author contributions

MAvdH, IME, KvS, and SCvV designed research; KvS collected data; KvS and IK analyzed data; and KvS, SCvV, MAvdH, IME, and IK wrote the paper, and have read and approved the final manuscript.

CHAPTER 5

EXPLORING EFFECTS OF DUAL-TASKS ON REDUCTION OF INTRUSIVE MEMORIES FOLLOWING ANALOGUE TRAUMA

K. VAN SCHIE

S. C. VAN VEEN

M. A. HAGENAARS

I. M. ENGELHARD

In preparation

Abstract

Performing a dual-task such as making eye movements (EM) while recalling the 'hotspot' of an unpleasant memory reduces subjective vividness and unpleasantness of the memory hotspot. Competition for limited working memory resources presumably leaves fewer resources available for maintaining a vivid memory, which affects how it is later recalled. We hypothesized that reduction of the memory hotspot's vividness and unpleasantness also reduces the number of intrusive memories. Additionally, we explored whether this was dependent on dual-task modality. To test this, participants watched a trauma film followed by one of three dual-task interventions (1) Recall+EM, (2) Recall+Counting, or (3) No-Task Control. The duration of these interventions was short (exp1) or long (exp2). Before and after each intervention, participants rated vividness and unpleasantness of the memory hotspot. In the following week, they recorded intrusive memories about the film in an intrusion diary. We found that prolonged dual-task interventions in experiment 2, regardless of modality, resulted in a lower number of intrusions compared to no-task control, but did not affect vividness and unpleasantness ratings. To confirm these tentative findings, and to rule out that intrusion modulation by dual-task interventions is not merely the result of memory recall (instead of memory recall+dual-task), a third experiment that incorporates a Recall Only control condition needs to be performed.

Keywords: intrusive memory; working memory; trauma film paradigm; dual-task; visual imagery; EMDR

Introduction

Most people experience a traumatic event at some point in their life; an event in which an individual is exposed to actual or threatened death, serious injury, or sexual violence (DSM-5; American Psychiatric Association, 2013). Approximately 7% of people that experience a traumatic event will develop post-traumatic stress disorder (PTSD) (e.g., de Vries & Olff, 2009; Kessler, Sonnega, Bromet, Hughes, & Nelson, 1995). One of the core symptoms of PTSD is intrusive, recurrent memories of the traumatic event, which typically take the form of visual (Ehlers et al., 2002; Hackmann, Ehlers, Speckens, & Clark, 2004) and other sensory images (e.g., auditory; Engelhard, van den Hout, Arntz, & McNally, 2002).

Eye Movement Desensitization and Reprocessing (EMDR) can be used to reduce intrusive imagery in PTSD. In EMDR, patients make eye movements (EM) by tracking a therapist's finger that moves horizontally in front of the patient's eyes while they simultaneously recall the "hotspot" of their traumatic memory. According to an influential 'working memory' theory (see Andrade Kavanagh, & Baddeley, 1997; Gunter & Bodner, 2008; van den Hout & Engelhard, 2012), EM and memory recall compete for limited working memory (WM) resources during the intervention. This competition impedes memory retrieval, which reduces its vividness and/or unpleasantness. Van den Hout and Engelhard (2012) have suggested that the memory, which is now reduced in vividness and emotional intensity, is re-stored in long-term memory (i.e., reconsolidation), and is subsequently recalled during future recalls.

A large body of research has shown that making voluntary, horizontal EM simultaneously with memory recall indeed reduces self-reported memory vividness and/or unpleasantness (e.g., Andrade et al., 1997; Engelhard, van Uijen, & van den Hout, 2010b; Kavanagh, Freese, Andrade, & May, 2001; Maxfield, Melnyk, & Hayman, 2008; van den Hout, Eidhof, Verboom, Littel, & Engelhard, 2014; van Schie, van Veen, Klugkist, Engelhard, & van den Hout, 2016; van Veen et al., 2016; for a meta-analysis of clinical studies and analogue studies, see Lee & Cuijpers, 2013). Making EM does not only affect the vividness and/or unpleasantness of autobiographical memories; effects are also present for novel non-idiosyncratic materials, such as neutral and negative pictures (Andrade et al., 1997, van den Hout, Bartelski, & Engelhard, 2013), prospective memories (e.g., Engelhard, van den Hout, Janssen, & van der Beek, 2010a), positive memories (Engelhard et al., 2010b), memories acquired during fear conditioning (Leer, Engelhard, Altink, & van den Hout, 2013; Leer et al., 2017), and memories acquired in virtual reality environments (Cuperus, Laken, van den Hout, & Engelhard, 2016). Furthermore, the effects are not specific to the horizontal EM intervention: other dual-tasks that sufficiently tax WM are equally effective, including making voluntary vertical EM (Gunter & Bodner, 2008), drawing a complex figure (Gunter & Bodner, 2008), playing Tetris (Engelhard et al., 2010b), counting backwards (van den Hout et al., 2010; Engelhard, van den Hout, & Smeets, 2011), attentional breathing (van den Hout et al., 2010), and complex tapping (Andrade et al., 1997).

One area of debate in the literature is whether visual processing specifically increases the effect of WM loads on visual imagery (e.g., Andrade et al., 1997), or whether load per se (Gunter & Bodner, 2008) is the important factor. According to Andrade et al. (1997) WM's visuospatial sketchpad (VSSP) stores visuospatial information, and this is the place where memories are held in mind during recall. Simultaneously making eye movements uses the same VSSP resources, which then results in less vivid memories. Alternatively, Gunter and Bodner (2008) argue that benefits of performing a dual-task (e.g., eye movements) occur at the central executive level, which is a WM component that is involved when attention needs to be divided. From the former explanation follows that dual-tasks need to have visuospatial properties to create competition in WM, while according to the latter a specific modality is not required; any generally taxing dual-task suffices.

There is evidence for both accounts; some studies only show effects for visuospatial dual-tasks (e.g., Andrade, Pears, May, & Kavanagh, 2012; May, Andrade, Panabokke, & Kavanagh, 2010), while others show it is not modality per se, but general WM taxation (e.g., Engelhard et al., 2011; Tadmor, McNally, and Engelhard, 2016; van den Hout et al., 2010). There is, however, no direct comparison yet that controlled for degree of cognitive load of visuospatial and non-visuospatial dual-tasks. Having participants perform a Random Interval Repetition (RIR) task (Vandierendonck, De Vooght, & Van der Goten, 1998) can assess cognitive load. In a RIR task participants are required to respond as fast as possible to a randomly administered stimulus (e.g., mild electrical stimulus) with or without performing a dual-task. The slowing down of reaction times in the dual-task condition compared to no dual-tasks provides a valid and highly sensitive measure of cognitive load. However, both accounts may not be mutually exclusive, because some studies have shown general dual-task load effects on visual and auditory images (Baddeley & Andrade, 2000), and modality specific effects on top of effects of general dual-task load (Kemps & Tiggeman, 2007; Matthijssen, van Schie, & van den Hout, 2017).

It is also unclear whether dual-task interventions affect *involuntary* experienced memories, such as intrusive memories. This is an important question, because how WM taxation causes PTSD symptom reduction is still a matter of debate (Gunter & Bodner, 2009).

Theoretically, dual-tasks may have the potential to reduce intrusive memory. Dual-tasks target a voluntarily recalled memory hotspot and this makes the memory hotspot less vivid and less unpleasant. Research with PTSD patients has shown that the majority of intrusive images reported are related to this hotspot (Holmes, Grey, & Young, 2005; Grey & Holmes, 2008). Therefore, dual-tasks that reduce the excitable nature of the hotspot may also reduce the number of intrusive images. According to influential theories of PTSD, highly visual intrusive memories are indeed easily triggered by other stimuli (e.g., Ehlers & Clark, 2000).

A paradigm that has been frequently used to experimentally investigate intrusion development is the trauma film paradigm (e.g., Horowitz, 1969; Holmes & Bourne, 2008; James et al., 2016a). In a typical experiment, participants view a film of approximately 10 minutes depicting traumatic events (e.g., injury or death), which results in a novel ´traumatic´ memory. Interventions to prevent intrusion development in the paradigm typically focus on the manipulation of cognitive processes during or after film viewing. It is important to note that interventions aimed at intrusion modulation within the trauma film paradigm have been exclusively performed *without* simultaneous memory recall (i.e., single-task interventions; Holmes et al., 2009; 2010; for a review see James et al., 2016a), while interventions aimed at hotspot modulation by use of dual-task interventions have always incorporated memory recall (e.g., van den H out & Engelhard, 2012). After the film and/or the manipulations, participants record their intrusive images of the film in a diary for a week.

Interventions in the trauma film paradigm initially focused on visuospatial modality of intrusive memories. Broadly speaking, theories of PTSD suggest that intrusions develop, because individuals selectively process sensory information of the event, with little verbal or conceptual processing (Brewin, 2001; Brewin & Holmes, 2003; Brewin, Dalgleish, & Joseph, 1996; Ehlers & Clark, 2000). In line with this, studies have shown that visuospatial interference tasks (e.g., playing Tetris, finger tapping) reliably reduce intrusive images (Deeprose, Zhang, DeJong, Dalgleish, & Holmes, 2012; Holmes, James, Kilford, & Deeprose, 2010). However, the effects of verbal-conceptual interference tasks (e.g., counting, playing a Pub quiz) are rather mixed. Compared to a no-task control condition some studies show more intrusions (Bourne, Frasquilho, Roth, & Holmes, 2010, exp 2; Holmes et al., 2010, exp 1), while other show less intrusions (Hagenaars, Holmes, Klaassen, & Elzinga, 2017; Krans, Näring, & Becker, 2009) or no difference between conditions (Deeprose et al., 2012; Holmes et al., 2010, exp 2; for overviews see Brewin, 2014; James et al., 2016a). Interestingly, a study by Pearson and Sawyer (2011) showed that regardless of modality (visuospatial vs. verbal), only high (vs. low) executive load tasks at the time of encoding of aversive pictures reduced intrusions of these images in the following week. Overall, the research on the modality specific modulation of intrusions suggests that decreasing intrusion frequency may rely more on a general effect of taxation, than on modality specific taxation.

Later adaptations of PTSD theories put less emphasis on modality and stress the importance of attentional resources that are relevant for conceptual processing of trauma information (e.g., Brewin & Burgess, 2014). Thus, both WM as PTSD theories suggest the presence of two distinct modalities (i.e., visuospatial/visual vs. verbal/conceptual), but it is not clear to what extent these modalities affect the development of involuntary trauma memories. Because research on dual-task interventions concentrates primarily on voluntary memory retrieval, the first aim of the current research was to investigate whether dual-task interventions reduce involuntary intrusive images using the trauma film paradigm. The second aim was to examine whether modality of the dual-task has an effect on intrusion frequency. Therefore we compared two dual-tasks: visuospatial (EM) and verbal (counting). Finally, we tested whether decreases in vividness and unpleasantness were stronger following dual-tasks interventions (compared to a no-task control condition) and whether these decreases predicted lowered intrusion frequency. Participants viewed a trauma film and were then randomly assigned to one of three interventions: Recall+EM, Recall+Counting, or No-Task Control. The nature of voluntary memory was assessed by participants' self-reported vividness and unpleasantness before and after the dual-task intervention. Involuntary memory was operationalized by the number of intrusions participants recorded in their diary in the week following the film.

Method, Experiment 1

Participants

Seventy-six participants partook in this study (22 males, 54 females; M_{age} = 21.1 years, Age range = 18 to 26 years) with 25, 26 and 25 participants, respectively, in Recall+EM, Recall+Counting, and No-Task Control groups. They were students from Utrecht University and the University of Applied Sciences (Hogeschool Utrecht). Individuals were excluded from participation in the study if they reported (1) knowledge about the working mechanism of EMDR, (2) participating in previous EMDR or trauma film experiments, (3) using medication that affected memory or concentration, (4) being diagnosed with a psychiatric disorder, (5) having been the victim or witness of extreme physical violence, (6) being prone to fainting when seeing blood. All participants gave written informed consent. The Ethical Committee of the Faculty of Social and Behavioral Sciences at Utrecht University (FETC15-104) approved this study.

Tasks

Trauma film. The trauma film consisted of a 9m24s-minute excerpt with a coherent narrative from 'Irréversible' (2002) produced by Gaspar Noé. The selected fragment showed a brutal murder in a dark night club with explicit violence using a fire extinguisher. It has been used in earlier studies (Nixon, Cain, Nehmy & Seymour, 2009; Weidmann, Conradi, Gröger, Fehm, & Fydrich, 2009; Verwoerd et al. 2008; 2009; 2011), and induced distress in a validation study (Arnaudova & Hagenaars, 2017). The film was shown on a 23-inch computer screen in a darkened room. Sound was played via over-ear headphones.

Dual-taxation task. Each participant was instructed to select the worst mental image from the movie, which served as the memory's hotspot. Then, the participant wrote down a label for the memory's hotspot, which served to refer to that specific mental image. After image selection, a participant was assigned to one of three groups: Recall+EM, Recall+Counting, or No-Task Control. Before the intervention, participants were presented with their label and were instructed to recall the mental image. They then rated their memory on two visual analogue scales (VAS) that ranged from 0 *not vivid/unpleasant* to 100 *very vivid/unpleasant*. In each group, participants recalled the memory for 6 intervals of 24 seconds separated by 10-second breaks (total duration: 3.23 minutes). Participants were seated approximately 60 cm from a 24-inch computer screen. In the Recall+EM group participants were instructed to recall the memory and simultaneously follow a 20 pixel dot that moved horizontally across the screen by moving their eyes and keeping their head still (600 pixel amplitude on a 1280×1024 pixel screen; i.e. a viewing angle of approximately 42 degrees) with a speed of 1.2Hz (van Schie et al., 2016; van Veen et al., 2015). A 1Hz cycle corresponds with one left-right-left movement within 1 second. The experimenter directly observed whether EM were made by means of a webcam (Logitech HD Pro C920) connected to a laptop that was placed out of the participant's sight. In the Recall+Counting group, participants counted back aloud from 450 in steps of 2 (van den Hout et al., 2010). Participants in the No-Task Control group were told that they had a short break, in which they were to stay seated and remain quiet. They were allowed to think about anything, without restrictions (cf. Holmes et al., 2009; 2010; James et al., 2015; 2016b; Hagenaars et al., 2017). After each intervention, participants again recalled the memory and rated its vividness and unpleasantness using a 0-100 VAS. The dual-taxation task was presented using OpenSesame Software (Mathôt, Schreij, & Theeuwes, 2012).

Random interval repetition task. To quantify WM taxation of two different dual-tasks, participants performed a RIR task (cf. Vandierendonck et al., 1998).The RIR task was adapted from van den Hout and colleagues (2011). During the RIR task mild electrical stimuli were administered to the participant's wrist of the non-dominant hand. Participants were asked to select a stimulus that was clearly discernable, but not painful. The intensity of the electrical stimulus was determined using a work-up procedure that started at 0.1Ma, and could be increase consecutively to 0.2, 0.5, 1.0, 1.5, 2.0, 2.5, 3.0, 3.5, 4.0, 5.0, 6.0, and 7.0. The intensity was increased with 1.0Ma each time, if participants required higher levels after 7.0Ma. The duration of the electrical stimulus was 2 ms and was administered with a Digitimer High Voltage Stimulator (Model DS7A). The experiment was presented using E-Prime 2.0 software (Psychology Software Tools, Pittsburgh, PA).

Participants were instructed to press the button of a response box with the index finger of their dominant hand as soon as they felt the electrical stimulus. Half of the time the interstimulus interval (ISI) was 900ms and the other half it was 1500ms. The order of ISI was quasi-random, with no more than four consecutive ISIs of the same duration in a row. First, participants received a short practice in responding to 24 electrical stimuli. After practice, participants performed the RIR task in three conditions: RIR+EM, RIR+Counting or RIR only (i.e., no dual-task). The order of conditions was counterbalanced. During each 3-min block they received 148 electrical stimuli. The instructions for the dual-task conditions (EM or counting) in the RIR task were identical to the instructions for the dual-task conditions in the dual-taxation task. The slowing down of reaction times during dual-task conditions compared to no dual-task provides a valid and very sensitive measure of central executive taxation.

Intrusion Diary. Participants were first told what we meant with an "intrusion" (e.g., James et al., 2015). They were specifically told that intrusive memories of the film appeared spontaneously in their mind, and that intrusions could be experienced as mental images (e.g., pictures in the mind's eye, sounds, etc.), verbal thoughts in the form of words and phrases, or as an image-thought combination. Participants were instructed to register any intrusive memory they had in a pen-and-paper diary in the following week. They were instructed to record the content intrusive memories as soon as possible after they experienced it. Only intrusive memories with image-based content were scored (e.g., Holmes et al., 2009; 2010).

Self-report questionnaires.

Beck-depression inventory-II (BDI-II). The BDI-II (Beck, Steer, & Brown, 1996; for the Dutch version see van der Does, 2002) is a 21-item self-report questionnaire that assesses depressive mood. Items measure different symptoms that are related to depression (e.g. sadness, hopelessness, self-blame, etc.). All items are scored on a scale ranging from 0 to 3, and higher sum total scores (ranging from 0 to 63) indicate higher severity of depressive mood.

State trait anxiety inventory – trait (STAI-T). The STAI-T measures trait anxiety and consists of 20 items, which are scored on a four-point scale from 1 *almost never* to 4 *almost always*. (Spielberger, Gorsuch, Lushene, Vagg, & Jacobs, 1983; For the Dutch version of the STAI-T, see van der Ploeg, Defares, & Spielberger, 2000). A number of items on the trait scale are anxiety absent items and need to be reverse scored. Sum scores on the STAI-T range from 20 to 80 with higher scores representing higher levels of anxiety.

Traumatic experiences questionnaire (TEQ). The TEQ is a 12-item checklist adapted from the Criterion A list of the Posttraumatic Diagnostic Scale (Foa, 1995; for the current study all items were translated to Dutch). For a list of 12 traumatic events participants indicated whether or not they experienced or witnessed the event (e.g., a serious accident, fire or explosion; a natural disaster; life-threatening illness). Total scores ranged from 0 *no traumatic event experienced or witnessed* to 12.

Impact of event scale (IES). The IES measures psychological stress reactions after a major life event (Horowitz, Wilner & Alvarez, 1979; For the Dutch version of the IES, see van der Ploeg, Mooren, Kleber, van der Velden, & Brom, 2004). The questionnaire has 15 items that are scored on a four-point scale with 0 indicating *not at all*, 1 *rarely*, 3 *sometimes*, and 5 *often*. The IES has two subscales: Intrusions (7 items) and Avoidance (8 items). Higher sum scores for the intrusion subscale reflect higher intrusions in the last 7 days. Participants responded specifically to the items in reference to the film.

Pre- and post-film mood. Participant's rated their levels of sadness, happiness, horror, fear, anxiety, and depression at this moment. The VAS ranged from 0 *not at all* to 100 *extremely*. A composite score was used to determine overall mood (e.g., Holmes et al., 2010; James et al., 2015). Happiness was reverse scored.

Film distress, Attention and Familiarity with the Film. On a VAS ranging from 0 *not at all* to 100 *extremely* participants rated distress ("how distressing did you find the film you just watched?") and attention paid to the film ("how much attention did you pay to the film you just watched?"). They indicated whether they had seen any part of the trauma film before by selecting *yes* or *no*.

Diary Compliance. Participants rated their diary compliance on a VAS ranging from 0 *not at all accurately* to 100 *extremely accurately*.

Procedure

On Day 1, participants gave written informed consent and completed the BDI-II, STAI-T, and TEQ. Next, they filled out the pre-film mood scales and were then instructed to watch the trauma film as if there were a bystander at the scene. Participants watched the trauma film alone in a darkened room, while the experimenter waited outside until the film was finished. Then they completed the post-film mood scales and indicated how distressing they found the film, how much attention they had paid to the film, and if there they had seen the film before. Then, all participants performed a standardized 10-minute music filler task, during which 15 short fragments of classical were rated for pleasantness (see Holmes et al., 2009; 2010; James et al., 2015). After the filler task, participants performed the dual-taxation task (before and afterward this intervention, they rated memory vividness and unpleasantness), which was followed up with instructions on how to complete the paper-and-pen diary the following week.

On Day 8, participants returned to the lab and indicated their diary compliance and filled out the IES. Next, participants in the Recall+EM or Recall+Counting groups performed the RIR task. Participants in the No-Task Control group did not perform this task. Then, all participants received an extensive debriefing, in which the background of the film was explained including how the most aversive scenes were shot using latex dolls and visual effects. Part of this was that we showed participants a clip in which the special effects supervisor shows how the effects were created (www.youtube.com/watch?v=y-PnrL-uw0w). A similar debriefing has been used in other research using a trauma film paradigm (Kindt, van den Hout, Arntz, & Drost, 2008). Finally, participants were reimbursed for their participation.

Data analysis

We used Bayesian statistics to analyze the data, because these statistics allow quantification of the amount of evidence in favor of the alternative hypothesis (H_1), but also in favor of the null hypothesis (H_0; Dienes, 2016), and we sometimes expect groups *not* to differ. Moreover, Bayesian statistics are increasingly used in the field of fear and traumatic stress (e.g., Krypotos, Blanken, Arnaudova, Matzke, & Beckers, 2016; Monden et al., 2016; Schweizer et al., 2017; van de Schoot, Broere, Perryck, Zondervan-Zwijnenburg, & Van Loey, 2015; Yalch, 2016).

The data were analyzed with the software JASP (Version 0.8.2; JASP team, 2017). JASP is a free, open-source package that can be used for Bayesian statistical analyses. This package determines a Bayes Factor (BF) per requested test, which expresses the relative likelihood of the data under H_1 and the H_0. Data in favor of the H_1 (relative to H_0) are presented as BF_{10}, which can be interpreted as the BF of H_1 against H_0. BF_{01} represents the reversed interpretation, where the evidence is in favor of the H_0 (relative to H_1). $BF_{01} = 5$ therefore means that the data are five times more probable under H_0 than under H_1. Because a BF is *relative*, the BF for the other hypothesis is easily determined by dividing 1 by a given BF (e.g., if BF_{01} is 5, BF_{10} is 1 divided by 5, hence 0.2). In JASP H_0 always states the absence of an effect. In all analyses, we used JASP's standard prior: a Cauchy distribution with scale r = 0.707 (i.e., medium prior).

Because generally three groups are compared, a Bayesian one-way ANOVA is performed first. Given sufficient evidence in favor of H_1, this is followed by three Bayesian *t*-tests comparing two out of three groups each time. In some instances, a two factor Bayesian ANOVA is performed (e.g., to test Group and Time effects). JASP then produces five models: a model for (1) the null, (2) a main effect of factor A, (3) a main effect of factor B (4), a two main effects model: A+B, (5), a two main effects model plus an A×B interaction. All models are compared to the null model; therefore all BFs state the relative probability against the null model. The relative probability amongst models 2-5 can be calculated by simply dividing model BFs. In these analyses, we used JASP's default prior model probabilities of 0.2.

Though the BF is a continuous scale without arbitrary cut-offs, BFs can also qualified by categories of evidence to facilitate scientific communication (Jeffreys, 1961; Wetzels & Wagenmakers, 2012). BFs around 1 represent (inconclusive) evidence that is not in favor of H_1 or H_0. BFs between 1-3 ($1- {}^1/_3$) represent anecdotal, 3-10 (${}^1/_3 - {}^1/_{10}$) substantial, 10-30 (${}^1/_{10} - {}^1/_{30}$) strong, or 30-100 (${}^1/_{30} - {}^1/_{100}$) very strong evidence relative to the other hypothesis. A BF above 100 (or below ${}^1/_{100}$) is interpreted as decisive evidence for one hypothesis relative to the other hypothesis.

Results, Experiment 1

Baseline measures

Overall, evidence shows that groups performed similarly on the baseline measures: BDI-II, $BF_{01} = 7.67$; STAI-T, $BF_{01} = 6.43$; and TEQ, $BF_{01} = 3.64$ (see Table 1), suggesting the groups were equivalent at the onset of the experiment.

Table 1. Experiment 1: Baseline Measures Sex, Age, Depression, Anxiety, and Trauma History Ratings; and Film Related Mood, Distress, Attention, and Familiarity for each Group. Means (SD) are reported.

	No-Task Control	Recall+EM	Recall+Counting
Sex: M/F (*n*)	7 / 18	6 / 19	9 / 17
Age	20.64 (1.96)	21.45 (1.94)	21.12 (1.95)
BDI-II	4.60 (3.57)	4.64 (4.32)	5.19 (3.83)
STAI-T	32.52 (5.91)	31.04 (7.47)	32.50 (6.43)
TEQ	0.88 (1.17)	0.80 (0.91)	0.50 (0.71)
Pre-Film Mood	9.84 (5.49)	14.4 (12.43)	14.15 (6.45)
Post-Film Mood	47.85 (21.43)	53.03 (22.47)	49.01 (14.87)
Distress	79.40 (20.76)	85.60 (14.96)	73.54 (15.97)
Attention	90.92 (8.85)	90.64 (11.91)	87.38 (10.92)
Familiarity (*n*)	1	1	2
Pre-Intervention Unpleasantness	86.59 (21.59)	85.91 (18.14)	81.29 (20.46)
Post-Intervention Unpleasantness	75.60 (24.08)	79.46 (15.31)	65.09 (22.95)
Pre-Intervention Vividness	59.97 (27.12)	69.56 (27.95)	69.44 (24.16)
Post-Intervention Vividness	52.95 (31.85)	64.76 (27.38)	55.89 (24.43)

Film Related Mood, Distress, and Attention

To assess whether negative mood ratings increased as a result of the film, a mixed Bayesian ANOVA was performed with factors Time and Group. The analysis shows evidence for a main effect model of Time (pre vs. post), BF_{10} = 7.591e+32. The main effect model of Group showed evidence in favor of groups performing equally, BF_{10} = 0.113. The two main effects models and the addition of a Time×Group interaction were less supported by the data relative to the main effect model of Time, respectively BF_{10} = 1.747e+32 and BF_{10} = 2.674e+31. This shows that changes in mood are predominantly an effect of the film itself, and not of the group or any interactions (see Table 1). Hence, the film successfully increased participants negative mood regardless of assigned group, which is in accordance with previous studies employing trauma films (e.g., Holmes et al., 2009; 2010),

The analysis on distress (in response to the film) showed weak evidence in favor of the null hypothesis, BF_{01} = 2.22. Likewise, reported attention paid to the film was similar for groups, BF_{01} = 4.5. Familiarity was not analyzed, because only four participants indicated they were familiar with the film, and these participants were equally divided over the three groups.

Memory Unpleasantness and Vividness

Memory unpleasantness and vividness were analyzed separately with a mixed Bayesian ANOVA with factors Time (pre vs. post-intervention) as within-subjects factor and Group as between-subjects factor, to assess whether changes in these ratings differed between groups. The analysis shows the largest BF for a main effect model of Time for unpleasantness (Figure 1, left panel); there is decisive evidence for this model compared to a null model, BF_{10} = 132,259. Inclusion of Group as additional main effect or as interaction with Time reduced the BF compared to the main effect model of Time, BF_{10} = 88,574 and BF_{10} = 58,081, respectively. A main effect model of only Group is less probable than the null model, BF_{10} = 0.617.

Vividness ratings displayed a similar pattern (see Figure 1, right panel). Again, there was decisive evidence for a main effect model of Time, BF_{10} = 27,924. Addition of Group as main effect or interaction with Time decreases the BF compared to the main effect model of Time, BF_{10} = 12,182 and BF_{10} = 3,261 respectively. Again, a main effect model of only Group performs worse compared to a null model, BF_{10} = 0.418. Overall, the results from unpleasantness and vividness ratings suggest that ratings decrease from pre to post-intervention irrespective of intervention type.

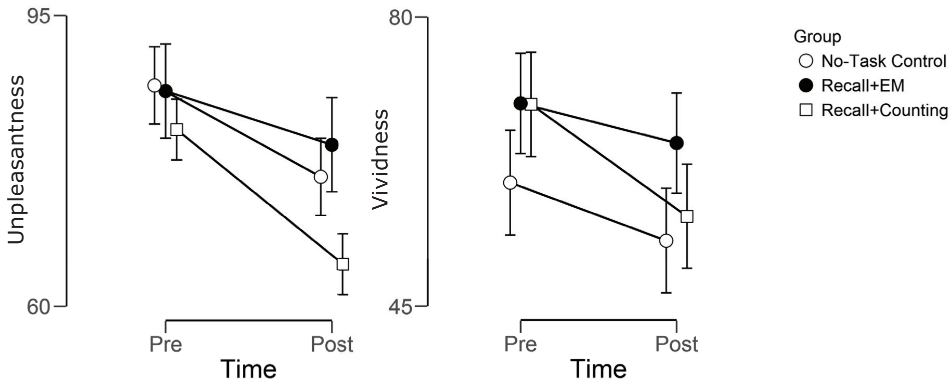


Figure 1. Experiment 1: Pre and Post-Intervention Ratings for Unpleasantness (left panel) and Vividness (right panel) for No-Task Control, Recall+EM, and Recall+Counting. Error bars represent 95% credible intervals.

Intrusive Memories

On average, participants reported 4.3 intrusions (*SD* = 4.15) in their diary (see Table 2). There were no differences between groups (BF_{01} = 7.54). Groups also reported similar re-experiences retrospectively on the IES, BF_{01} = 2.85, though the evidence here was weak. Diary compliance was high (*M* = 75.18, *SD* = 13.80) and did not differ between groups, BF_{01} = 5.87.

Table 2. Experiment 1: Intrusive Memories Reported in the Diary and on the IES for each Group. Means (SD) are reported.

	No-Task Control	Recall+EM	Recall+Counting
Diary Intrusions	4.44 (3.86)	3.84 (4.10)	4.54 (4.57)
IES	8.52 (6.46)	7.32 (5.39)	5.92 (4.20)
Diary compliance	75.95 (16.06)	76.76 (12.66)	72.96 (12.73)

WM Taxation

Taxing of WM, as evidenced by RT performance on the RIR task, showed evidence for differences between conditions, BF_{10} = 1.083e+31. Follow-up tests showed that cognitive load of RIR+EM (*M* = 394.2, *SD* = 74.07) and RIR+Counting (*M* = 426.1, *SD* = 87.68) was higher compared to RIR Only (*M* = 234.5, *SD* = 65.33), both BF_{10} > 4.781e+17.

Additionally, there was weak evidence that cognitive load was higher for RIR+Counting than for RIR+EM, $BF_{10} = 2.64$ (see Figure 2).

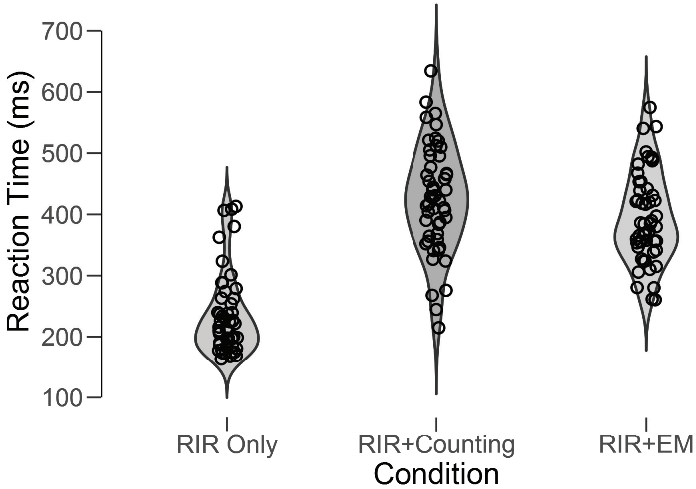


Figure 2. Kernel Density Distributions (i.e., Violin Plots) and Individual Reaction Times (circles) during RIR Only, RIR+EM, and RIR+Counting in Experiment 1.

5

Relationship between WM Taxation, Vividness/Unpleasantness and Intrusive Memories

According to WM theory, the degree of slowing down on the RIR task during Counting or during EM (relative to RIR Only) should be positively related to drops in scores for vividness and unpleasantness. To assess this, we first calculated a pre-post difference score for each participant on vividness and unpleasantness scores and then correlated this with the index of WM taxation. For each participant, this index was the difference between RIR only performance and the relevant experimental RIR condition (EM or Counting). WM taxation was not related to decreases in unpleasantness, r = -.01, $BF_{01} = 6.16$, or vividness, r = -.20, $BF_{01} = 12.72$.

Additionally, we hypothesized that decreases in unpleasantness and vividness ratings should be related to fewer intrusive memories assessed by the intrusion diary and by the IES. However, diary intrusions were not related to unpleasantness, r = .04, $BF_{01} = 9.15$, or vividness, r = .06, $BF_{01} = 10.03$; and intrusion subscale of the IES was also not related to unpleasantness, r = -.04, $BF_{01} = 5.20$, or vividness, $r < .01$, $BF_{01} = 7.18$.

Discussion Experiment 1 & Introduction Experiment 2

Dual-tasks have been proven to be successful in reducing the unpleasantness and vividness of voluntary memories. Experiment 1 was set up to test whether they would also affect *involuntary* memory (i.e., intrusive images of analogue trauma), and whether changes in vividness and unpleasantness of voluntary memories would predict intrusion frequency. Against expectations, two dual-tasks did not affect intrusion frequency. Surprisingly, the dual-tasks also did not result in drops in vividness and unpleasantness, although each task was clearly more taxing on WM than no task. The most likely explanation for the lack of effects may be the intervention's duration. While the duration was in line with previous studies employing dual-tasks (e.g., Kearns & Engelhard, 2015; Leer, Engelhard, Dibbets, & van den Hout, 2013; van Schie et al., 2016; van Veen et al., 2015), it was significantly shorter than non-dual-task interventions within the trauma film paradigm, which usually last for about 10 minutes (e.g., Holmes et al., 2009; 2010; James et al., 2016b). Also, longer durations for dual-tasks resulted in stronger effects on vividness and unpleasantness in previous studies (Leer, Engelhard, & van den Hout, 2014; van Veen, Engelhard, & van den Hout, 2016). Alternatively, it is possible that dual-tasks, indeed, do not affect the number of intrusions, but only affect qualitative characteristics of an intrusion. After all, dual-task interventions are known to affect memory characteristics that are voluntarily retrieved and assessed such as vividness and unpleasantness (e.g., Lee & Cuijpers, 2013). Therefore, in Experiment 2, we tested the effects of a longer dual-task (of approximately, 9 min), and added measures for intrusion vividness, unpleasantness, and distress to the intrusion diary.

Method, Experiment 2

Experiment 2 was similar to Experiment 1 with one crucial methodological change. We prolonged the intervention duration from 6×24s (total duration: 3.23 minutes) to 16×24s (total duration: 8.9 minutes; van Veen et al., 2016), which resembles the duration of interventions used in the trauma film paradigm. As a consequence, participants in the Recall+Counting group were instructed to count back in steps of 2 from 1000 (cf. Engelhard et al., 2011), instead of 450, to avoid that that they would reach 0. Additionally, participants were asked to rate each intrusion in the diary on distress, unpleasantness, and vividness with a number between 0 *not at all* and 10 *extremely* (e.g., Hagenaars et al., 2017). They were also asked to rate intervention compliance, pleasantness, and difficulty on a VAS ranging from 0 *not at all* to 100 *extremely* to explore whether effects might be moderated by task performance related variables.

Participants

Seventy-five participants took part in this study (25 males, 50 females; M_{age} = 20.85 years, Age range = 18 to 29 years), and were equally divided over the groups: Recall+EM, Recall+Counting, No-Task Control. One participant from the Recall+EM group was later excluded, because of overall non-compliance with instructions.

Results, Experiment 2

Baseline measures

Overall, groups performed similarly on all baseline measures: BDI-II, BF_{01} = 5.94, STAI-T, BF_{01} = 2.85, and TEQ, BF_{01} = 7.8 (see Table 3). This suggests that the groups were equivalent at the onset of the experiment.

Table 3. Experiment 2: Baseline Measures Sex, Age, Depression, Anxiety, and Trauma History Ratings; and Film Related Mood, Distress, Attention, and Familiarity for each Group. Means (SD) are reported.

	No-Task Control	Recall+EM	Recall+Counting
Sex: M/F (*n*)	6 / 19	8 / 17	10 / 14
Age	20.40 (2.06)	20.96 (2.67)	21.29 (2.37)
BDI-II	5.57 (3.61)	6.28 (4.62)	6.79 (4.68)
STAI-T	33.56 (7.33)	32.16 (7.47)	36.00 (9.00)
TEQ	0.44 (0.71)	0.48 (0.65)	0.55 (0.78)
Pre-Film Mood	12.18 (5.99)	15.52 (8.86)	13.91 (7.59)
Post-Film Mood	52.11 (17.64)	46.22 (19.77)	44.12 (12.38)
Distress	68.84 (24.66)	73.32 (20.18)	73.38 (17.35)
Attention	88.32 (8.86)	89.48 (7.00)	87.75 (10.46)
Familiarity (*n*)	0	2	0
Pre-Intervention Unpleasantness	82.96 (19.90)	78.65 (30.16)	88.88 (12.11)
Post-Intervention Unpleasantness	62.52 (24.84)	64.25 (30.60)	71.35 (25.07)
Pre-Intervention Vividness	73.16 (21.92)	74.00 (20.28)	68.64 (20.93)
Post-Intervention Vividness	62.70 (22.63)	56.98 (22.12)	47.20 (21.70)

Film Related Mood, Distress and Attention

Negative mood ratings increased as a result of the film, as evidenced by a main effect model of Time (pre vs. post), BF_{10} = 1.436e+32 (see Table 3). All other models received lower

BFs compared to the main effect model of Time: BF_{10} = 0.090 (Group), BF_{10} = 2.074e+31 (Time + Group), and BF_{10} = 2.077e+31 (Time+Group + Time×Group interaction). This shows that all groups experienced a similar increase in negative mood as a result of watching the film.

This strong film effect was also apparent on overall high distress scores in response to the film (M = 71.82, SD = 20.81), that were comparable for the groups, BF_{01} = 6.46. Reported attention paid to the film was high (M = 88.53, SD = 8.77) and showed group similarity, BF_{01} = 7.18. Only two participants indicated that they were familiar with the film; therefore, familiarity was not analyzed.

Memory Unpleasantness and Vividness

Memory unpleasantness and vividness were analyzed separately with a mixed Bayesian ANOVA with factors Time (pre vs. post-intervention) as within-subjects factor and Group as between-subjects factor, to assess whether changes in these ratings differed between groups (see Figure 3 and the Appendix). The analysis shows conclusive evidence for a main effect model of Time for unpleasantness, BF_{10} = 1.585e+9, showing that ratings decreased from pre to post-intervention regardless of Group. All other models received lower BFs compared to the main effect model of Time: BF_{10} = 0.338 (Group), BF_{10} = 7.028e+8 (Time+Group), and BF_{10} = 1.556e+8 (Time+Group + Time×Group interaction). Overall, this shows that memory unpleasantness ratings decreased equally for all groups.

Vividness ratings display a similar pattern compared to unpleasantness. Again, there was decisive evidence for a main effect model of Time, BF_{10} = 375,031, showing that all groups experienced similar decreases in vividness ratings from pre to post-intervention., and all other models received lower BFs compared to the main effect model of Time: BF_{10} = 0.47 (Group), BF_{10} = 220,065 (Time+Group), and BF_{10} = 78,804 (Time+Group + Time×Group interaction). Overall, the hypothesized interaction was not observable in the ratings of unpleasantness or vividness.

Task Compliance, Pleasantness, and Difficulty

Participants reported high task compliance overall (M = 83.41, SD = 12.77), but, compliance differed between the groups, BF_{10} = 4.57 (see Table 4). Follow-up tests showed that No-Task Control had higher compliance ratings than Recall+Counting, BF_{10} = 15.84. There is insufficient evidence for meaningful conclusions about the other follow-up test for No-Task Control vs. Recall+EM ,BF_{10} = 1.04, and for Recall+EM vs. Recall+Counting, BF_{10} = 0.61.

Additionally, on task pleasantness, the evidence is inconclusive about relative support for the null or alternative hypothesis, $BF_{10} = 0.715$. Groups did differ on task difficulty, $BF_{10} = 3497$. Recall+EM and Recall+Counting are considered to be more difficult than No-Task Control, both $BF_{10} > 97$. Recall+EM and Recall+Counting did not seem to differ from each other in reported difficulty, $BF_{01} = 2.1$, but the evidence is weak.

Table 4. Experiment 2: Task Compliance, Pleasantness, & Difficulty for each Group. Means (SD) are reported.

	No-Task Control	Recall+EM	Recall+Counting
Task Compliance	88.80 (8.46)	83.28 (12.87)	77.96 (14.41)
Task Pleasantness	37.72 (19.19)	27.32 (18.54)	29.58 (14.86)
Task Difficulty	22.04 (19.20)	46.60 (24.48)	54.00 (21.49)

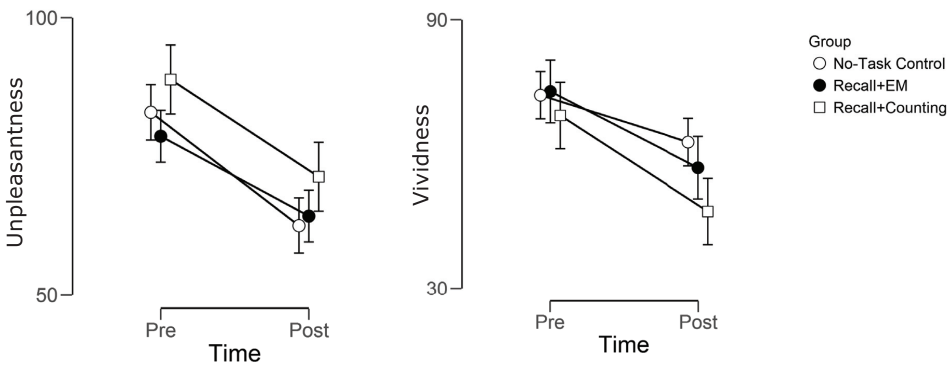


Figure 3. Experiment 2: Pre and Post-Intervention Ratings for Unpleasantness (left panel) and Vividness (right panel) for No-Task Control, Recall+EM, and Recall+Counting. Error bars represent 95% credible intervals.

Intrusive Memories

There were group differences in the number of diary memories, $BF_{10} = 7.25$ (see Table 5 and Figure 4). Follow-up tests show that both Recall+EM and Recall+Counting reported fewer intrusions compared to No-Task Control, $BF_{10} = 3.695$ and $BF_{10} = 9.384$, respectively. There was no difference between Recall+EM and Recall+Counting in the number of intrusions, $BF_{01} = 3.46$. Moreover, participants reported high compliance in reporting intrusions in their diary ($M = 74.38$, $SD = 12.05$), but the evidence about the absence or presence of any group differences in compliance was inconclusive, $BF_{01} = 1.29$.

Finally, we did not observe sufficient evidence for group differences on retrospectively reported intrusions on the IES, $BF_{10} = 1.66$.

Participants with at least one intrusion in the diary were included in the analyses of intrusion characteristics. For each participant the average score for an intrusion characteristic was used. We observed no group differences on distress ($BF_{01} = 5.93$) or unpleasantness ($BF_{01} = 4.62$). The evidence for vividness was inconclusive ($BF_{01} = 1.6$) (see averages in Table 5).

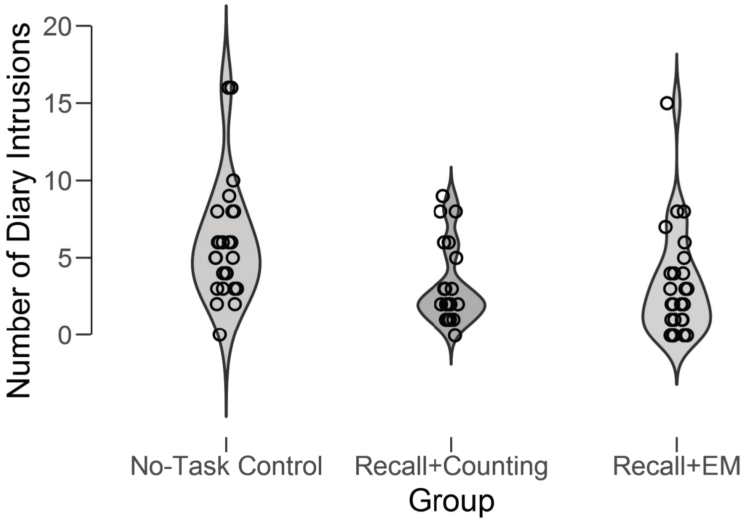


Figure 4. Kernel Density Distributions (i.e., Violin Plots) and Individual Reported Number of Intrusions (circles) for No-Task Control, Recall+EM, and Recall+Counting

Table 5. Experiment 2: Intrusion Characteristics: Distress, Unpleasantness, and Vividness Ratings for each Group. Means (SD) are reported.

	No-Task Control	Recall+EM	Recall+Counting
Diary Intrusions	5.92 (3.85)	3.28 (3.47)	3.13 (2.51)
IES	9.47 (5.11)	7.28 (5.05)	5.91 (3.94)
Intrusion Distress[1]	2.55 (1.60)	2.91 (2.10)	3.03 (2.11)
Intrusion Unpleasantness[1]	3.89 (2.11)	4.74 (2.80)	4.45 (2.27)
Intrusion Vividness[1]	5.58 (1.58)	5.27 (1.86)	4.45 (2.23)

[1] Seven participants were excluded from these analyses, because they reported no intrusions. This changed the number of participants in each group slightly: No-Task Control (*n* = 24), Recall+EM (*n* = 20), and Recall+Counting (*n* = 23).

5

WM Taxation

Taxing of WM, as evidenced by RT performance on the RIR task, showed decisive evidence for differences between groups, $BF_{10} > 10{,}000$. Follow-up showed that RIR+EM (*M* = 390, *SD* = 86.32) and RIR+Counting (*M* = 438.7, *SD* = 89.23) were more taxing compared to RIR Only (*M* = 267.7, *SD* = 88.31), $BFs_{10} > 1.275e+7$. Additionally, RIR+Counting was more taxing than RIR+EM, $BF_{10} = 28.7$ (see Figure 5).

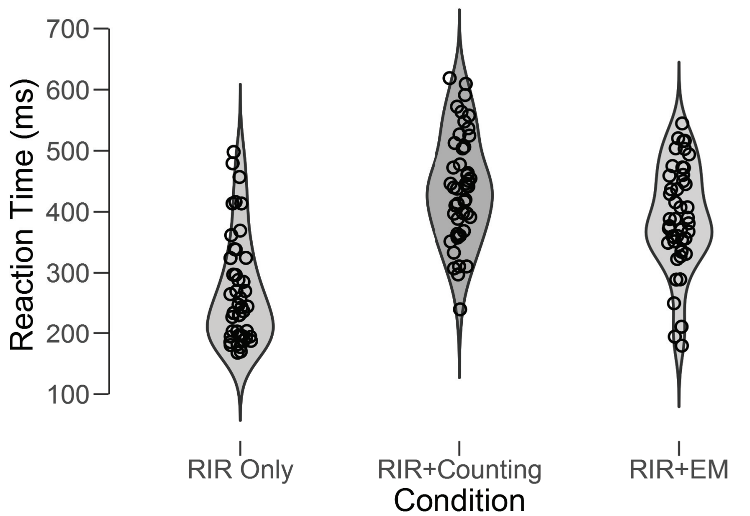


Figure 5. Kernel Density Distributions (i.e., Violin Plots) and Individual Reaction Times (circles) during RIR Only, RIR+EM, and RIR+Counting in Experiment 2

Relationship between WM Taxation, Vividness/Unpleasantness and Intrusive Memories

WM taxation should be positively related to drops in scores for vividness and unpleasantness. However, again, we did not find correlations between WM taxation and unpleasantness, $r = .08$, $BF_{01} = 3.49$, or vividness, $r = -.03$, $BF_{01} = 6.27$. Additionally, changes in unpleasantness and vividness ratings were inconsistently related to intrusive memories in general. Some correlations showed evidence in favor of the null hypothesis, such as unpleasantness and diary intrusions, $r = -.06$, $BF_{01} = 4.54$, and vividness and intrusions on the IES, $r = -.13$ $BF_{01} = 2.18$. The other correlations showed evidence that did not support the alternative or the null hypothesis: vividness and diary intrusions, $r = -.18$, $BF_{01} = 1.25$, and unpleasantness and intrusions on the IES $r = -.15$, $BF_{01} = 1.8$.

We also calculated correlations between drops in scores for vividness and unpleasantness, and,self-reported intrusion characteristics of vividness, unpleasantness, and distress. Larger drops should be related to lower scores on intrusion characteristics. For both vividness and unpleasantness, we observed evidence that was largely in favor of the null hypothesis, *r*s between -.12 and .08, BFs_{01} between 2.39 and 10.39.

General discussion

The current studies tested whether dual-task interventions reduce involuntary intrusive images and whether this depends on the dual-task's modality. Additionally, we examined if reductions in voluntarily recalled memory vividness and unpleasantness predict fewer intrusions (Experiment 1 and 2) and lower scores on intrusion qualities (Experiment 2). To test this, we induced intrusions by a trauma film and used dual-task interventions of different modalities (visuospatial and verbal). In experiment 1, we found that the two dual-task interventions did not differ from the control condition on voluntary or involuntary measures, despite the fact that WM was clearly taxed more in the dual-task conditions compared to the control condition. In experiment 2, we prolonged the dual-task intervention and observed an effect on involuntary memory: each dual-task resulted in a lower number of intrusions compared to the no-task control condition. However, performing a dual-task still did not affect voluntary memory (i.e., vividness and unpleasantness ratings), compared to control, although it again taxed WM more than control.

Our results suggest that dual-tasks after a trauma film reduce involuntary intrusive images, but only when a prolonged intervention is used. The dual-task intervention in experiment 2 was prolonged to match the duration of typical 10-minute single-task interventions within the trauma film paradigm, such as playing Tetris or finger tapping (e.g., Deeprose et al., 2012; Holmes et al., 2009; 2010; James et al., 2016b). This suggests that interventions following the trauma film need to be of a sufficient duration to modulate intrusions. Interestingly, in the current study, intrusion modulation occurred regardless of modality; the visuospatial and the verbal dual-task intervention both effectively reduced the number of intrusions compared to a no-task control condition.

The effectiveness of a visuospatial dual-task is in line with previous research that shows that visuospatial single-task interventions (i.e., interventions without simultaneous memory recall) reduce intrusions (for overviews see Brewin, 2014; James et al., 2016a). However, earlier studies show that effects for verbal single-tasks on intrusion modulation are rather mixed. The comparable effects for a verbal and visuospatial dual-task in our study gives the impression that intrusion modulation is dependent on a general taxation effect of the intervention (also see Hagenaars et al., 2017; Krans et al., 2009; Pearson & Sawyer, 2011). However, there may be involvement of modality specific processing as well. There were similar effects on intrusion modulation, yet the visuospatial task was less taxing on WM compared to the verbal task. This suggests that there was something other than general WM load about this condition that increased its efficacy. This may have been the fact that the dual-task EM intervention was visuospatial in nature (and therefore taxed WM's VSSP specifically). Previous research on hotspot modulation showed that modality specific effects and general dual-task load can co-occur, and that modulation is most efficient when memory and intervention are modality congruent (Kemps & Tiggeman, 2007; Matthijssen, van Schie, & van den Hout, 2017; but see Tadmor et al., 2017).

Unexpectedly, WM taxation with a visuospatial or verbal dual-task in both experiments, did not result in stronger decreases in vividness and unpleasantness of the memory hotspot compared to no-task control. This is at odds with previous studies showing that visuospatial dual-tasks specifically (e.g., Andrade et al., 1997, 2012; May et al., 2010), or any dual-task that taxes WM (e.g., Engelhard et al., 2011; Gunter & Bodner, 2008; van den Hout et al., 2010) decrease memory ratings compared to recall only control conditions. It is the more surprising, because in both experiments, making EM and counting both slowed down reaction times compared to no-task control, as evidenced by RIR task performance. The reaction times for the dual-task conditions are comparable to previous studies testing similar samples (Engelhard et al., 2011; van den Hout et al., 2010; 2011), which suggests that overall these dual-tasks were effective in taxing WM.

One explanation for not finding differential effects for vividness or unpleasantness between the dual-task and no-task control conditions may have been inherent to the control condition itself. The no-task control condition mirrored previous work with the trauma film paradigm (e.g., Holmes & Bourne, 2008; James et al., 2016a), but this particular condition has never been used within a dual-task framework. In dual-task paradigms the control condition traditionally is a 'recall only' condition, in which the hotspot is recalled continuously, but *without* performing a task such as making EM. The typical effect of a recall only condition on vividness or unpleasantness ratings is no change (e.g., Leer et al., 2013; Littel, Remijn, Tinga, Engelhard, & van den Hout, 2017; van Veen et al., 2016), or an increase from pre to post intervention (e.g., Engelhard et al., 2010a; 2010b; Leer et al., 2014; Maxfield et al., 2008; van den Hout et al., 2010; 2011; van den Hout, Muris, Salemink, & Kindt, 2001). However, in the no-task control condition in the current study we observed a decrease. This suggests two key things. First, participants probably did not engage in memory recall that was analogous to a recall only intervention (though this cannot be ruled out). Second, a no-task control condition reflects effects that go above and beyond processes that affect all conditions uniformly (e.g., natural decay), and any meaningful interactions (of Time and Group) are − at least − not partly explained by increases in memory ratings as are expected as a result of a Recall Only condition. Therefore, including a passive no-task control condition may be recommended for future research.

A different explanation for the vividness and unpleasantness data pattern could be related to characteristics of our stimulus material. The trauma film we used induced a memory that was novel and non-idiosyncratic in nature, and also highly aversive. This is different from the majority of experiments that used dual-task interventions, which were applied to self-selected autobiographical memories (e.g., Engelhard et al., 2010b; van den Hout et al., 2014; van Schie et al., 2016; van Veen et al., 2016). It is, however, unlikely that merely the memory's novelty or non-self-referential nature would obviate finding differential effects. After all, based on the WM model there is no a priori reason to assume that dual-task interventions would only be effective for autobiographical materials (Baddeley, 2012). In fact, several studies induced an aversive memory using novel and non-idiosyncratic materials (e.g., negative pictures, memories of an aversive film clip, or memories acquired in virtual reality environments) and showed that dual-task interventions were effective in reducing memory vividness and/or unpleasantness (Andrade et al, 1997; Cuperus et al., 2015; Leer et al., 2013; 2017; van den Hout et al., 2013).

The highly aversive nature of the movie, and subsequently the memory hotspot, may be a reason why vividness and unpleasantness ratings did not decrease more than no-task control. Perhaps the aversive nature of the hotspot caused a process of continuous hypermnesic recall. Memory recall then occupies many WM resources, while none to very little remain for the dual-task. This prevents competition effects and thereby hampers changes in vividness and unpleasantness. It is also possible, however, that the aversive nature of the hotspot circumvents competition in WM in an entirely different way. There may have been lack of competition, but not because of too much memory recall, but of too little. Participants may have been motivated to avoid thinking about the hotspot. Effectively, this may have changed a dual-task intervention into a single-task intervention, which then modulated involuntary re-experiencing, but left voluntary memory intact (e.g., Bourne et al., 2010; Holmes et al., 2009; 2010). Compared to other studies (using novel or autobiographical materials), the hotspot in the current experiment was rated higher on vividness, but especially on unpleasantness (e.g., Engelhard et al., 2010a; Leer et al., 2013, 2017), and compared to other affective films, the trauma film employed in this study was rated as highly aversive and distressing (Arnaudova & Hagenaars, 2017). Currently, however, either explanation is speculative and can only be substantiated in future experiments if participants are asked to (retrospectively) report to what extent they recalled the memory when they performed a dual-task.

Although prolonged dual-task interventions − regardless of modality − modulated intrusions, given the current design, an alternative explanation for the results cannot be ruled out. It is possible that continuous memory recall was the driving force behind intrusion modulation; both active conditions differed from the no-task control condition not only in the presence of a dual-task, but also in memory recall. This limitation clearly necessitates an additional experiment that incorporates a Recall Only control condition. At the same time, memory recall performance in all conditions should be assessed. This could be achieved by having participants report *after* the intervention the extent to which they recalled the memory *during* the intervention. If this third experiment produces results that are similar to the second experiment (with recall only performing similar to no-task control), only then can we conclude that effects of the dual-task interventions are not merely the result of memory recall.

For now, we found that only prolonged dual-task interventions, regardless of modality, decreased intrusions, but did not result in larger decreases in vividness and unpleasantness compared to no-task control. The discrepancy between voluntary and involuntary memory casts doubt on the idea that one sequential process affects both measures. Moreover, it has been argued that involuntary and voluntary memories are experienced differently (e.g., Berntsen, 1998; Schlagman & Kvavilashvili, 2008), and it is possible that different mechanisms affect them. However, sound conclusions about supposed mechanisms and their effects cannot be drawn until a third experiment is performed containing a recall only condition. Until that point, the jury is still out on whether dual-task interventions, indeed, modulate intrusive memories.

Acknowledgments

This research was supported by a TOP grant (dossier number: 40-00812-98-12030) from the Netherlands Organization for Health Research and Development (ZonMw). IME is supported with a Vici grant (453-15-005) from the Netherlands Organization for Scientific Research (NWO). We are grateful to Emily Holmes and her group for training us in using the trauma film paradigm. We thank Johan Verwoerd and Ineke Wessel for providing us with the trauma film, Arne Leer for his help with the Random Interval Repetition task, and Luke de Kruif, Bart Strijbos, Lisa van Impelen, and Lindi Wolbert for their assistance in testing.

Author contributions

KvS and IME developed the study concept. All authors contributed to the study design. KvS was responsible for data collection and analysis. KvS drafted the manuscript and SCvV, MAH, and IME provided critical revisions.

CHAPTER 6

REDUCING INTRUSIVE MEMORIES: THE ROLE OF DELIBERATE SIMULTANEOUS RECALL DURING TAXATION OF WORKING MEMORY

K. VAN SCHIE
H. KESSLER
M. A. VAN DEN HOUT
I. M. ENGELHARD

Submitted

Abstract

Background and objectives: Intrusive memories are a core symptom of posttraumatic stress disorder. Reducing the adversity and/or frequency of the former may be accomplished by two interventions that appear similar, but are mechanistically different: selectively taxing visuospatial working memory resources with or without deliberate simultaneous memory recall. We tested whether the addition of simultaneous memory recall while taxing visuospatial working memory (by playing Tetris) augments reducing involuntary intrusive memories.

Methods: Participants viewed a trauma film followed by one of three interventions (1) Tetris+ Recall, (2) Tetris Only, or (3) No-Task Control. Before and after each intervention, participants rated vividness and unpleasantness of the memory hotspot. In the following week, they recorded intrusive memories about the film in an intrusion diary.

Results: Tetris+ Recall resulted in the largest decrease in memory vividness. Memory unpleasantness showed decreases from pre to post-intervention for all conditions. There were no differences between conditions on the number of intrusions.

Limitations: The current study was an analogue study on intrusive re-experiencing and results may not necessarily generalizes to interventions directly following real life trauma.

Conclusions: Lack of effects on intrusive re-experiencing necessitates critical reflection on what specific elements of Tetris are supposed to make Tetris gameplay (with or without simultaneous recall) an intrusion modulating intervention. We suggest that future research should focus on two complementary approaches: (1) assessing the robustness of the 'Tetris effect' by setting up multilab pre-registered independent replications, and (2) carefully testing the working (neuro)cognitive mechanism(s) underlying intrusion modulation.

Keywords: intrusive memory, involuntary memory, dual-task, trauma film, working memory, PTSD

Highlights

- Visuospatial Working Memory can be taxed by playing the game Tetris
- Playing Tetris with and without simultaneous recall supposedly reduces intrusions
- Playing Tetris with and without memory recall were directly compared
- Playing Tetris with, but not without, memory recall reduced memory vividness
- These two interventions have the same number of intrusions compared to control

Introduction

The majority of people experience or witness a potentially traumatic event at one point in their lives (Breslau et al., 1998; Breslau, 2009). A significant minority of them subsequently develop post-traumatic stress disorder (PTSD; Santiago et al., 2013), which is characterized by persistent re-experiencing of the trauma, avoidance of trauma-related reminders, negative thoughts, and increased arousal or reactivity (American Psychiatric Association, 2013). Re-experiencing of the trauma can take the form of flashbacks of the traumatic event that are typically visual in nature (Ehlers et al., 2002; Hackmann, Ehlers, Speckens, & Clark, 2004). These intrusions are debilitating and have a strong "here-and-now" quality, which heightens the sense of threat (Kleim, Graham, Bryant, & Ehlers, 2013). To limit the occurrence and adversity of intrusions, recent studies have employed different behavioral interventions (e.g., Holmes, James, Coode-Bate, & Deeprose, 2009; Horsch et al., 2017; Iyadurai et al., 2017; van Schie, van Veen, Hagenaars, & Engelhard, 2017).

One approach to targeting intrusive memories is selectively taxing visuospatial working memory (WM) resources quickly after exposure to analogue trauma (i.e., a trauma film). At this point, the memory has not yet consolidated and is still in a labile state and susceptible to memory modification (e.g., McGaugh, 1966). Taxing visuospatial WM resources is typically done by performing a visuospatial task, such as playing the computer game Tetris (e.g., Hagenaars, Holmes, Klaassen, & Elzinga, 2017; Holmes et al., 2009; Holmes, James, Kilford, & Deeprose, 2010; Horsch et al., 2017; Iyadurai et al., 2017; James et al., 2015). Playing Tetris supposedly engages visuospatial WM resources that otherwise would have been used for excessive perceptual processing of the traumatic event (e.g., Brewin, 2014; Ehlers & Clark, 2002). Because taxing visuospatial WM resources selectively disrupts sensory aspects of memory that give rise to intrusive images, it limits future intrusion development. Indeed, a number of studies have shown that playing Tetris after watching a trauma film reduces the number of intrusions compared to a No-Task Control condition (e.g., Deeprose, Zhang, DeJong, Dalgleish, & Holmes, 2012; Hagenaars et al., 2017; Holmes et al., 2009; 2010; Horsch et al., 2017; Iyadurai et al., 2017; James et al., 2015). Note that in these 'trauma film' studies, Tetris is played without doing anything else (i.e., a single-task intervention).

A different intervention to target intrusions is deliberately recalling the 'hotspot' of an aversive memory while *simultaneously* performing a secondary task (i.e., a dual-task intervention). Effective secondary tasks include making horizontal or vertical eye movements by following a dot that moves on a computer screen (Gunter & Bodner, 2008; for an overview, see van den Hout & Engelhard, 2012), playing Tetris (Engelhard, van Uijen, & van den Hout, 2010b), and drawing a complex figure (Gunter & Bodner, 2008). It is hypothesized that simultaneous memory recall and secondary task performance results in competition for limited WM resources, which impedes memory retrieval and subsequently reduces memory vividness and/or unpleasantness. Van den Hout and Engelhard (2012) hypothesized that after this process, the temporarily labile memory *re*consolidates into long-term memory, and that during future recalls the changed memory will be retrieved. These findings offer an explanation for the effects of making eye movements in Eye Movements Desensitization and Reprocessing (EMDR; see Lee & Cuijpers, 2013). Moreover, because the majority of intrusions in PTSD are related to the hotspot (Holmes, Grey, & Young, 2005; Grey & Holmes, 2008), reductions in memory vividness/unpleasantness may also reduce the number of intrusions (van Schie et al., 2017).

Single and dual-task interventions have largely been applied during different stages of memory consolidation and reconsolidation processes. Single-task interventions have frequently been applied during consolidation of novel memories (e.g., Hagenaars et al., 2017; Holmes et al., 2009; 2010), and dual-task interventions have been applied during reconsolidation of already existing memories (e.g., Engelhard et al, 2010b; van Schie et al., 2016; van Veen et al., 2015). Recently, James and colleagues (2015) showed that single-task interventions can also effectively modulate intrusions during reconsolidation, and some studies on dual-task interventions showed that dual-task interventions are effective in reducing memory vividness/unpleasantness during consolidation (e.g., Cuperus, Laken, van den Hout, & Engelhard, 2016; Leer, Engelhard, Altink, & van den Hout, 2013; Leer et al., 2017; van den Hout, Bartelski, & Engelhard, 2013). Thus, both single-task as dual-task interventions possibly can be applied during memory consolidation and reconsolidation to modulate aversive memories.

A major difference between single-task and dual-task interventions is that during dual-task interventions the memory hotspot is deliberately recalled *while* simultaneously performing the intervention. That is, in single-task interventions, the novel memory is primed briefly before the intervention by showing participants static neutral images of the film (e.g., Holmes et al., 2009, 2010; James et al., 2015), and participants are not instructed to recall the memory while performing the single-task intervention. In dual-task interventions, participants are instructed to simultaneously recall the memory during the intervention, which initiates the memory´s labile state (van den Hout & Engelhard, 2012). Therefore, a labile memory state is a necessary component in both approaches. Interestingly, each approach is effective in reducing intrusive memories (e.g., Holmes et al., 2009; 2009; van Schie et al., 2017), but the two approaches have never been directly compared.

The aim of this study was to directly compare a single-task and a dual-task intervention and to examine their effects on vividness and unpleasantness of the hotspot and the number of intrusive memories. To induce intrusions we employed the trauma film paradigm (see Holmes & Bourne, 2008). Participants first watched a film with aversive content and were then assigned to one of three conditions: No-Task Control, Tetris+Recall, or Tetris Only. We used Tetris as intervention, because it has been used in both single-task (e.g., Holmes et al., 2009) and dual-task interventions (Engelhard et al., 2010b). Before and after each intervention, participants rated memory vividness and unpleasantness. In the week following the film, participants recorded intrusions about the film in their diary.

Materials & Methods

Participants

Ninety students participated in this study (22 males, 68 females; M_{age} = 20.93 years, Age range = 17 to 28 years). Individuals were excluded from participation if they reported (1) knowledge about the working mechanism of EMDR, (2) participating in previous EMDR or trauma film experiments, (3) using medication that affected memory or concentration, (4) being diagnosed with a psychiatric disorder, (5) having been the victim or witness of extreme physical violence, (6) being prone to fainting when seeing blood.

Tasks

Trauma film. The trauma film consisted of a 9m24s-minute excerpt from 'Irréversible' (2002). The selected fragment shows a brutal murder in a dark night club with explicit violence using a fire extinguisher (Nixon, Cain, Nehmy, & Seymour, 2009; Weidmann, Conradi, Gröger, Fehm, & Fydrich, 2009; van Schie et al., 2017; Verwoerd et al., 2008; 2009; 2011). The film was shown on a 23-inch computer screen in a darkened room. Sound was played via over-ear headphones.

Tetris. In the tile-matching puzzle game Tetris (Version 1.6.00, Electronic Arts Inc.), participants see seven differently colored geometric blocks (i.e., tetrominoes) that fall from the top of the screen to the bottom. Using the touch screen, participants are instructed to move the blocks left or right, rotate them 90 degrees, or move them down to complete as many horizontal lines across the bottom of the playing area as possible. When a horizontal line is completed, it disappears from the screen and points are awarded. All participants start at level

1. Over the course of the game, blocks descend faster as more rows are completed (Marathon mode). In line with previous experiments (e.g., Holmes et al., 2009; 2010; James et al., 2015), participants were instructed to focus on the three tetrominoes that were to appear at the top of the playing field after the one that was currently in play. These blocks were clearly visible at the upper right-hand side of the screen. Also, participants were encouraged to mentally rotate these blocks and to figure out where to best place them. Visual mental rotation was further encouraged by playing Tetris without a ghost tetromino. Participants played Tetris on a 9.7-inch iPad without sound. Each participant´s cumulative Tetris score was noted.

Intrusion intervention task. Each participant selected the worst mental image from the movie (i.e., the memory's hotspot). Then, the participant wrote down a label for the memory's hotspot, which served to refer to that specific mental image. After image selection, each participant was assigned to one of three conditions: Tetris Only (i.e., a single task intervention), Tetris+Recall (i.e., a dual-task intervention), or No-Task Control. Assignment to conditions was random with the restriction that after three participants all conditions were tested equally often. Before the intervention, participants were instructed to recall the hotspot of the memory and to rate the hotspot on two visual analogue scales (VAS) that ranged from 0 *not vivid/unpleasant* to 100 *very vivid/unpleasant*. In the Tetris Only condition, participants played Tetris for 10 minutes consecutively (Holmes et al., 2009; 2010; James et al., 2015; James et al., 2016). In the Tetris+Recall condition, participants recalled the memory and simultaneously played Tetris for 25 intervals of 24 seconds. The intervals were separated by 10-second breaks which ended with a brief reiteration of the instructions. The duration of the Tetris game play was the same in these two active conditions. Participants in the No-Task Control condition were told that they had a short 10-minute break, in which they were instructed to stay seated and remain quiet. They were told that they could think about anything, without restrictions (e.g., Holmes et al., 2009; 2010). After each intervention, participants again recalled the hotspot and rated its vividness and unpleasantness using the two VASs. The task was presented using E-Prime 2.0 software (Psychology Software Tools, Pittsburgh, PA).

Intrusion Diary. Participants were first explained what we meant with an "intrusion" (e.g., James et al., 2015). We told them that intrusions appeared spontaneously in their mind, and that these intrusions could be mental images (e.g., pictures in the mind's eye), or verbal thoughts in the form of words and phrases. Participants were instructed to record these intrusions in a pen-and-paper diary for a week. They were asked to write down the intrusion's content as soon as possible after they experienced it. Moreover, we asked them to score that intrusion on distress, vividness, unpleasantness, and suddenness on scales that ranged from 0 *not at all* to 10 *extremely*. Only intrusive memories with image-based content were scored (e.g., Holmes et al., 2009; 2010).

Intrusion-provocation task (IPT). The IPT was adapted from James and colleagues (2015) and was included as an exploratory measure. To construct the IPT, we selected nine static images of the film. Each image was blurred using GIMP 2.8 (Gaussian Blur set 20.0). During the IPT, participants were seated in front of the computer and then viewed the blurred images, which were presented for 2s each according to the film's chronology. After viewing the blurred images, participants were instructed to close their eyes for 2 minutes and to press a selected key on a computer board whenever they experienced an intrusive memory of the film. Intrusive memories were defined as in the intrusion diary. The IPT was administered one week after viewing the film clip (see below).

Self-report measures

Beck-depression inventory-II (BDI-II). The BDI-II (Beck, Steer, & Brown, 1996; for the Dutch version see van der Does, 2002) is a 21-item self-report questionnaire that assesses depressive mood. Items measure different symptoms that are related to depression. All items are scored on a scale ranging from 0 to 3, and higher sum total scores (ranging from 0 to 63) indicate higher severity of depressive mood.

State trait anxiety inventory – trait (STAI-T). The STAI-T measures trait anxiety and consists of 20 items, which are scored on a four-point scale from 1 *almost never* to 4 *almost always*. (Spielberger, Gorsuch, Lushene, Vagg, & Jacobs, 1983; For the Dutch version of the STAI-T, see van der Ploeg, Defares, & Spielberger, 2000). A number of items need to be reversed scored. Sum scores on the STAI-T range from 20 to 80 with higher scores representing higher levels of anxiety.

Traumatic experiences questionnaire (TEQ). The TEQ is a 12-item checklist adapted from the Criterion A list of the Posttraumatic Diagnostic Scale (Foa, 1995; for the current study all items were translated to Dutch). For a list of 12 traumatic events participants indicated whether or not they experienced or witnessed the event (e.g., a serious accident, life-threatening illness). Total scores ranged from 0 *no traumatic event experienced or witnessed* to 12.

Impact of event scale (IES). The IES measures psychological stress reactions after a major life event (Horowitz, Wilner & Alvarez, 1979; For the Dutch version of the IES, see van der Ploeg, Mooren, Kleber, van der Velden, & Brom, 2004). The questionnaire has 15 items that are scored on a four-point scale with 0 indicating *not at all*, 1 *rarely*, 3 *sometimes*, and 5 *often*. The IES has two subscales: Intrusions (7 items) and Avoidance (8 items). Higher sum scores reflect higher intrusions and avoidance in the last 7 days. Participants were instructed to rate each item with respect to the film.

PTSD checklist for DSM-5 (PCL-5), cluster B. The PCL-5 is 20 item self-report measure that directly maps onto DSM-5 symptom criteria for PTSD (Weathers et al., 2013; for the Dutch version, see Boeschoten, Bakker, Jongedijk, & Olff, 2014). Only the five items that specifically map onto the symptom cluster of re-experiencing were used. These items were adjusted to refer specifically to re-experiencing of the film (as traumatic event) in the last week. Items are scored from 0 *not at all* to 4 *extremely*, and a total score can be calculated by summing up all items.

Pre- and post-film mood. Before and after the film, participants rated their current levels of sadness, happiness, horror, fear, anxiety, and depression. Each VAS ranged from 0 *not at all* to 100 *extremely*. A composite score was used to determine overall mood; happiness was reversed scored (e.g. Holmes et al., 2010, James et al., 2015).

Film distress, Attention and Familiarity with the Film. On a VAS ranging from 0 *not at all* to 100 *extremely,* participants rated distress and attention paid to the film. A participant indicated whether they were familiar with any part of the trauma film by selecting *yes* or *no*.

Task compliance, Task difficulty, Task pleasantness, Memory Recall. On a VAS ranging from 0 *not at all* to 100 *extremely,* participants rated their general task compliance, task difficulty, and task pleasantness. They also rated the extent to which they recalled the memory was simultaneously recalled during the intervention on a VAS that ranged from 0 *not at all* to 100 *always*.

Diary Compliance. Participants rated their diary compliance on VAS ranging from 0 *not at all accurately* to 100 *extremely accurately*.

Procedure

The experiment consisted of two laboratory sessions, which were separated by one week, in which participants were individually tested. Session 1 lasted approximately 90 minutes, and session 2 approximately 30 minutes. In between these sessions participants completed the paper-and-pen diary at home and recorded the number of intrusive memories.

On Day 1, all participants gave written informed consent, and completed the BDI-II, STAI-T, and TEQ. Next, the experimenter explained how Tetris needed to be played and participants practiced playing Tetris for 3 minutes. Then they filled out the pre-film mood scales and were subsequently instructed to watch the trauma film as if there were a bystander at the scene. They watched the trauma film alone in a darkened room. The experimenter waited outside until the film was finished. After the film, participants filled out the post-film mood scales and indicated levels of film distress, attention paid to the film and familiarity with the film. Then, participants performed a standardized 10-minute music filler task that consisted of 15 short fragments of classical that participants rated on pleasantness (e.g. Holmes et al., 2009; 2010; James et al., 2015). After the filler task participants performed the intrusion intervention task and afterwards rated their task compliance, task difficulty, and task pleasantness (Emily A. Holmes, personal communication, 29 June 2016). They also rated to what extent they recalled the memory hotspot during the intervention. They received diary instructions directly after these ratings.

On Day 8, participants visited the lab and rated their diary compliance, filled out the IES and PCL-5, completed the IPT, and handed in the diary. Participants received an extensive debriefing, in which a clip was shown in which the special effects supervisor explained extensively how the effects in the most aversive scenes were shot using latex dolls and visual effects (www.youtube.com/watch?v=y-PnrL-uw0w). A similar debriefing has been used in other research using a trauma film paradigm (Kindt, van den Hout, Arntz, & Drost, 2008). Finally, participants were reimbursed for participating.

Data analysis

For data analysis we used JASP; a free, open-source software package for Bayesian statistical analyses (JASP team, 2016). JASP determines a Bayes Factor (BF) per requested test. A BF expresses the relative likelihood of the data under H_1 and the H_0. BF_{10} represents data in favor of H_1. This can be interpreted as the BF of H_1 against H_0. BF_{01} represents the reversed interpretation: evidence is in favor of the H_0. $BF_{01} = 10$ therefore means that the data are ten times more probable under H_0 than under H_1. BF is always *relative*, which means that the BF for the other hypothesis is easily determined by dividing 1 by a given BF (e.g., if BF_{01} is 10, BF_{10} is 1 divided by 10, hence 0.1). We used Bayesian statistics because, contrary to frequentist statistics, it allows for quantifying evidence in favor of the H_1 and H_0 (Dienes, 2016; Krypotos, Blanken, Arnaudova, Matzke, & Beckers, 2016).

When comparing three conditions, a Bayesian one-way ANOVA is performed first. Given sufficient evidence in favor of the H_1, this is followed up by three Bayesian *t*-tests comparing two out of three conditions each time. When Time is included as an additional within-subjects factor, a mixed Bayesian ANOVA is performed. JASP determines BFs for five models: (1) a null model, (2) a main effect of Time, (3) a main effect of Condition (4), a two main effects model: Time+Condition, (5), a two main effects model plus an Time×Condition interaction. All models are compared to the null model; therefore all BFs state the relative probability against the null model. JASP can directly provide a BF for a Time×Condition interaction, when the main effects of Time and Condition are added into the model as 'nuisance' factors. Additionally, the relative probability amongst models 2-5 can be calculated by dividing model BFs.

Although the BF is a continuous scale without arbitrary cut-offs, BFs can also be qualified by categories of evidence to facilitate scientific communication (Jeffreys, 1961; Wetzels & Wagenmakers, 2012). BFs around 1 represent evidence neither in favor of H_1 nor H_0. BFs between 1-3 ($1 - {}^{1}/_{3}$) represent anecdotal, 3-10 (${}^{1}/_{3} - {}^{1}/_{10}$) substantial, 10-30 (${}^{1}/_{10} - {}^{1}/_{30}$) strong, or 30-100 (${}^{1}/_{30} - {}^{1}/_{100}$) very strong evidence relative to the other hypothesis. A BF above 100 (or below ${}^{1}/_{100}$) is interpreted as decisive evidence for a hypothesis relative to the other hypothesis.

Results

Baseline measures

Overall, conditions showed similar performance on the baseline measures: Age, BF_{01} = 7.53; BDI-II, BF_{01} = 3.32; and TEQ, BF_{01} = 10.05 (see Table 1), which suggests that the conditions were equivalent in depression levels and number of traumatic events reported at the onset of the experiment. There was insufficient evidence for trait anxiety to draw meaningful conclusions, STAI-T, BF_{01} = 1.53.

Table 1. Baseline Measures: Sex, Age, Depression, Anxiety, and Trauma History Ratings; and Film Related Mood, Distress, Attention, and Familiarity for each Experimental Condition. Means (SD) are reported.

	No-Task Control	Tetris+Recall	Tetris Only
Sex: M/F (*n*)	7 / 18	6 / 19	9 / 17
Age	20.53 (2.623)	21.10 (2.66)	20.87 (2.46)
BDI-II	4.97 (3.11)	5.80 (3.53)	6.67 (4.92)
STAI-T	33.87 (6.10)	34.93 (6.40)	37.97 (9.60)
TEQ	0.63 (0.93)	0.63 (1.07)	0.63 (1.03)
Pre-Film Mood	14.51 (8.18)	14.86 (9.06)	15.08 (7.37)
Post-Film Mood	56.58 (16.60)	53.47 (19.39)	51.44 (17.28)
Distress	76.80 (21.13)	78.33 (20.68)	78.40 (18.92)
Attention	90.92 (8.85)	90.64 (11.91)	87.38 (10.92)
Familiarity (*n*)	0	0	1

6

Film Related Mood, Distress, and Attention

Negative mood ratings show no evidence for a Time×Condition interaction (when main effects of Time and Condition are entered as nuisance factors), $BF_{10} = 0.196$. Negative mood ratings increased after the film as evidenced by a main effect model of Time (pre vs. post), $BF_{10} > 10{,}000$ (see Table 1 and the Appendix). All other models decreased the amount of evidence compared to the main effect model of Time by a factor > 10,000 (Condition), 9.96 (Time+Condition), or 52.28 (Time+Condition + Time×Condition interaction). This shows that participants in all three conditions reported a similar increase in negative mood as a result of watching the film. Scores for distress ($M = 77.51$, $SD = 20.07$) and for attention paid to the film ($M = 90.7$, $SD = 8.94$) were high. Analyses on distress after the film and attention paid to the film showed evidence in favor of the null hypothesis, $BF_{01} = 8.85$ and $BF_{01} = 5.47$, respectively. This indicates that were no differences between conditions. Only one participant had seen the film fragment before participating in the study.

Memory Unpleasantness and Vividness

Memory unpleasantness and memory vividness of the hotspot were analyzed separately with a mixed Bayesian ANOVA with factors Time (pre vs. post-intervention) as within-subjects factor and Condition as between-subjects factor, to assess whether changes in these ratings differed between conditions (exact BFs are given in the Appendix).

For memory unpleasantness, there is no evidence for a Time×Condition interaction (with main effects of Time and Condition as nuisance factors), $BF_{10} = 0.791$. The best model is a two main effect model of Time+Condition (i.e., a main effect of Time and of Condition) closely followed by a two main effect model of Time+Condition supplemented with a Time×Condition interaction, both $BF_{10} > 10{,}000$. This shows no support for an interaction, although unpleasantness seems to decrease more for Tetris+Recall than the other interventions (see Figure 1, left panel). A main effect model of Time reduced the BF substantially by a factor 2.46 compared to the best model. A main effect model of Condition received little support, $BF_{10} = 2.026$. In summary, the results for memory unpleasantness show that all conditions decrease equally from pre to post-intervention, and that there are differences between conditions. However, decreases are not moderated by intervention type.

For memory vividness, there is evidence for a Time×Condition interaction (with Time and Condition main effects entered as nuisance factors), $BF_{10} = 7.283$. Logically, the model that is most supported by the data is the two main effects model of Time+Condition supplemented with the Time×Condition interaction, $BF_{10} > 10{,}000$. A one factor model of Time, is also supported by the data, but less compared to the best model. The best model outperformed the two main effect model of Time and Condition by a factor 6.03. Again, a main effect model of Condition received very little support, $BF_{10} = 0.132$. The analyses for memory unpleasantness support a Time×Condition interaction; there is a larger decrease for Tetris+Recall compared to the two other interventions (see Figure 1, right panel).

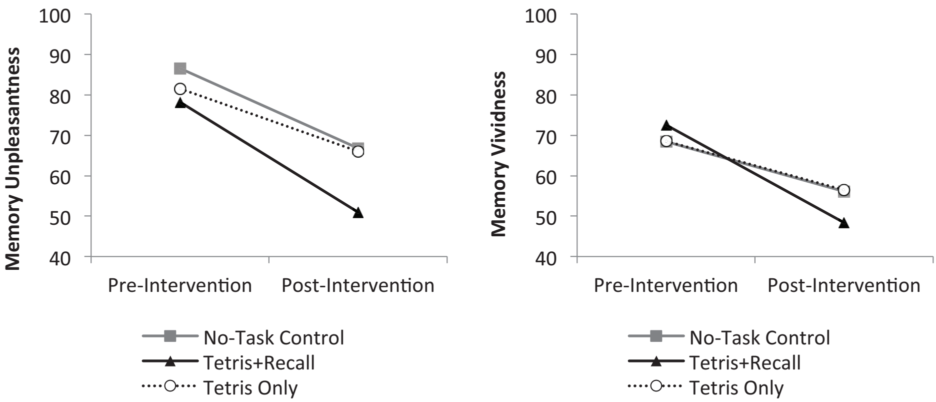


Figure 1. Pre and Post-Intervention Ratings for Memory Unpleasantness (left panel) and Memory Vividness (right panel) for No-Task Control, Tetris+Recall, and Tetris Only.

Task Performance

After the intervention-task participants reported retrospectively how they performed on the intrusion intervention task. They rated task compliance, task pleasantness, and task difficulty (see Table 2). With regard to task compliance, there were differences between conditions, BF_{10} = 11.98. Follow-up analyses showed that No-Task Control reported higher compliance compared to Tetris+Recall, BF_{10} = 32.3. The comparison No-Task Control vs. Tetris Only received minimal evidence, BF_{10} = 1.94, as did the comparison Tetris+Recall vs. Tetris Only, BF_{10} = 0.63. For task pleasantness, there were overall differences, BF_{10} = 311.66. Tetris Only was rated as more enjoyable compared to No-Task Control, BF_{10} = 1747. The other follow-up analyses, No-Task Control vs. Tetris+Recall and Tetris+Recall vs. Tetris Only, received minimal evidence, BF_{10} = 1.88 and BF_{10} = 2.38 respectively. Conditions also differed on task difficulty, BF_{10} = 8.71. While Tetris+Recall and Tetris Only seemed equivalent in reported task difficulty, BF_{01} = 2.94, they both were rated as being more difficult compared to No-Task Control, BF_{10} = 23.57 and BF_{10} = 4.15, respectively. Finally, participants also reported to what extent they recalled the memory during the intrusion intervention task. Here, there was convincing evidence for differences in memory recall, BF_{10} > 10,000. All conditions differed from each other, all BF_{10} > 66.5: memory recall was highest in Tetris+Recall, followed by No-Task Control, and then Tetris Only. This clearly shows that our manipulation of simultaneous memory recall was successful. With regard to how well participants played Tetris (i.e., their cumulative Tetris score), there was insufficient evidence for a meaningful conclusion, BF_{10} = 1.59.

Table 2. Task Compliance, Task Pleasantness, and Task Difficulty. Memory Recall (during the intervention) and Cumulative Tetris Score for each Experimental Condition. Means (SD) are reported.

	No-Task Control	Tetris+Recall	Tetris Only
Task Compliance	93.53 (7.96)	83.23 (14.11)	88.00 (11.20)
Task Pleasantness	44.8 (24.11)	58.53 (24.54)	71.37 (17.85)
Task Difficulty	19.43 (18.64)	39.53 (26.95)	34.27 (25.05)
Memory Recall	44.63 (26.84)	74.33 (19.04)	18.47 (27.15)
Cumulative Tetris Score	n/a	13292 (9210)	17182 (11467)

Intrusive Memories

In the diary collected one week after watching the film, participants reported on average 3.64 intrusions (SD = 3.21) (see Table 3). There were no differences between conditions in the number of intrusions, BF_{01} = 5.60, and in diary compliance, BF_{01} = 3.36. Participants who reported at least one intrusion were included in the analyses of intrusion characteristics of distress, vividness, unpleasantness, and suddenness. In case a participant reported multiple intrusions, we averaged over ratings for a given intrusion characteristic. We did not observe differences between conditions on intrusion characteristics of distress (BF_{01} = 3.81), vividness (BF_{01} = 7.84), unpleasantness (BF_{01} = 2.36), or suddenness (BF_{01} = 5.36). Participants also made retrospectively ratings of intrusive memories related to the film in the previous week on the IES (intrusion subscale) and PCL-5 (re-experiencing subscale). The ratings did not differ between the conditions, as evidenced by BFs in favor of the null hypothesis: BF_{01} = 9.05 and BF_{01} = 3.48, respectively. There were also no differences between conditions on the IPT, BF_{01} = 9.43.

Table 3. Number of Intrusive Memories Reported in the Diary (including ratings on Intrusion Characteristics), and scores on the IES (intrusion subscale), PCL-5 (re-experiencing subscale), IPT, and Diary Compliance for each Condition. Means (SD) are reported.

	No-Task Control	Tetris+Recall	Tetris Only
Diary Intrusions	3.1 (2.34)	4.1 (4.01)	3.73 (3.1)
Intrusion Distress[1]	3.51 (1.37)	2.93 (1.94)	2.88 (1.81)
Intrusion Vividness[1]	5.34 (1.61)	5.25 (1.96)	5.07 (1.57)
Intrusion Unpleasantness[1]	4.99 (1.87)	3.98 (2.29)	4.22 (1.98)
Intrusion Suddenness[1]	6.19 (1.93)	6.69 (1.76)	6.61 (1.26)
IES (intrusion subscale)	7.46 (4.34)	6.91 (4.37)	7.10 (4.72)
PCL-5 (re-experiencing subscale)	2.17 (1.76)	1.67 (1.83)	2.43 (1.92)
IPT	2.3 (2.4)	2.57 (2.87)	2.47 (2.50)
Diary compliance	79.20 (12.3)	79 (12.25)	74.67 (11.16)

[1] Thirteen participants were excluded from these analyses, because they reported no intrusions. The remaining participants per condition were: No-Task Control (n = 26), Tetris+Recall (n = 26), and Tetris Only (n = 25).

Discussion

The current study compared two approaches that target intrusive memories. In one approach visuospatial WM resources are taxed when participants play the game Tetris after an aversive memory is briefly primed. In the other approach WM resources are also taxed when participants perform Tetris gameplay, but simultaneously these resources are taxed because a participant recalls an aversive memory. By comparing these interventions, we tested whether the task of memory recall while playing Tetris adds to the effects of Tetris itself on reducing involuntary memories. The findings showed that the number of intrusive memories was comparable between these two conditions, and was also comparable to a No-Task Control condition. Drops in ratings of memory vividness were largest in the Tetris+Recall condition, though this pattern was absent for memory unpleasantness.

In line with literature on dual-task interventions (Engelhard et al., 2010b; see van den Hout & Engelhard, 2012) we also observed that a dual-task Tetris intervention reduced memory vividness more than a 'no intervention' control condition. We did not find that memory unpleasantness was reduced more for the Tetris+Recall condition. Though this discrepancy between memory vividness and unpleasantness is not unique (e.g., Andrade et al., 1997; Engelhard et al., 2010b; Engelhard, van den Hout, & Smeets, 2011; Maxfield, Melnyk, & Hayman, 2008), it is currently unclear why both measures did not change consistently. Based on work by Smeets, Dijs, Pervan, Engelhard, and van den Hout (2012) it is possible that changes in memory unpleasantness may only occur after memory vividness has dropped sufficiently.

Contrary to a Tetris dual-task intervention, a Tetris single-task intervention did not seem to reduce vividness and/or unpleasantness more than the 'no intervention' control condition. A likely explanation for this difference is how control conditions were set up. In control conditions in dual-task paradigms, participants usually recall the memory hotspot *without* simultaneously performing an additional intervention such as playing Tetris (i.e., Recall Only; van den Hout & Engelhard, 2012). The typical effect is no change in vividness or unpleasantness ratings (e.g., Leer, Engelhard, Altink, & van den Hout, 2013; Littel, Remijn, Tinga, Engelhard, & van den Hout, 2017; van Veen et al., 2016), or an increase from pre to post intervention (e.g., Engelhard van den Hout, Janssen, & van der Beek, 2010a; Engelhard et al., 2010b; Leer, Engelhard, & van den Hout, 2014; Maxfield et al., 2008; van den Hout et al., 2010; 2011). For the No-Task Control condition in the current experiment, vividness and unpleasantness simply seem to decay, probably as a result of time passing (see also, van Schie et al., 2017). It is difficult to ascertain whether this effect generally occurs, because previous work on intrusion modulation did not use measures of vividness or unpleasantness, but only different measures of intrusive re-experiencing (e.g., Holmes et al., 2009; 2010).

Regarding intrusive memory, our results are in contrast with earlier trauma film studies that clearly show that playing Tetris reduces the number of intrusions compared to No-Task Control (e.g., Hagenaars et al., 2017; Holmes et al., 2009; 2010; Horsch et al., 2017; Iyadurai et al., 2017; James et al., 2015). Why were we unsuccessful in finding those effects for the Tetris interventions? One possible explanation relates to the film we used. The film has been validated as aversive and intrusion provoking (Arnaudova & Hagenaars, 2017; Verwoerd et al. 2008; 2009; 2011), but the number of intrusions reported in our study is lower compared to studies using other trauma films (e.g., Hagenaars et al., 2017; Holmes et al., 2009; 2010, for an overview see Brewin, 2014). However, the film we used was not necessarily unsuccessful in inducing a sufficient number of intrusions. The same film was used in two other experiments, in which participants reported between 4.5 and 5.8 intrusions on average (in the No-Task Control condition; van Schie et al., 2017), which was slightly higher than the average of 3.64 intrusions in the current study. In other studies, fewer intrusive memories were reported, which did not prevent the primary intervention from being effective in intrusion modulation (e.g., Hagenaars, van Minnen, Holmes, Brewin, & Hoogduin, 2008; Krans et al., 2009; Stuart, Holmes, & Brewin, 2006).

Perhaps our results may be attributed to the intrusion intervention task, and Tetris specifically. Tetris was used in both active interventions, but it is possible that the active interventions may differ in WM taxation overall during the intrusion intervention task (though this was not tested). Another difference with other studies is that in the current study Tetris was played on an iPad, and in earlier studies Tetris was performed on a regular desktop computer (e.g., Holmes et al., 2009, 2010). This may suggest that larger screens are essential for effective interventions. However, recent studies show that playing Tetris on much smaller devices, such as a Nintendo DS XL, is effective as well (Horsch et al., 2017; Iyadurai et al., 2017). It could still be that playing Tetris on an iPad, in one way or another, is qualitatively different from playing on a computer or Nintendo DS XL. Perhaps because of a difference in operation: an iPad is touch sensitive, while the others are operated by keys or buttons. However, this reasoning is farfetched and if it is true, this would cast serious doubts about the robustness of Tetris interventions.

The lack of effects of Tetris without any clear methodological limitations, beckons us to question what specific elements of Tetris are supposed to make it an intrusion modulating intervention. It has been suggested that specifically its engagement of visuospatial processing (e.g., mental rotation, image formation) prevents processing of visual images of the trauma (e.g., Holmes et al., 2009). In line with this idea, other tasks, such as clay modelling (Stuart et al., 2006) and complex tapping (Holmes, Brewin, & Hennessy, 2004), also are effective in reducing intrusions. This, indeed, suggests taxation of visuospatial WM resources to some extent. Based on this idea, Asselbergs and colleagues (2017) developed two gamified interventions modelled after Tetris that should target intrusions after (analogue) trauma. Unfortunately, these interventions were not successful in reducing intrusions. Moreover, some researchers have argued that it is not visuospatial taxation per se that is necessary for intrusion modulation, but general taxation (e.g., Gunter & Bodner, 2008; van den Hout & Engelhard, 2012; Pearson & Sawyer, 2011; van Schie et al., 2017). This would also explain studies finding intrusion reducing effects for non-visuospatial tasks (e.g., Hagenaars et al, 2017; Krans et al., 2009; van Schie et al., 2017). At the very minimum these findings, taken together, raise questions concerning the supposed mechanisms underlying Tetris interventions.

These questions could give an impulse to future research in two important and complementary ways. First, the current non-replication of a Tetris intervention necessitates other independent replications of this effect to assess its robustness, preferably by setting up a multilab pre-registered replication (see Hagger et al., 2016). Admittedly, all published studies employing a Tetris intervention show favorable results, and published work with null effects for a Tetris intervention is non-existent, but this does not have to mean that this is a genuine effect per se. Independent replications remain crucial, because some effects in psychology have proven difficult to reproduce by other research groups (e.g., Open Science Collaboration, 2015). Second, research should increasingly focus on *why* interventions work (e.g., Holmes, Craske, & Graybiel, 2014; McNally, 2007; van den Hout, Engelhard, & McNally, 2017), zooming in on the exact (neuro)cognitive mechanism(s) underlying effective interventions, and also examine whether an effect is dependent on minor changes in research design. Ultimately, this should elucidate if and under what conditions a Tetris intervention is effective. Understanding the basic mechanisms that underlie intrusion modulation is an important step in properly translating findings from the laboratory to the clinic where, ultimately, these interventions may be used to treat PTSD symptoms that follow traumatic events.

Acknowledgments

This research was supported by a TOP grant (dossier number: 40-00812-98-12030) from the Netherlands Organization for Health Research and Development (ZonMw) awarded to MAvdH and IME. IME is supported with a Vici grant (453-15-005) from the Netherlands Organization for Scientific Research (NWO). We are grateful to Emily Holmes and Ella James for providing materials and helpful instructions on the trauma film paradigm, intrusion diary, and Tetris Only condition. We thank Johan Verwoerd and Ineke Wessel for providing us with the trauma film, Sarah Goerres and Derk Wisman for their assistance in testing, and Muriel Hagenaars for useful suggestions regarding study design and discussion.

Author contributions

KvS and IME developed the study concept. KvS, IME, and HK contributed to the study design. KvS was responsible for data collection and analysis. KvS drafted the manuscript and HK, MAvdH, and IME provided critical revisions.

SECTION 2

MEMORY DISRUPTION AND UPDATING DURING RECONSOLIDATION

CHAPTER 7

MODIFICATION OF EPISODIC MEMORIES BY NOVEL LEARNING: A FAILED REPLICATION STUDY

K. VAN SCHIE

S. C. VAN VEEN

M. A. VAN DEN HOUT

I. M. ENGELHARD

European Journal of Psychotraumatology, 8(sup1), 1315291 (2017)

doi:10.1080/20008198.2017.1315291

Abstract

Background: After reactivation, memories can become unstable and sensitive to modification before they are restored into long-term memory. Using behavioral manipulations, reactivated memories may be disrupted via the mechanism of interference (i.e., novel learning). In a laboratory study, Wichert et al. (2013a) showed that new learning after reactivation changed episodic memory, while new learning alone or reactivation alone did not. Given the potential clinical application of such a procedure in trauma-focused psychological treatments, such as CBT or EMDR, the aim of this study was to replicate Wichert et al.

Method: On day 1, participants ($N = 96$) viewed and recalled a series of emotional and non-emotional pictures. Then, participants were randomized to one of four groups. One week later, on day 8, Group 1 reactivated the previously learned pictures and learned new pictures. To control for specific effects of reactivation or new learning, Group 2 only reactivated the previously learned pictures, and Group 3 only learned new pictures. Group 4 received no reactivation and no new learning. On day 9, all groups indicated for each picture out of a series whether they had seen it on day 1.

Results: The data were analyzed using Bayesian hypothesis testing, which allows for quantifying the evidence in favor of the alternative and the null hypothesis. In general, results showed that Group 1 recognized fewer pictures from day 1 compared to Group 2 and 4 on Day 9. However, the expected difference between new learning following reactivation (i.e., Group 1) and new learning alone (i.e., Group 3) was not substantially supported by the data, for any of our dependent measures.

Conclusions: We replicated some of the findings by Wichert et al., but did not find substantial support for the critical difference between new learning following reactivation and new learning alone.

Keywords: Memory reconsolidation, episodic memory, replication, Bayesian statistics

Highlights

- Consolidated memories can enter an unstable phase after reactivation before being restabilized (called 'reconsolidation').
- Wichert et al. (2013a) found that only reactivated episodic memories can be altered through learning new information.
- Replication is necessary before translation to potential clinical application.

- This study could not fully replicate Wichert et al. (2013a): memory change occurred after novel learning, regardless of prior reactivation of the original memory
- The findings do not support that reconsolidation is the underlying mechanism.

Introduction

For the past 20 years, psychological science has seen a fast-growing interest in memory reconsolidation (Nader, 2015). Memory reconsolidation is the process in which reactivated, consolidated memories require a stabilization phase during which they are temporarily sensitive to amnesic agents (for an overview see Ågren, 2014; Besnard, Caboche, & Laroche, 2012; Schwabe, Nader, & Pruessner, 2014). In cognitive psychology, it has long been known that episodic memories are malleable (e.g., Loftus & Palmer, 1974), but, prior to their groundbreaking work on reconsolidation (e.g., Nader, Schafe, & LeDoux, 2000), behavioral neuroscientists believed that emotional memories were indelible (LeDoux, Romanski, & Xagoraris, 1989). A seminal study by Nader et al. (2000) brought these two research disciplines closer together by demonstrating that consolidated memories can indeed be changed. Nader et al. showed that memories are impaired when the reconsolidation process is disrupted by injecting rodents with a pharmacological agent (i.e., anisomycin) shortly after reactivation of consolidated memories. Because this agent blocks the protein-synthesis that is necessary for long-term memory formation, there was amnesia for these memories.

Reconsolidation research was extended quickly from animals to humans and started focusing on two methods of reconsolidation manipulation to test the boundary conditions of reconsolidation (Ågren, 2014). One line of research remained close to the animal studies and demonstrated experimentally that human memories can be changed via pharmacological manipulations (e.g., Brunet al., 2008; Kindt, Soeter, & Vervliet, 2009; Sevenster, Beckers, & Kindt, 2012; 2013; Soeter & Kindt, 2012), while the other line showed that memories can be altered with behavioral manipulations (e.g., Forcato et al., 2007; Hupbach, Gomez, Hardt, & Nadel, 2007; James et al., 2015; Schwabe & Wolf, 2009; Schiller et al., 2010; Wichert, Wolf, & Schwabe, 2011, 2013a, 2013b, see also van den Hout & Engelhard, 2012). Using behavioral manipulations, reactivated memories can be updated or disrupted during reconsolidation via novel learning (Ågren, 2014). An example of such a study is the one by Wichert et al. (experiment 1, 2013a), who used the canonical three-day design and normative emotional stimuli. On day 1, all four groups viewed and recalled a series of pictures (i.e., Set 1). On day 8, one week later, one of four groups reactivated the previously learned pictures and learned new pictures (i.e., Set 2). To control for specific effects of reactivation or new learning, a second group only learned new pictures without reactivation, and a third only reactivated the previously learned pictures without new learning. A fourth group received neither reactivation nor new learning. On day 9, all groups performed a recognition test in which they classified whether pictures from Set 1 and Set 2, and a set of pictures that were never seen (i.e., Set 3) were seen on day 1. On this test, only the reactivation + new learning group showed a memory impairment (i.e., lower recognition scores) compared to all other groups. Thus, Wichert et al. indeed showed that reactivation followed by interference affected consolidated memories.

Changing a memory in the lab via behavioral reconsolidation manipulations is a necessary first step in translating these findings to the psychological treatment of psychiatric disorders (Kindt & van Emmerik, 2016; Lane, Ryan, Nadel, & Greenberg, 2015). Theoretically, this could mean that in clinical practice patients first recall (i.e., reactivate) an emotionally distressing memory that is central to the psychiatric disorder. Subsequently they receive an appropriate behavioral intervention that modifies the memory. Afterwards, patients may suffer less from such an emotionally distressing memory (e.g., which may manifest itself as less intrusions in PTSD). Focusing on novel behavioral interventions is especially important, because recent studies show that frequently used pharmacological agents do not consistently affect reconsolidation in patients and healthy participants (Wood et al., 2015; for a meta-analysis see, Lonergan, Olivera-Figueroa, Pitman, & Brunet, 2013). Given the potential of reconsolidation manipulations in psychological treatments, an important question is whether changing memories during reconsolidation by use of behavioral manipulations is a finding that can be replicated reliably.

Replication of reconsolidation findings is also crucial, because recently a substantial number of studies in psychological science failed to replicate the original results (e.g., Hagger et al., 2015; Maes et al., 2016; Maslany & Campbell, 2013; Matzke et al., 2015; Zwaan & Pecher, 2012; for a large-scale collaborative attempt to replicate 100 psychological experiments see: Open Science Collaboration, 2015). This emphasizes the importance of independent replications; an appeal that has been made repeatedly in recent years to ensure the self-correcting nature of psychological science (e.g., Asendorpf, 2013; Koole & Lakens, 2012; Makel, Plucker, & Hegarty, 2012; Nosek, Spies, & Motyl, 2012; Pashler & Wagenmakers, 2012). In this spirit, we attempted to replicate Wichert et al.'s (experiment 1, 2013a) findings using similar procedures, manipulation, measures and population.

We used Bayesian statistics to analyze the data, because these statistics allow for quantifying the amount of evidence in favor of the alternative hypothesis (H_1), but also in favor of the null hypothesis (H_0; Dienes, 2016). Moreover, Bayesian statistics are increasingly used in the field of fear and trauma (e.g., Krypotos, Klugkist, & Engelhard, 2017; van de Schoot, Broere, Perryck, Zondervan-Zwijnenburg, & Van Loey, 2015). Contrary to Bayesian statistics, frequently used Null Hypothesis Significance Testing (NHST; e.g., Fisher, 1935) does not allow for a comparison of different hypotheses and only tests the evidence against the H_0. As a consequence, *p*-values above .05 cannot be interpreted as evidence in favor of the H_0. Yet, 9 out of 10 replication studies are currently evaluated almost exclusively using NHST to test whether the effect is different from zero (i.e., testing evidence against, but never in favor of, H_0; Simonsohn, 2015).

Given that a number of researchers have expressed their concerns that a large number of published research findings may be false-positive findings (e.g., Ioannidis, 2005; Simmons, Nelson, & Simonsohn, 2011), testing whether the data are evidence either for H_1 or H_0 was our primary reason to use Bayesian hypothesis testing in this replication. We tested whether or not a memory impairment is indeed specific to the reactivation + new learning group compared to the three other groups (Reactivation, New Learning, and No Reactivation + No New Learning).

Method

Participants and Design

Ninety-six students (48 men, 48 women; age: M = 21.2 years, *range* = 18-30) participated for course credit or a monetary compensation. Participants were excluded if they reported a current or chronic mental disorder, drug abuse, or current treatment with medication, or if they were younger than 18 or older than 30 years. All participants gave written informed consent. The Ethical Committee of the Faculty of Social and Behavioral Sciences at Utrecht University (FETC15-001) approved this study.

Participants were randomly assigned to one of four groups with the restriction that men and women were equally assigned to each of the groups: (1) reactivation + new learning (Re+NL); (2) reactivation (Re); (3) new learning (NL); or (4) no reactivation + no new learning (no Re+no NL, see Figure 1).

Day 1

Initial Learning

Set 1 (→ Set 1) →

Day 8

Reactivation and/or New Learning

Group	**Intervention**
Re+NL:	Re + Set 2 (→ Set 2)
Re:	Re
NL:	Set 2 (→ Set 2)
no Re+no NL:	-

→

Day 9

Recognition Testing
- old/new (Set 1, 2, and 3)
- remember/know

Picture Rating

Figure 1. Experimental design. Day 1: initial learning of Set 1; Day 8: reactivation (Re) of initially learned pictures and/or new learning (NL) of Set 2 (depending on the condition). If participants did not reach the learning criterion on Day 1 or Day 8, the set was repeated once: (→ Set 1) or (→ Set 2). Set 3 was intermixed with Set 1 and Set 2 in the recognition test, and was a set of pictures that was never seen. Day 1 and Day 9 were identical for all participants.

Stimulus materials

Because the stimulus set from Wichert et al. (2013a) was relatively small (16 pictures), which makes ceiling effects in learning pictures likely, we replaced this set and used a larger stimulus set from Wichert et al. (2013b) (L. Schwabe, personal communication, January 26, 2015). We used three sets of 60 IAPS pictures (Lang, Bradley, & Cuthbert, 2005; 30 negative and 30 neutral) that were matched for valence and on arousal based on IAPS scores (see the Appendix for specific IAPS numbers). To validate these scores in our current sample, all participants rated one picture set at the end of the experiment. Valence and arousal were rated on a visual analogue scale ranging from 0 *neutral/not arousing* to 100 *negative/very arousing*. These ratings confirmed the original IAPS classification: negative pictures were rated as more negative (M = 66.01, SD = 14.09) than neutral pictures (M = 6.81, SD = 5.72), and they were rated as more arousing (M = 54.51, SD = 16.47) than neutral pictures (M = 9.57, SD = 7.11). Wichert et al. used neutral and negative pictures, but did not find any interactions effects. For replication purposes, we kept the materials' valence matched with their study.

Procedure

In accordance with Wichert et al. (experiment 1, 2013a) testing was divided over three days: Day 1, initial learning; Day 8, reactivation and/or new learning; Day 9, recognition testing and picture rating (see Figure 1). A different set of pictures was used on each day; these sets were counterbalanced over the three days. We made small changes to the procedure of Wichert et al. (2013a) based on the study of Wichert et al. (2013b), which we detail below.

On Day 1, participants saw each picture from set 1 presented individually on a computer screen. Picture presentation was the same as in the study Wichert et al. (2013b) who used a standardized viewing time of 2s (1s intertrial interval) per picture. After the picture presentation, participants verbally recalled as many pictures in as much detail as they could. There was no time limit for this free recall test. The experimenter scored the number of recalled pictures out of the participant's sight and without giving any feedback. At least 20 out of 60 images needed to be correctly recalled in sufficient detail (e.g., 'A man pointing a gun at a woman', and not 'A gun'), so that the picture could be uniquely identified. The number of pictures that needed to be correctly recalled was comparable with what participants recalled on average in Wichert et al. (2013b) on Day 1. If this criterion was not reached, presentation and recall was repeated once. Participants continued with the experiment regardless of the number of recalled pictures after the second recall. The learning session took approximately 25 minutes.

On Day 8, the experimental procedure was different for each group. Participants in the reactivation groups (Re+NL, Re) were brought back to the same spatial context as learning on Day 1 (cf. Hupbach, Hardt, Gomez, & Nadel, 2008). During reactivation participants in these groups had two minutes to think back on the pictures that were presented on Day 1. Then, they verbally recalled the pictures they remembered. Directly after reactivation, participants in the Re+NL group were presented with new pictures from Set 2, following the same procedure and learning criterion as on Day 1. To control for specific effects of reactivation, the NL group learned and recalled Set 2 without reactivation of Set 1. This group recalled Set 2 in a spatial context that was different from Day 1. A final group (No Re+No NL) did not reactivate previously learned pictures or learned new pictures; they omitted a visit to the lab on Day 8.

On Day 9, all participants returned to the same lab as on Day 1 at approximately the same time as Day 8 (no more than 2 hours before or after the time on Day 8). On average, the time between Day 8 and Day 9 was 24 hours (SD = 1.04). On this day, participants completed a recognition test in which 180 pictures were randomly shown: 60 from Day 1, 60 from Day 8, and 60 never seen pictures (Set 3). In line with a two-step procedure (e.g., Elridge, Sarfatti, & Knowlton, 2002) participants first indicated whether they had seen the picture on Day 1 by pressing a yes or no button. If participants pressed yes, they were required to judge whether they 'remembered' seeing the picture on Day 1 or whether they 'knew' so because of a feeling of familiarity (Yonelinas, 2002). We added the remember/know distinction to explore if the subjective feeling of remembering would be affected by a behavioral reconsolidation manipulation (cf. Schwabe, Nader, & Pruessner, 2013). After the recognition test, participants gave valence and arousal ratings for pictures from set 1. Contrary to Wichert et al. (2013a), participants in our study rated valence and arousal at the end of the experiment to avoid confounds as a result of differences in encoding strength of the three sets (see also Wichert et al., 2013b). Finally, participants rated whether pictures from Day 1 were spontaneously or deliberately retrieved between Day 1 and Day 8 (i.e., strengthening of initial learning) and whether pictures from Day 1 were spontaneously or deliberately retrieved in the hours before the Day 8 appointment (i.e., reactivation opening the reconsolidation window). Ratings for all four questions were given on a 5-point Likert scale with labels: *never, rarely, sometimes, often,* and *very frequently*.

Data Analysis

All data were analyzed using the BayesFactor package (Morey & Rouder, 2015) in R (R Core Team, 2015). This package determines a Bayes Factor (BF) per requested test, which expresses the relative likelihood of the data under H_1 and the H_0. Data in favor of the H_1 are presented as BF_{10}, which can be interpreted as the Bayes Factor of H_1 against H_0. BF_{01} represents the reversed interpretation, where the evidence is in favor of the H_0. These BF representations are used when Bayesian ANOVAs are performed. When the hypothesis is directional and one group is expected to perform better or worse than another group on a given variable, a Bayesian *t*-test is used. Here, the directional alternative hypothesis (e.g., A performs better than B) is compared to a complementary null hypothesis (e.g., A performs equal or worse than B). Because we hypothesized that memory change is specific to the Re+NL group, this group was always compared to the other three groups in follow-up analyses using Bayesian *t*-tests, whenever the Bayesian ANOVA indicated there was evidence in favor of the groups being different. As a prior we used a Cauchy distribution with scale r = 0.707, which is the standard (i.e., medium) prior in the BayesFactor package. Bayesian sensitivity analyses were performed with different priors to check the robustness of the results. Though the BF is a continuous scale, BFs can also qualified by categories of evidence to facilitate scientific communication (Jeffreys, 1961; Wetzels & Wagenmakers, 2012). BFs around 1 represent evidence neither in favor of H_1 nor H_0. BFs between 1-3 represent anecdotal, 3-10 substantial, 10-30 strong, or 30-100 very strong evidence relative to the other hypothesis. A BF above 100 is interpreted as decisive evidence for a hypothesis relative to the other hypothesis. A BF_{01} of 2 therefore means that the data are twice as probable under H_0 than under H_1. Because a BF is relative, the BF for the other hypothesis is easily determined by dividing 1 by a given BF (e.g., if BF_{01} is 2, BF_{10} is 1 divided by 2, hence 0.5). Analyses for negative (BF_{neg}), neutral (BF_{neu}) or all pictures combined (BF_{all}) are presented separately.

We first present the results on initial learning on Day 1, and memory reactivation and new learning on Day 8. However, the variable crucial to our research question is recognition accuracy on Day 9, and the further break-down of that variable in false alarms and hits (for recalled and non-recalled pictures).

Results

Initial Learning on Day 1

Overall groups performed similarly and recalled a comparable total number of pictures (M = 25.34, SD = 5.34) during initial learning ($BF_{01\ all}$ = 11.13). Participants recalled more negative (M =15.04, SD = 3.53) than neutral pictures (M = 10.3, SD = 3.07). Groups did not differ in their recall of negative pictures ($BF_{01\ neg}$ = 12.49); the evidence was indecisive for neutral pictures ($BF_{01\ neu}$ = 1). Participants required on average 1.74 trials to reach the learning criterion of 20 out of 60 images (see Table 1 for overall averages).

Memory Reactivation and New Learning on Day 8

The two groups that reactivated pictures from Day 1 recalled, on average, 19.1 pictures (SD = 4.57). Again, more negative pictures (M = 11.79, SD = 2.71) were recalled, than neutral pictures (M = 7.31, SD = 2.98). The analyses that compared the two reactivation groups show that the evidence does not unambiguously favor the null or the alternative hypothesis, regardless of stimulus valence ($BF_{01\ all}$ = 0.89, $BF_{01\ neg}$ = 1.20 , $BF_{01\ neu}$ = 1.72). The two groups that viewed and recalled new pictures on Day 8 seemed to score similarly on the number of recalled pictures during new learning, regardless of stimulus valence ($BF_{01\ all}$ = 1.76, $BF_{01\ neg}$ = 3.00, $BF_{01\ neu}$ = 1.24), but the evidence in favor of the null hypothesis remains anecdotal. More negative pictures (M = 15.29, SD = 4.29) were learned than neutral pictures (M = 10.48, SD = 3.4). On average, these groups recalled 25.77 pictures (SD = 6.18).

Memory Performance on Day 9

Our primary interest is recognition accuracy. This is the percentage of correctly recognized pictures from set 1 (hits) minus the percentage of pictures from set 2 or set 3 that were incorrectly identified as being from set 1 (false alarms), and the further break-down of that variable in hits and false alarms from set 2. With this break-down, we are able to investigate whether general memory change in the different groups is the result of participants incorporating new information (i.e., false alarms of pictures from set 2) into the original memory (Hupbach et al., 2007; 2008) or whether new information only impairs the memory, but is not incorporated into the original memory (i.e., a lower percentage of hits).[3] See Table 1 for overall averages and Table 2 and 3 for averages separate for negative and neutral materials.

The analysis on recognition accuracy showed that the four groups differed ($BF_{10\ all}$ = 7.48). Interestingly, this effect was present in the negative pictures ($BF_{10\ neg}$ = 112.27), but not in the neutral pictures ($BF_{01\ neu}$ = 2.12). Follow-up analyses revealed that the evidence is prominently in favor of the hypothesis that the Re+NL group shows lower accuracy compared to the No Re+No NL group ($BFs_{10\ all,\ neg,\ neu}$ > 1398) and Re group ($BFs_{10\ all,\ neg,\ neu}$ > 7). There was only anecdotal evidence that the Re+NL group differs from the NL group ($BF_{10\ all}$ = 1.99, BF_{10neg} = 1.17, BF_{10neu} = 2.66).

A Bayesian ANOVA for the percentage of false alarms from set 2 showed that the groups differed ($BFs_{10\ all,\ neg,\ neu}$ > 7.5). Follow-up tests revealed that – in accordance with our hypotheses – the Re+NL group had a higher percentage of false alarms than the No Re+No NL group ($BFs_{10\ all,\ neg,\ neu}$ > 38.9) and Re group ($BFs_{10\ all,\ neg,\ neu}$ > 85). However, the percentage of false alarms does not seem to differ between the Re+NL group and NL group, with the strongest evidence for neutral pictures ($BF_{01\ all}$ = 1.93, $BF_{10\ neg}$ = 1.49, $BF_{01\ neu}$ = 4.91). This suggests that the extent of memory updating is similar for the Re+NL and NL groups.

For the percentage of hits analyses show that groups differ. The effect seems to be specific for negative materials learned on Day 1 ($BF_{10\ all}$ = 2.75, $BF_{10\ neg}$ = 81.9, $BF_{01\ neu}$ = 2.67). Follow-up analyses show that hits are lower in the Re+NL group compared to the No Re+No NL group ($BFs_{10\ all,\ neg,\ neu}$ > 62) and Re group ($BFs_{10\ all,\ neg,\ neu}$ > 6.9). The analysis comparing the NL group and the Re+NL group revealed that the Re+NL group showed a reduced number of hits, but only for neutral pictures ($BF_{10\ all}$ = 2.52, $BF_{10\ neg}$ = 0.84, $BF_{10\ neu}$ = 4.84). However, this final result has to be interpreted with caution since the Bayesian ANOVA did not provide evidence for an overall group difference.

Because, according to reconsolidation memories can only change when they are reactivated, we also performed analyses on hits for recalled pictures only. Re+NL and Re were directly compared, and indeed showed that Re+NL showed a reduced number of hits ($BFs_{10\ all,\ neg,}$ > 3.28), yet not for neutral pictures ($BF_{01\ neu}$ = 5.54). Though, this comparison is interesting, a more crucial comparison would be between NL and Re+NL Unfortunately, it is impossible to make this comparison with the current data. Alternatively, Wichert et al. (2013a)

[3] To foreshadow some of our findings: Bayesian sensitivity analyses were performed for all dependent variables of primary interest; recognition accuracy, false alarms, hits (recalled and non-recalled), and remember percentage. Different priors (i.e., medium, wide or ultra-wide) did not influence these results and therefore confirm that the reported results are robust.

performed an analysis on hits for non-recalled pictures only in the Re+NL and Re groups. Because non-reactivated memories do not become labile, no change is expected. Hence, these two groups should perform similar. A direct comparison, however, revealed that Re+NL scores lower than Re ($BF_{10\ all} = 9$, $BF_{10\ neg} = 96.39$, $BF_{10\ neu} = 1.48$).

Finally, we tested whether participants in the Re+NL group would display a reduced feeling of subjectively 'remembering' pictures from Day 1 compared to the other groups. Overall, analyses showed that there is substantial evidence that groups do not differ ($BF_{01\ all} = 4.6$, $BF_{01\ neg} = 5.27$, $BF_{01\ neu} = 7.03$), which shows that the behavioral reconsolidation manipulation did not influence the self-reported source of memorizing: remembering vs. knowing.

Table 1. Total Performance on Day 1 after initial learning, on Day 8 after reactivation and/or new learning, and on Day 9 for recognition testing. Means and standard deviations are presented.

Group	Day 1	Day 8		Day 9			
All Pictures	Pictures recalled after initial learning	Pictures recalled after Reactivation	Pictures recalled after New Learning	Recognition Accuracy (%)	Hits (%)	False Alarms from Day 8 (%)	Remember (%)
Reactivation + New Learning	25.96 (5.50)	20.29 (4.39)	24.63 (7.00)	63.96 (12.2)	71.87 (11.2)	11.74 (5.64)	50.39 (21.63)
Reactivation	24.83 (3.27)	17.92 (4.53)	-	73.19 (13.74)	79.51 (11.34)	6.25 (5.86)	47.4 (17.62)
New Learning	25.96 (6.33)	-	26.92 (5.12)	65.97 (16.62)	74.24 (14.3)	13.06 (13.02)	53.16 (17.42)
No Reactivation+ No New Learning	24.63 (5.92)	-	-	76.39 (10.98)	82.01 (11.14)	5.14 (6.43)	55.72 (19.3)

Table 2. Performance for Negative Pictures on Day 1 after initial learning, on Day 8 after reactivation and/or new learning, and on Day 9 for recognition testing. Means and standard deviations are presented.

Group	Day 1	Day 8		Day 9			
Negative Pictures	Pictures recalled after initial learning	Pictures recalled after Reactivation	Pictures recalled after New Learning	Recognition Accuracy (%)	Hits (%)	False Alarms from Day 8 (%)	Remember (%)
Reactivation + New Learning	15.17 (4.26)	12.42 (2.89)	15.37 (4.12)	67.50 (13.25)	76.94 (10.49)	12.92 (8.06)	63.40 (17.39)
Reactivation	14.88 (1.70)	11.17 (2.43)	-	78.89 (11.12)	86.11 (7.78)	7.08 (7.57)	63.18 (13.46)
New Learning	14.58 (3.48)	-	15.67 (3.36)	67.99 (16.18)	76.53 (13.42)	12.08 (12.50)	63.91 (13.59)
No Reactivation+ No New Learning	15.54 (4.23)	-	-	80.69 (10.15)	86.39 (8.78)	5.00 (5.73)	69.95 (15.42)

7

Table 3. Performance for Neutral Pictures on Day 1 after initial learning, on Day 8 after reactivation and/or new learning, and on Day 9 for recognition testing. Means and standard deviations are presented.

Group	Day 1	Day 8		Day 9			
Neutral Pictures	Pictures recalled after initial learning	Pictures recalled after Reactivation	Pictures recalled after New Learning	Recognition Accuracy (%)	Hits (%)	False Alarms from Day 8 (%)	Remember (%)
Reactivation + New Learning	10.79 (2.83)	7.88 (2.77)	10.17 (2.73)	60.42 (16.76)	66.81 (15.62)	10.56 (6.50)	50.39 (21.63)
Reactivation	9.96 (2.88)	6.75 (3.14)	-	67.50 (19.18)	72.92 (17.29)	5.42 (6.20)	47.40 (17.62)
New Learning	11.68 (3.29)	-	11.25 (3.26)	63.96 (19.52)	71.94 (17.80)	14.03 (14.48)	53.16 (17.43)
No Reactivation+ No New Learning	9.08 (2.93)	-	-	70.08 (13.80)	77.64 (15.15)	5.28 (9.68)	55.72 (19.30)

Self-reported Spontaneous and Deliberate Memory Retrieval

At the end of the experiment, participants indicated retrospectively to what extent the pictures came or were brought to mind between Day 1 and Day 8. Analyses show substantial evidence in favor of the null hypothesis for either spontaneous (BF_{01} = 6.85) or deliberate strengthening (BF_{01} = 8.68) of initial learning. This suggests that, overall, memories of pictures were *rarely* strengthened spontaneously (M = 2.48, SD = 0.92), or deliberately (M = 2.02, SD = 0.95). Analyses for picture reactivation show similar results: substantial evidence for the null hypothesis either for spontaneous (BF_{01} = 8.05) or deliberate reactivation (BF_{01} = 8.41). Again, in general, memories were *rarely* recalled spontaneously (M = 2.51, SD = 1.17), or deliberately (M = 2.06, SD = 1.1).

Discussion

During reconsolidation consolidated memories are temporarily sensitive to interventions that modify or update the original memory (e.g., Ågren, 2014). Memory modifications in the lab as a result of behavioral interventions are an important first step before translating these findings to clinical practice. Research into behavioral interventions during reconsolidation is especially important, because current psychological treatments are grounded in cognitive and behavioral interventions (e.g., cognitive-behavioral therapy; Rothbaum, Meadows, Resick, & Foy, 2000). Therefore, in the present study, we attempted to replicate one of the first studies using pictorial stimuli showing evidence for memory modification in episodic memory as a result of novel learning during reconsolidation, which controlled for reactivation and new learning (experiment 1, Wichert et al., 2013a). We were able to replicate some of the findings of the original study: relative to the groups Re and no Re+no NL, the group Re+NL, showed memory impairment and memory updating. However, we did not find that the crucial expected difference between Re+NL and NL alone was *substantially* supported by the data, for any of our dependent measures. Taken together with the analyses on hits for recalled and non-recalled pictures, the findings pose a challenge to predictions derived from reconsolidation theory.

The findings in our study can be explained without reconsolidation theory. The similarities between Re+NL and NL alone in our study may be the result of competition between items from the original memory and items from the new learning memory; an explanation that is in line with interference theory. The other groups, however, did not experience interference, because they did not learn something new on Day 8. Consequently, performing a recognition test for these groups should have been fairly easy. Moreover, the Wichert et al. (2013b) picture sets may have made performing the recognition test even more difficult for the new learning groups, because these picture sets had been matched not only on valence and arousal, but also on type of pictures in each set (e.g., each set contained a picture of a gun related event, a modern building, etc.), reducing the relative distinctiveness between sets. Furthermore, it is not surprising that interference was especially pronounced in neutral items (e.g., high false alarms), as non-emotional information is usually not as well remembered as emotional information is. Subsequently, this gives more opportunity for difficulties to arise in old/new recognition for neutral items (e.g., Dolan, 2002; Johnson, Hashtroudi, & Lindsay, 1993).

Our results can be explained by other theoretical accounts, such as interference theory, but it remains unclear why we did not observe effects in line with reconsolidation. Perhaps our failure to replicate the original results is related to the minor changes we made to the original design. To avoid ceiling effects we used three pictures sets of 60 pictures instead of 16 pictures, and we had participants rate one of these sets on valence and arousal at the end of the experiment instead of on Day 1. We also standardized picture-viewing time to 2s on all days, while in the original the viewing time on Day 1 was determined by how long it took to perform picture ratings. Theoretically, any of these changes might be the sole or joint cause for the absence of effects in the current study. However, this seems rather unlikely, because all these changes were derived from another study by Wichert et al. (2013b), which also tested reconsolidation and showed that new learning after reactivation changed the original memory, even with all those modifications in the method. This suggests that the changes we made probably did not influence our results.

If small deviations from the original design cannot account for why our effects deviate from the original effects, then perhaps the presence of a moderating variable can. One straightforward explanation for lack of a clear difference between the Re+NL and NL groups is unintended reactivation in the NL group. Reactivation in the NL group would also have made the consolidated memory in this group sensitive to modification. Although, this is a possibility, we made particular efforts to avoid spontaneous reactivation in the NL group on Day 8. To achieve this, participants were tested by another experimenter and in a spatial context that was different from Day 1. This context switch was necessary, because memories are directly reactivated when participants return to the original learning context (Hupbach et al., 2008). Moreover, self-reports suggest that participants in all groups rarely reactivated pictures they originally learned before they came back to the lab, which renders it unlikely that unintended reactivation can account for the Re+NL group and NL group performing similarly.

It is unclear whether our replication is a false-negative finding or the finding from Wichert et al. (2013a) was a false-positive one. There is, of course, always a possibility of a non-replication, because a given number of replication attempts are inevitably doomed to be unsuccessful due to chance. However, our study joins a growing number of other studies showing failures to replicate the reconsolidation effect using behavioral manipulations (e.g. Golkar, Bellander, Olsson, & Öhman, 2012; Hardwicke, Taqi, & Shanks, 2016; Kindt & Soeter, 2013; Soeter & Kindt, 2011; Wichert et al., 2011), or pharmacological manipulations (e.g., Bos, Beckers, & Kindt, 2014; Wood et al., 2015). The study of Hardwicke et al. (2016) makes a particularly strong case against reconsolidation effects in humans, because it attempted to directly and conceptually replicate the results of a study that has been frequently referred to as a convincing demonstration of human reconsolidation in procedural memory (Walker, Brakefield, Hobson, & Stickgold, 2003), but was unsuccessful in 7 experiments. Hardwicke et al. did not find any evidence for the impairment effects predicted by reconsolidation theory.

Still, lack of substantial support in our study clearly cannot nullify effects found in the study of Wichert et al. (2013a). Moreover, it is difficult to disregard memory reconsolidation as a whole based on a few studies that did not find reconsolidation effects; also given the considerable animal and human literature on this topic (e.g., Ågren, 2014; Schiller & Phelps, 2012). However, these studies can cast doubt on the reliability and effect sizes of previous results and subsequently on the theory itself, or at least on the boundary conditions and working mechanisms that are implied by the theory. By now, it is clear that reconsolidation is an intricate process, not merely dependent on reactivation followed by an intervention. It is a process that is suggested to be conditional on a number of boundary conditions, such as the context in which reactivation takes place (Hupbach et al., 2008), the original memory's age and strength (Wichert et al., 2011), or whether something new is learned after reactivation (i.e., prediction error; Sevenster et al., 2012, 2013). The increasing number of boundary conditions raises the question whether other yet to be empirically uncovered boundary conditions may have dampened finding indisputable reconsolidation effects in this study.

Taken together, the current failed replication study highlights a number of critical points. First, the reconsolidation process may not be as reliable or as robust as previous reconsolidation studies have suggested. To increase our confidence in which findings are trustworthy, more direct or conceptual replications will be necessary (e.g., Nosek, Spies, & Motyl, 2012; Pashler & Wagenmakers, 2012). These replications will ultimately be essential in determining whether our and previous findings on reconsolidation are reliable. They will also be crucial for uncovering genuine boundary conditions of the reconsolidation process. The fact that we were unable to induce reliable change in relatively simple memories in well-controlled environments using behavioral reconsolidation manipulations, relates to a fundamentally important question for trauma-focused therapies like cognitive-behavioral therapy, EMDR, and imagery rescripting; how can complex and strong memories related to psychiatric disorders in real-life situations be changed reliably and safely? This remains an especially pertinent question since the boundary conditions that have been found, for instance the strength of the memory, differentiates memories created in the lab from memories related to psychiatric disorders, such as PTSD. Given that changing troubling real-life memories is one of the goals in clinical practice, it is crucial that future research on reconsolidation further advances our understanding, so that we will eventually be able to bridge the gap from lab research to clinical practice.

Acknowledgements

This research was supported by a TOP grant (dossier number: 40-00812-98-12030) from the Netherlands Organization for Health Research and Development (ZonMw) awarded to

MAvdH. IME is supported with a Vici grant (453-15-005) from the Netherlands Organization for Scientific Research (NWO). We are grateful to Angelos-Miltiadis Krypotos for his help with the data-analysis in R. We thank Vera Aben, Merel van Benthem, Tim van Breda, Eunice Kasse, Tim van Lier, Liza van der Linden, Lotte Oldenhof, Marie Wischnewski, and Harmen Zoet for their assistance in testing.

Author contributions

KvS, SCvV, IME, and MAvdH designed research; KvS collected data; KvS analyzed data; and KvS, SCvV, IME, and MAvdH wrote the paper, and have read and approved the final manuscrip

CHAPTER 8

INTERVENTION STRENGTH DOES NOT DIFFERENTIALLY AFFECT MEMORY RECONSOLIDATION OF STRONG MEMORIES

K. VAN SCHIE
S. C. VAN VEEN
Y. R. HENDRIKS
M. A. VAN DEN HOUT
I. M. ENGELHARD

Neurobiology of Learning and Memory, *144*, 174-185 (2017)
doi:10.1016/j.nlm.2017.07.011

Abstract

Recently, it has become clear that retrieval (i.e., reactivation) of consolidated memories may return these memories into a labile state before they are restored into long-term memory ('reconsolidation'). Using behavioral manipulations, reactivated memories can be disrupted via the mechanism of novel learning. In the present study, we investigated whether changing a strong memory during reconsolidation depends on the strength of novel learning. To test this, participants (N = 144) in six groups acquired a relatively strong memory on Day 1 by viewing and recalling a series of pictures three times. On Day 8, these pictures were reactivated in three groups, and they were not reactivated in the other three groups. Then, participants viewed and recalled new pictures once (weak new learning) or three times (strong new learning), or they did not learn any new pictures. On Day 9, participants performed a recognition test in which their memory for Day 1 pictures was assessed. Two main results are noted. First, the groups that reactivated pictures from Day 1 and received weak or strong new learning did not differ in memory performance. Second, these two groups consistently performed similar to groups that controlled for new learning without reactivation. Because these results contradict what was expected based on the reconsolidation hypothesis, we discuss possible explanations and implications.

Keywords: Memory reconsolidation, episodic memory, novel learning; updating

Highlights

- Strong memories were changed by novel learning regardless of prior reactivation
- Changing reactivated strong memories did not depend on novel learning strength
- Novel learning impaired and updated the original strong memory

Introduction

More than a century ago, Müller and Pilzecker (1900) proposed the memory consolidation hypothesis. It stated that memories of newly learned information are initially in a dynamic, labile state before they become fixed in long-term memory ('consolidation'), and that once they are consolidated, they are resistant and insensitive to interference by distracting stimuli, injuries, or toxins (see McGaugh, 2000; see Dudai, 2004; though see, Loftus & Palmer, 1974). Recently, however, it has become clear that retrieval (i.e., reactivation) of consolidated memories may return these memories into a labile state before they are re-stabilized ('reconsolidation') (Nader, Schafe, & LeDoux, 2000). During reconsolidation, the memory trace can be strengthened, weakened, or updated (for an overview see Ågren, 2014; Besnard, 2012; Schwabe, Nader, & Pruessner, 2014).

Disruption of memory during the reconsolidation process was first observed in animals that received amnesic agents (i.e., anisomycin) shortly after memory reactivation of conditioned threat memories (e.g., Dębiec, LeDoux, & Nader, 2002; Nader et al., 2000). These agents blocked protein-synthesis, which is necessary for long-term memory formation, and caused amnesia for the original threat memory. Research on human reconsolidation quickly followed and primarily targeted conditioned threat memories with pharmacological manipulations that were safe for humans such as propranolol (e.g., Brunet al., 2008; Kindt, Soeter, & Vervliet, 2009; Sevenster, Beckers, & Kindt, 2012; 2013; 2014; Soeter & Kindt, 2012). However, recent studies – including a meta-analysis – have shown that the effects of pharmacological interventions on human reconsolidation are not consistent (Lonergan, Olivera-Figueroa, Pitman, & Brunet, 2013; Wood et al., 2015).

An alternative approach to disrupting memories during the reconsolidation process is by use of behavioral manipulations (Forcato et al., 2007; Hupbach, Gomez, & Nadel, 2009; Hupbach, Gomez, Hardt, & Nadel, 2007; Hupbach, Hardt, Gomez, & Nadel, 2008; James et al., 2015; Kredlow & Otto, 2015; Schiller et al., 2010; Schwabe & Wolf, 2009; Wichert, Wolf, & Schwabe, 2011, 2013a, 2013b). Behavioral manipulations that change memories during reconsolidation typically focus on the mechanism of memory updating or interference after reactivation, and usually follow a three-day design that mirrors research with amnesic agents. On Day 1 encoding of a novel memory takes place. On Day 2, participants in the experimental group reactivate the memory, and then encode new, yet comparable memories. Typically, there are also conditions that only reactivate the memory, only receive the manipulation, or do not get tested on this day (i.e., control conditions). On the third day, there is a test of memory strength of the original encoded material. Overall, these studies have shown that new learning after reactivation on Day 2 yields effects on Day 3. They found deteriorated recall of the original encoded material, which was interpreted as updating of the original memory by incorporating the newly encoded information (e.g., Schwabe et al., 2014).

Several boundary conditions have been identified which preclude memories from being disrupted during reconsolidation (e.g., Schwabe et al., 2014). For instance, older and stronger memories are more resistant to post-reactivation modification compared to younger and weaker memories (Eisenberg, Kobilo, Berman, & Dudai, 2003; Wichert et al., 2011). Some studies show that these boundary conditions can be overcome and that post-reactivation modification during reconsolidation can still take place (e.g., Wang, Alvares, & Nader, 2009; Winters, Tucci, & DaCosta-Furtado, 2009). For instance, Wichert, Wolf, and Schwabe (2013a) showed that, compared to weak episodic memories, strong episodic memories (i.e., memories that were repeatedly reactivated without subsequent new learning) are more resistant to the effect of new learning during reconsolidation. The effects of new learning interventions resulted primarily in loss of the original memory for those memories that were weak. This happened to a smaller extent for memories that were strong. There were no signs of updating (i.e., incorporation of new information into the original memory) for either weak or strong memories. Overall, this shows that strong memories are still sensitive to modification during reconsolidation, but to a lesser extent than weak memories.

It is likely that the extent to which memories can be changed is not solely dependent on the strength of the initial memory, but also on the strength of the post-reactivation manipulation. In animal research, the impact of the post-reactivation manipulation is dose-dependent: increasing the dose of the amnesic agent increases memory impairment (Duvarci, Nader, & LeDoux; 2008; Nader et al., 2000). Recently, Wichert, Wolf, and Schwabe (2013b) investigated the effect of different doses of behavioral post-reactivation manipulations on the memory reconsolidation process in humans. On Day 1 in their experiment, participants (in six groups) acquired a relatively weak memory by viewing and recalling a series of pictures once. One week later (Day 8), these pictures were reactivated in three groups, and they were not reactivated in the other three groups. Then, of these three groups, one group learned new pictures once (weak manipulation), another learned new pictures three times (strong manipulation), and another did not learn new pictures. Hence, the number of exposures to new materials primarily determined the intervention's strength. Two weeks after the testing first day (Day 15), participants performed a recognition test in which they saw the original pictures from Day 1, the newly learned pictures from Day 8, and a completely novel set of pictures. These pictures were intermixed and for each picture participants were asked to classify whether they had seen it on Day 1 by pressing a yes or no button. In line with findings from pharmacological manipulations, Wichert et al. showed that, following reactivation, learning new pictures once had no effect on episodic memory while learning these pictures three times did. That is, the reactivation + strong manipulation group showed memory change compared to all other groups. This change reflects incorporation of new information into the original memory, but no loss of the original memory. This shows that the extent of modification of a relatively weak memory is affected by the strength of new learning after reactivation. However, it is currently unclear how post-reactivation manipulations affect *strong* memories.

Knowledge about how to change strongly encoded memories can have important implications for the behavioral treatment of psychiatric disorders where dysfunctional memories are a core feature (Dębiec, 2012; Parsons & Ressler, 2013; Schwabe et al., 2014). For instance, core symptoms of post-traumatic stress disorder are intrusive memories of a traumatic event, which are memories that are overconsolidated due to the release of stress hormones in reaction to the traumatic event (Pitman, 1989), and the intrusive and repetitive nature of the traumatic memory (Hackmann, Ehlers, Speckens, & Clark, 2004). Using behavioural interventions during reconsolidation in clinical practice could entail that patients recall (i.e., reactivate) their memory and receive an appropriate intervention that modifies the memory (Beckers & Kindt, 2017; Lane, Ryan, Nadel, & Greenberg, 2015). Given this potential clinical application, an important question is how strong memories can be changed during reconsolidation by post-reactivation behavioral manipulations that differ in number of exposures.

In the present study, we investigated whether reconsolidation of *strong* memories is dependent on the strength of the post-reactivation manipulation. To test this, we used the three-day paradigm developed by Wichert et al. (2013b), but made three modifications: the strength of the initial memory on Day 1, the time between the final two days, and we added the remember/know distinction. Therefore, in our study on Day 1, participants in all six groups acquired a relatively strong memory by viewing and recalling a series of pictures three times (instead of once). On Day 8 – one week later –, these pictures were reactivated in three groups, and they were not reactivated in the other three groups. Then, participants viewed and recalled new pictures once (weak novel learning) or three times (strong novel learning), or they did not learn any new pictures. On day 9, which was one day instead of one week later, participants performed a recognition test in which they saw the picture sets from day 1 and day 8, and a novel picture set. For each picture, they decided whether they had seen it on day

1 by pressing a yes or no button, and whether they associated it with a feeling of remembering or knowing. The shortened time interval of 24h between intervention and test was based on Wichert et al. (2013a). Theoretically, the duration of this interval is largely irrelevant as long as the test on the final day is placed outside the reconsolidation window (approximately 4 to 6 hours after reactivation), which is the case in the current study. Moreover, with a shortened interval reconsolidated memories are less prone to general memory decay. Given that strong memories remain sensitive to modification during reconsolidation (see earlier pharmacological studies) and that post-reactivation manipulations are dose-dependent, we expected that strongly consolidated memories would be altered more by applying the strong post-reactivation manipulation after reactivation than by the weak manipulation.

Method

Participants and Design

Students from Utrecht University and the University of Applied Sciences ($N = 144$, 72 men, 72 women; $M_{age} = 21.14$ years, $SD = 2.27$) participated for course credit or a monetary compensation. Participants were excluded if they reported a current or chronic mental disorder, drug abuse, current treatment with medication, or if they were younger than 18 or older than 30 years. All participants provided written informed consent. The Ethical Committee of the Faculty of Social and Behavioral Sciences at Utrecht University (FETC15-001) approved this study.

Participants were assigned to one of six groups using block randomization. Groups were constituted by whether there was reactivation (yes vs. no) and new learning strength (strong vs. weak vs. none). The reactivation + strong new learning (RE+SNL) and the reactivation + weak new learning (RE+WNL) were of primary interest in our study. The other four groups were control groups that control for effects of reactivation or new learning: reactivation + no new learning (RE+NNL), no reactivation + strong new learning (NRE+SNL), no reactivation + weak new learning (NRE+WNL), or no reactivation + no new learning (NRE+NNL, see Figure 1). An equal number of men and women were assigned to each group.

Day 1

Initial Learning

Set 1 → Set 1 → Set 1 →

Day 8

Reactivation and/or Strong, Weak or No New Learning

Group	**Intervention**
RE+SNL:	Re + Set 2 → Set 2 → Set 2
RE+WNL:	Re + Set 2
RE+NNL	Re
NRE+SNL:	Set 2 → Set 2 → Set 2
NRE+WNL:	Set 2
NRE+NNL:	

→

Day 9

Recognition Testing
- old/new (Set, 1, 2, and 3)
- remember/know

Picture rating

Figure 1. Experimental design. Day 1: initial learning by viewing and recalling Set 1 three times; Day 8: reactivation (RE) or no reactivation (NRE) of initially learned pictures and/or new learning (i.e., Strong, Weak or None) of Set 2 (depending on the condition). Pictures in Set 2 were viewed and recalled three times (Strong New Learning; SNL), once (Weak New Learning; WNL), or not at all (No New Learning; NNL); Day 9: participants judged whether a picture was seen on day 1 and if so, whether this was associated with a feeling of remembering or knowing. Set 3 was intermixed with Set 1 and 2 in the recognition test, and was new for participants. Day 1 and Day 9 were identical for all participants.

Stimulus materials

We used 180 pictures (90 negative, 90 neutral), divided into three equal sets of 60 pictures (30 negative and 30 neutral), that were also used by Wichert et al. (2013b)[4]. These pictures were all originally taken from the International Affective Picture System (IAPS; Lang, Bradley, & Cuthbert, 2008) and were matched based on IAPS standard scores for valence and on arousal (see the Appendix for specific IAPS numbers). To confirm these scores, all participants rated one set of pictures after the experiment. Valence and arousal were rated on two visual analogue scales ranging from 0 *neutral/not arousing* to 100 *negative/very arousing* (Wichert et al., 2013b). These ratings confirmed the picture classification: negative pictures were rated as more negative (M = 63.74, SD = 15.71) than neutral pictures (M = 7.21, SD = 5.46) and were rated as more arousing (M = 53.61, SD = 16.48) than neutral pictures (M = 12.45, SD = 9.82).

Procedure

Testing took place divided over three days (see Figure 1). On each day, a different set of pictures was used; these sets were fully counterbalanced over days.

[4] We received the IAPS numbers that were used in Wichert et al. (2013b). However, 39 out of 180 pictures were no longer in the IAPS database. We replaced these with other pictures depicting similar content.

On Day 1, participants saw 60 negative and neutral pictures (Set 1) on screen for 2s each (1s intertrial interval). Afterwards, they performed a free recall test without a time limit and verbally described as many pictures in as much detail as they could. The experimenter scored a picture as correct when a participant's description involved a unique picture identification. This was done out of the participant's sight and without giving feedback. The experimenter only intervened by briefly repeating the instructions when the participant did not recall the pictures in detail. Then, participants viewed and recalled the same set of pictures two additional times to ensure strong encoding (Karpicke & Roediger, 2008). The learning session took approximately 35 minutes.

On Day 8, the experimental procedure differed per group. Participants in the reactivation conditions (RE+SNL, RE+WNL, and RE+NNL) were given 2 minutes to think back of the pictures that were presented on Day 1; they then verbally recalled as many pictures in as much detail as they could. Reactivation always took place in the same spatial context as learning on Day 1. Directly after reactivation, participants in two RE+NL conditions learned and recalled the second picture set (Set 2) once (RE+WNL) or three times (RE+SNL) following the same procedure as on Day 1. To control for the effects of new learning alone, two groups learned and recalled the second picture set, either once (NRE+WNL) or three times (NRE+SNL) without reactivation of the first picture set. For these two groups, new learning was performed in a spatial context that was different from Day 1 and with an experimenter that was different from Day 1. This context switch was necessary, because memories are directly reactivated when participants return to the original learning context (Hupbach et al., 2008). A final group (NRE+NNL) did not reactivate previously learned pictures or learned new pictures; they did not come to the lab on Day 8.

On Day 9 – as close as possible to 24 hours after Day 8 (M = 24.13 hours, SD = 1.98) – participants completed a recognition test in which 180 pictures were shown: 60 from Day 1, 60 from Day 8, and 60 new pictures. Following a two-step procedure (e.g., Elridge et al., 2002) participants first had 3s to indicate whether they had seen the picture on Day 1 by pressing a yes or no button. If they pressed yes, they had another 3s to judge whether they 'remembered' seeing the picture on Day 1 or whether they 'knew' this because of a feeling of familiarity (Yonelinas, 2002). We explained that remembering meant that a picture evoked specific memories of what was experienced during its presentation (e.g., how it was presented or what the participant was thinking at that time). Knowing meant being confident that a picture appeared, but without recollection of any aspects of its presentation. The remember/know instructions were taken from Elridge et al. (2002, p140) and were translated to Dutch. We added the remember/know distinction in the current study to explore whether the memory manipulation would affect the subjective feeling of remembering (cf. Schwabe, Nader, & Pruessner, 2013). After the recognition test, participants gave valence and arousal ratings for pictures from Set 1. Lastly, participants rated on a 5-point Likert scale whether pictures from Day 1 were spontaneously or deliberately retrieved between Day 1 and Day 8 (i.e., strengthening of initial learning) and whether pictures from Day 1 were spontaneously or deliberately retrieved in the hours before the Day 8 appointment (i.e., opening the reconsolidation window). Ratings for all four questions were given with the labels: *never (1), rarely (2), sometimes (3), often (4),* and *very often (5)*.

Data preparation

Data were normally distributed, except for three variables: recognition accuracy, hits, and false alarms. Because of substantial negative skewness, the data for recognition accuracy and hits were sign-reversed (i.e., negative skewness was transformed to positive skewness) and then log transformed (Tabachnik & Fidell, 2013). As a result, the interpretation of these variables needs to be reversed as well (e.g., a higher score on recognition accuracy needs to be interpreted as *lowered* recognition accuracy). The data for false alarms were only log transformed because of positive skewness, which does not change that variable's interpretation. See Appendix B for the non-transformed scores of recognition accuracy, hits, and false alarms.

Data Analysis

The BayesFactor package in R was used for data-analysis (Morey & Rouder, 2015; R Core Team, 2015). With this package a Bayes Factor (BF) is determined per requested test, which expresses the relative likelihood of the data under H_1 and H_0. A BF represented as BF_{10} can therefore be interpreted as the Bayes Factor of H_1 relative to H_0 (i.e., evidence is in favor of the H_1 compared to H_0), while BF_{01} represents the reversed interpretation. $BF_{01} = 4$ therefore means that the data are four times more probable under H_0 than under H_1. BF is always *relative*, which means that the BF for the other hypothesis is easily determined by dividing 1 by a given BF (e.g., if BF_{01} is 4, BF_{10} is 1 divided by 4, hence 0.25).

First, Bayesian ANOVAs were performed to test for group differences. Because we hypothesized that memory change is specific to the RE+WNL and RE+SNL groups, these groups were always compared to the other four groups in follow-up analyses using Bayesian *t*-tests. A final comparison was made between the RE+SNL group and RE+WNL group to test whether memory change was indeed strongest in the RE+SNL group. In the follow-up analyses, the directional alternative hypothesis (e.g., A performs better than B) is always compared to a complementary null hypothesis (e.g., A performs equal *or* worse than B). In all analyses, we used the BayesFactor package's standard prior: a Cauchy distribution with scale r = 0.707 (i.e., medium prior). Additionally, wide and ultra-wide were used in the Bayesian sensitivity analyses to check the results' robustness.

The BF originally is a continuous scale and is best interpreted this way, but to facilitate scientific communication, BFs can also be interpreted qualitatively by use of categories of evidence (Jeffreys, 1961; Wetzels & Wagenmakers, 2012). The cut-offs of these arbitrary categories should not be viewed as absolute (e.g., a BF of 2.9 is equally valuable as a BF of 3.0). BFs around 1 represent evidence that is not in favor of H_1 or H_0. BFs between 1-3 (1-$1/3$) represent anecdotal, 3-10 ($1/3 - 1/10$) substantial, 10-30 ($1/10 - 1/30$) strong, or 30-100 ($1/30 - 1/100$) very strong evidence. A BF above 100 (or below $1/100$) is interpreted as decisive evidence for a given hypothesis. Analyses for negative (BF_{neg}), neutral (BF_{neu}), or all pictures combined (BF_{all}) are presented separately.

Results

Initial Learning on Day 1

Overall, there was an increase in the number of pictures recalled during the first (M = 13.76, SD = 5.04), second (M = 24.24, SD = 7.98) and third repetition (M = 31.67, SD = 9.31). On the third and final repetition, participants in the six groups recalled, as expected, a comparable total number of pictures during initial learning ($BF_{01\ all}$ = 31.62, $BF_{01\ neg}$ = 26.46, $BF_{01\ neu}$ = 24.61; see Table 1 for total averages). Participants recalled more negative (M = 17.75, SD = 4.99) than neutral pictures (M = 13.92, SD = 5.17), BF_{10} > 10,000. Overall, this shows that there are no differences between the groups in encoding the initial memories.

Table 1. Total Performance on Day 1 after initial learning, on Day 8 after reactivation and/or new learning, and on Day 9 in the recognition test. Means and standard deviations are presented.

Group	Day 1	Day 8		Day 9			
	Number of pictures recalled after initial strong learning	Number of pictures recalled after Reactivation	Number of pictures recalled after New Learning	Recognition Accuracy	Hits	False Alarms from Day 8	Remember (%)
RE+SNL	31.13 (9.35)	22.88 (9.84)	30.00 (10.67)	1.27 (0.34)	1.13 (0.34)	0.79 (0.46)	56.15 (17.64)
RE+WNL	32.13 (7.80)	23.50 (8.55)	15.46 (4.86)	1.34 (0.26)	1.17 (0.35)	0.85 (0.44)	60.72 (22.95)
RE+NNL	33.04 (10.02)	24.63 (8.03)	-	0.83 (0.40)	0.64 (0.46)	0.31 (0.38)	66.47 (18.02)
NRE+SNL	32.79 (10.77)	-	33.00 (8.42)	1.22 (0.27)	1.07 (0.26)	0.75 (0.42)	53.42 (15.59)
NRE+WNL	30.88 (9.52)	-	15.29 (5.68)	1.36 (0.20)	1.14 (0.23)	1.11 (0.31)	66.43 (15.02)
NRE+NNL	30.08 (8.72)	-	-	0.88 (0.43)	0.75 (0.45)	0.32 (0.37)	55.97 (22.56)

Note. Recognition accuracy and Hits are sign-reversed and log transformed scores. For these scores the interpretation of such a variable needs to be reversed as well; for instance, a higher score on recognition accuracy needs to be interpreted as *lowered* recognition accuracy. False alarms from Day 8 were log transformed only; a higher value reflects higher false alarms. Remember (%) reflects the percentage of Day 1 pictures correctly identified as being from Day 1 that were said to be remembered. RE = Reactivation, NRE = No Reactivation, SNL = Strong New Learning., WNL = Weak New Learning, NNL = No New Learning

8

Memory Reactivation and New Learning on Day 8

The three reactivation groups recalled, on average, 23.67 pictures (*SD* = 8.75) from Day 1. As expected, the three groups did not differ in the number of pictures that were recalled ($BF_{01\ all}$ = 7.06, $BF_{01\ neg}$ = 5.87, $BF_{01\ neu}$ = 6.74). Again, they recalled more negatives pictures (*M* = 14.04, *SD* = 5.19) than neutral pictures (*M* = 9.63, *SD* = 4.34), BF_{10} > 10,000.

Four groups viewed and recalled new pictures on Day 8: two groups received the WNL intervention, and two groups received the SNL intervention. Regarding new learning, there was decisive evidence that these four groups differed from each other on the number of recalled pictures (all BF_{10} > 10,000). We checked whether the two SNL groups recalled more pictures than the two WNL groups (i.e., a successful manipulation of new learning). These analyses showed that this was the case, regardless of valence: the RE+SNL and NRE+SNL groups recalled more pictures (respectively, *M* = 30.00, *SD* = 10.67, and *M* = 33.00, *SD* = 8.42) than the RE+WNL and NRE+WNL groups (respectively, *M* = 15.46, *SD* = 4.86, and *M* = 15.29, *SD* = 5.68; all BF_{10} > 4350). There appeared to be no differences in the number of recalled pictures between RE+SNL and NRE+SNL ($BF_{01\ all}$ = 5.20, $BF_{01\ neg}$ = 2.28, $BF_{01\ neu}$ = 2.46) or between RE+WNL and NRE+WNL ($BF_{01\ all}$ = 3.46, $BF_{01\ neg}$ = 3.28, $BF_{01\ neu}$ = 2.89). These comparisons do not show strong support for H_0, but they still indicate that this hypothesis is more plausible than the H_1. Overall, more negative (*M* = 14.04, *SD* = 5.19) than neutral pictures (*M* = 9.63, *SD* = 4.34) were recalled during new learning, BF_{10} > 10,000. Taken together, these results show that our Day 8 manipulations of reactivation and different strengths of new learning were successful.

Memory Performance on Day 9

Central to this study is recognition accuracy. This is the percentage of correctly recognized pictures from Set 1 (hits) minus the percentage of pictures from both Set 2 and Set 3 that were incorrectly identified as being from Set 1 (false alarms). We also individually analyzed hits and only the false alarms from Set 2 to investigate whether new information (i.e., false alarms of pictures from Set 2) was incorporated into the original memory (Hupbach et al., 2007; 2008), or whether new information after reactivation only impaired the memory (i.e., a lower percentage of hits). This break-down allowed us to investigate whether the hypothesized reduced recognition accuracy in the Re+NL groups is related to higher memory impairment or more memory updating.[5] See Table 1 and Figure 2 for visual representations (See Appendix C and D for these data split out by picture valence)

[5] Bayesian sensitivity analyses were performed for all dependent variables of primary interest: recognition accuracy, false alarms, hits, and remember percentage. Different priors (i.e., medium, wide

The analysis on recognition accuracy showed that the six groups differed regardless of picture valence (all BF_{10} > 10000). Follow-up analyses revealed that the RE+WNL group displayed a lowered accuracy compared to three control groups, namely RE+NNL, NRE+SNL, and NRE+NNL (all BF_{10} > 5.25), but not compared to the NRE+WNL group ($BF_{10\ all}$ = 0.68, $BF_{10\ neg}$ = 0.36, $BF_{10\ neu}$ = 0.98). RE+SNL also displayed a lowered accuracy compared to the same three control groups (all BF_{10} > 3.15), but compared to NRE+SNL the BF provided some support for it coming from the H_1 ($BF_{10\ all}$ = 2.21, $BF_{10\ neg}$ = 1.34, $BF_{10\ neg}$ = 3.15). Furthermore, the RE+SNL did not differ from NRE+WNL ($BF_{10\ all}$ = 0.20, $BF_{10\ neg}$ = 0.06, $BF_{10\ neu}$ = 2.87). The crucial comparison between RE+SNL and RE+WNL showed evidence, in contrast to the hypothesis, for comparable recognition accuracy ($BF_{01\ all}$ = 3.18, $BF_{01\ neg}$ = 5.41, $BF_{01\ neu}$ = 2.61). Overall, the results of recognition accuracy show that both RE+SNL or RE+WNL generally differ from three out of four control groups, except NRE+WNL. Both RE+SNL and RE+WNL are similar to NRE+WNL in recognition accuracy. RE+SNL and RE+WNL did not differ from each other.

For the percentage of hits, analyses show that the groups differed (all BF_{10} > 4780). As expected, the RE+WNL group had a lower number of hits compared to three control groups regardless of stimulus valence (all BF_{10} > 3.08). There is some evidence that RE+WNL differed from the control group NRE+WNL ($BF_{10\ all}$ = 1.82, $BF_{10\ neg}$ = 1.23, $BF_{10\ neu}$ = 1.45), but this evidence is weak. RE+SNL displayed lowered hits compared to two control groups (RE+NNL, and NRE+NNL, all BF_{10} > 24). Again, when comparing RE+SNL and NRE+SNL the evidence was in the expected direction: it was almost 2.5 more likely that the data came from H_1 than from H_0 ($BF_{10\ all}$ = 2.47, $BF_{10\ neg}$ = 0.61, $BF_{10\ neu}$ = 2.30). The comparisons between RE+SNL and NRE+WNL hardly provided any support for the H_1 or the H_0 ($BF_{01\ all}$ = 1.23, $BF_{01\ neg}$ = 3.57, $BF_{01\ neu}$ = 1.33). RE+SNL was expected to show lowered hits compared to RE+WNL, but the percentage of hits showed a tendency to be the same in these two groups ($BF_{01\ all}$ = 1.99, $BF_{01\ neg}$ = 3.68, $BF_{01\ neu}$ = 0.60). In conclusion, both RE+SNL and RE+WNL groups differed consistently from two control groups (RE +NNL and NRE+NNL). RE+SNL and RE+WNL groups do not differ or only very little from NRE+WNL. Also, there is some evidence that RE+SNL and RE+WNL do not differ from each other.

For the percentage of false alarms, we focused specifically on false alarms from Set 2. According to the concept of memory reconsolidation, new information should be

or ultra-wide) did not influence these results and therefore confirm the robustness of our results. Therefore, we only report the BFs for the BayesFactor package's standard prior.

incorporated into the original memory after reactivation (i.e., memory updating; Lee, 2009). False alarms from Set 2 are indicative of this process (Hupbach et al., 2007). We included NRE+NNL and RE+NNL as a comparison to other experimental groups to ascertain the possible interfering effect of Set 2 on memory for Set 1. Overall, the six groups differed in the percentage false alarms from Set 2 (all BF_{10} > 10,000). RE+WNL indeed scored higher than three control groups, namely RE+NNL, NRE+SNL, and NRE+NNL (all BF_{10} > 3.33). However, RE+WNL did not have a higher percentage of false alarms compared to NRE+WNL ($BF_{10\ all}$ = 0.02, $BF_{10\ neg}$ = 0.01, $BF_{10\ neu}$ = 0.09). This can also be clearly observed in Figure 1 (right panel), where, unexpectedly, NRE+WNL has more false alarms from Set 2 than any of the five other groups. As hypothesized, RE+SNL showed a similar pattern and had more false alarms compared to NRE+NNL and RE+NNL (all BF_{10} > 72). However, compared to NRE+SNL, there was hardly any support provided for the H_1 or the H_0 ($BF_{10\ all}$ = 1.63, $BF_{10\ neg}$ = 0.57, $BF_{10\ neu}$ = 4.15), except for neutral items where the evidence supports the H_1. Here as well and contrary to our hypothesis, NRE+WNL displayed more false alarms than RE+SNL ($BF_{01\ all}$ = 138, $BF_{01\ neg}$ = 791, $BF_{01\ neu}$ = 10.95). Unexpectedly, there was a tendency for RE+SNL and RE+WNL to score similarly on false alarms ($BF_{01\ all}$ = 1.91, $BF_{01\ neg}$ = 6.21, $BF_{01\ neu}$ = 0.98), though this effect seemed to vary with stimulus valence. Again, collectively this shows that SNL or WNL show similar or worse effects compared to RE+SNL and RE+WNL. Differences between RE+WNL and RE+SNL seem to be small.

Finally, we tested whether the feeling of subjectively 'remembering' pictures from Day 1 was reduced in participants the RE+WNL and RE+SNL compared to the other groups. Overall, the analyses showed that there is some tendency that the groups were similar ($BF_{01\ all}$ = 1.67, $BF_{01\ neg}$ = 6.17, $BF_{01\ neu}$ = 0.83). This could suggest that a behavioral reconsolidation manipulation may not influence the self-reported source of memorizing: remembering vs. knowing. However, note that there currently is unconvincing evidence to come to this conclusion.

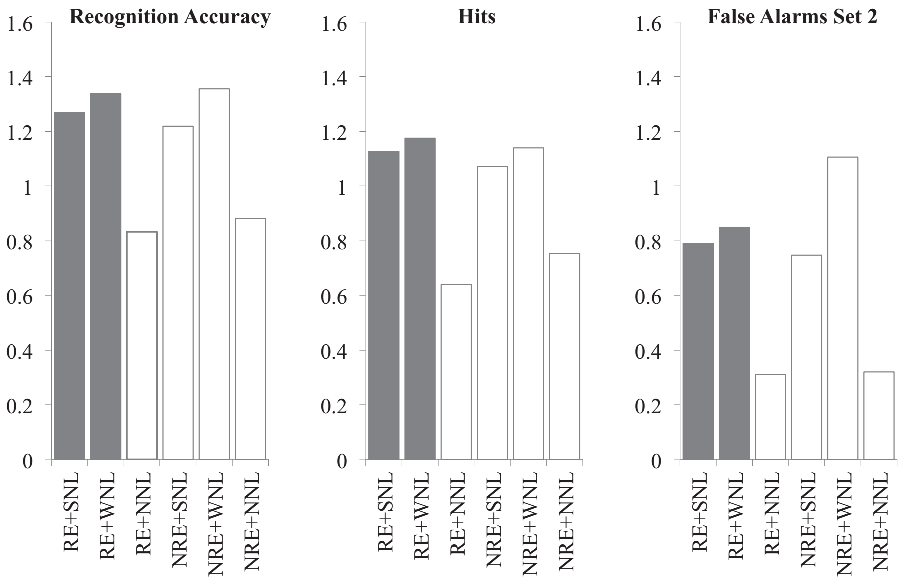


Figure 2. Total Performance on Day 9 for recognition testing. The left panel depicts sign-reversed and transformed recognition accuracy scores (hits minus all false alarms). The center panel depicts the sign-reversed and transformed percentage of hits. Sign-reversed scores for recognition accuracy and hits should be reversed interpreted; a higher bar reflects a lower score. The right panel depicts the transformed percentage of false alarms from Set 2 pictures. Control groups are depicted in white. Note. RE = Reactivation, NRE = No Reactivation, SNL = Strong New Learning., WNL = Weak New Learning, NNL = No New Learning

Self-reported Spontaneous and Deliberate Memory Retrieval

At the end of the experiment, participants indicated retrospectively to what extent the pictures spontaneously came to mind or were brought to mind deliberately between Day 1 and Day 8 (i.e., strengthening of initial learning) and a couple of hours before the Day 8 session (i.e., opening the reconsolidation window). The evidence is in favor of groups reporting similar strengthening of initial learning, which shows that overall memories of pictures were *rarely* strengthened spontaneously (BF_{01} = 5.06, M = 2.63, SD = 0.93), or deliberately (BF_{01} = 4.72, M = 2.07, SD = 1.05). Analyses for picture reactivation show comparable results: the different groups show similar spontaneous reactivation (BF_{01} = 19.47, M = 2.56, SD = 1.15) and deliberate reactivation (BF_{01} = 8.67, M = 2.02, SD = 1.17). Also, for reactivation, memories were *rarely* recalled spontaneously or deliberately in general.

Exploratory analyses

Individual between group comparison seems to suggest that reactivation followed by weak or strong new learning is not necessary per se to achieve memory impairment, but that weak or strong new learning itself is sufficient. As a purely post-hoc exploration, we examined whether the data (of negative and neutral pictures combined) of the six groups *as a whole* were in line with one of these ideas. Each idea can be modeled by specific (in)equality constraints between the six different groups. This constrained model can then be tested against another constrained model. Model 1 states that RE+WNL as well as RE+SNL groups should perform more poorly than the four control groups (i.e. worse recognition accuracy, lower on hits, higher on false alarms). Model 2 states that the four NL groups should perform equally, but all worse than RE+NNL and NRE+NNL. We used BIEMS (Mulder et al., 2009) which is a software package that computes a BF per model taking all (in)equality constraints into account simultaneously. Competing constrained models can then be compared directly. Model 1 received BFs of 2, 4.5 and 0.02 for recognition accuracy, hits, and false alarms from Set 2 respectively. Model 2 received BFs of 44, 99.7 and 1.01. Although this clearly is a post-hoc exploration of the data – which should be interpreted with caution – it suggests that new learning is sufficient to change memory, and that reactivation may not be necessary.

Discussion

In the present study, we investigated whether changing strong memories during reconsolidation was dependent on the strength of the behavioral post-reactivation manipulation, which was operationalized as the number of exposures to that manipulation. To test this, participants first acquired relatively strong memories, which were reactivated (or not) one week later followed by a new learning intervention that was either weak or strong. All participants were tested a day later. The results show quite consistent patterns over our dependent variables: recognition accuracy, hits, and false alarms. Overall, both RE+NL groups differed reliably from two to three out of four control groups; these groups controlled for reactivation or types of new learning. An unexpected, but prominent finding is that both RE+WNL and RE+SNL groups consistently performed similar to the control group that crucially controlled for weak new learning without reactivation (NRE+WNL). Moreover, RE+SNL and another important control group NRE+SNL differed only very little. Finally, we did not observe the strongest memory changes for RE+SNL compared to RE+WNL; both interventions seemed to have similar effects on memory. Thus, these results contradict what was expected based on the reconsolidation hypothesis. On the premise of reconsolidation, memory change should have been specific to the RE+SNL and RE+WNL compared to the four control groups, and a dose-dependent relationship between RE+SNL and RE+WNL should have been present.

Though we expected memory change to be specific to the groups that received interventions that entailed both explicit reactivation and new learning (i.e., RE+SNL and RE+WNL), we observed that interventions that only entailed novel learning without explicit reactivation resulted in similar memory change (i.e., NRE+SNL, NRE+WNL). Interestingly, the RE+SNL group in which the greatest memory change was expected was comparable with NRE+WNL, and in some instances also with NRE+SNL; two groups without explicit reactivation. One explanation for the new learning groups to perform similarly is that there was unintended reactivation in the former groups, which would have also made these memories sensitive to modification during reconsolidation. However, this seems unlikely because participants in NRE+SNL and NRE+WNL were tested in a different room on Day 8, and by a different experimenter (of the opposite sex). Specifically, switching spatial context limits spontaneous reactivation of the original memory and therefore does not open the reconsolidation window (Hupbach et al., 2008) Also, there were no differences between self-reports regarding reactivation, which suggest that participants in all groups rarely reactivated pictures they originally learned before they came back to the lab. Taken together, it is unlikely that unintended reactivation initiating the reconsolidation process explains the observed changes in the NRE+WNL and NRE+SNL groups.

A different process that is not grounded in memory reconsolidation may alternatively explain the effects on memory in the four groups that learned something new on day 8. This explanation is supported by interference theory (e.g., McGeoch, 1942). According to this theory, original memories were impaired because there was competition between different memory traces during the final test on Day 9. Accordingly, the groups that learned something new on Day 8 (regardless of whether there was reactivation of the original memory) seem to have experienced interference caused by competition, because memories from Day 1 and Day 8 need to be disambiguated (i.e., source identification of Day 1 or Day 8 encoding is required) in order to successfully complete the recognition test on Day 9. Alternatively, it may also be possible that source misattribution is responsible for our data patterns. Correct source identification may have been especially difficult because picture sets were largely matched on type of pictures (e.g., each set contained a picture of a dead body, a tool, etc.). If, during the recognition test, participants made their decisions in part on the basis of whether they saw the picture, instead of on which day it was seen, this could have affected test performance.

Specifically, this would have reduced hits for groups that studied novel materials on Day 8, while it increases false alarms. Groups that did not study novel materials would be largely unaffected because they are able to perform the recognition simply based on whether the picture was seen before. Interestingly, when using stringent source-monitoring paradigms, Hupbach et al. (2009) showed that memory updating during reconsolidation is not simply a matter of source confusion. Overall, interference theory and source confusion can both explain our results more parsimoniously than reconsolidation: memory impairments and errors occur, because of memory traces that are relatively similar and that compete which each other. This is exactly what happened as a result of novel learning on Day 8 in the four NL groups.

Based on the reconsolidation hypothesis, we also expected a dose-dependent relationship between RE+SNL and RE+WNL, with the largest effect for the RE+SNL group (Duvarci et al., 2008; Nader et al., 2000; Wichert et al., 2013b). However, both groups displayed comparable memory change. Perhaps memory change was hampered by how the original strong memories were created. We intended to create 'strong' memories on Day 1 by viewing and recalling pictures three times. The idea that three repetitions of the pictures induces relatively strong memories was inferred from two studies by Wichert et al. (2013a, 2013b) that showed that to create strong memories three (compared to one) reactivations and recalls are sufficient. However, the strength of the memory after the three repetitions may have been too weak for the present purposes and it may have remained susceptible to behavioral post-reactivation interventions, regardless of the intervention's strength. Still, there is reason to consider the initial memory as relatively strong. Our SNL interventions on Day 8 were identical to initial learning on Day 1 and showed that, after these interventions, participants recalled twice the number of pictures compared to participants in the WNL interventions. Also, compared to Wichert et al. (2013b), the memory participants acquired on Day 1 in the current study can at least be considered as stronger compared to the memory in their study as participants in our study viewed and recalled pictures three times, while in Wichert et al. (2013b) this happened only once.

Alternatively, it is also possible that the initial learning memory was actually too strong, leaving little room for memory change and consequently for differentiation in memory change between RE+SNL and RE+WNL. After all, items that are recalled early during initial learning are likely to be recalled during the second or third repetition, which increases their strength (Karpicke & Roediger, 2008) and potential resistance to change during reconsolidation. Items in memory that primarily remained susceptible to modifications during reconsolidation probably received little retrieval practice. The idea of resistance to change with increased memory strength is also in line with a study by Hupbach (2015), who showed that multiple testing strengthens memories for the original materials, which does indeed prevent newly learned information from intruding into the original memory. It is possible that memory strength did not only impose a limit on differential change between RE+SNL and RE+WNL, but also on whether these two groups could differ from the control groups by means of changing memories during reconsolidation, specifically from SNL and WNL. However, theoretically, it may still be possible to alter strong change-resistant memories during reconsolidation in future research. Rodent studies have shown that certain boundary conditions can be overcome given the right approach, such as delay in time (Wang et al, 2009), presenting new information (Winters et al., 2009), or prolonged reactivation (Suzuki et al., 2004). Clearly then, the question how to operationalize memories as either strong or weak in experiments like the present one warrants further study.

Our study also sheds light on the feasibility of capitalizing on this process in clinical practice. Here, reconsolidation has frequently been put forward as a new and promising avenue for treating clinical disorders (e.g., Beckers & Kindt, 2017; Besnard, 2012; Kindt & Emmerik, 2016; Schiller & Phelps, 2010; Schwabe et al., 2014). Indeed, there have been interventions during reconsolidation that successfully showed that a reactivated consolidated memory can be changed with a behavioral manipulation (Hupbach et al., 2007; 2008; 2009; James et al., 2015; Schiller et al., 2010; Schwabe & Wolf, 2009; Wichert et al., 2011; 2013a; 2013b). However, our study and other studies seem to suggest that reconsolidation is a sensitive process that does not always take place, even in well-controlled lab environments optimized to study this phenomenon (Bos, Beckers, & Kindt, 2014; Golkar, Bellander, Olsson, & Öhman, 2012; Hardwicke, Taqi, & Shanks, 2016; Kindt & Soeter, 2013; Klucken et al., 2016; Soeter & Kindt, 2011; Wichert et al., 2011;Wood et al., 2015). It is especially disconcerting that "strong" memories in the lab already show a certain resistance to change during supposed reconsolidation (as evidenced by similar changes in non-reactivation conditions). These memories are lab analogues of real life traumatic memories, but can still be considered relatively mild and controllable. Clearly, traumatic memories are strong and complex, and they differ from lab memories in a number of crucial aspects, such as the presence of strong emotions and stress hormones (see Pitman, 2011). The pathogenic nature of traumatic memories may pose a challenge to bridging the gap between the lab and the clinical practice. Of course, it is still possible that the hypothesized effects exist, but that finding these effects requires specific (currently unknown) parameters. Because of the volatility of the reconsolidation process, translation of results to clinical practice should therefore be done with the utmost caution.

Taken together, our study shows that behavioral new learning interventions of different strengths, regardless of whether the original memory was reactivated first, changed an original memory to an equal extent. This contradicts previous research on reconsolidation in two ways. First, behavioral interventions should only have been able to change original memories after these were reactivated. Second, behavioral interventions of different strengths should have had a dose-dependent relationship with memory change. These results suggest that a strong original memory may be difficult to change during the reconsolidation process, or that a strong memory can be changed, but that reactivation starting the reconsolidation process is not required. Whether memory strength may be a genuine boundary condition, or whether memory changes simply requires retroactive interference remains to be investigated in future research.

Acknowledgements

This research was supported by a TOP grant (dossier number: 40-00812-98-12030) from the Netherlands Organization for Health Research and Development (ZonMw) awarded to MAvdH and IME. IME is supported with a Vici grant (453-15-005) from the Netherlands Organization for Scientific Research (NWO). We are grateful to Ivar van de Burgt and Anouschka Welsing for their assistance in testing. We thank Lars Schwabe for providing us with their study materials, Angelos-Miltiadis Krypotos for assisting with the analyses in R, and Irene Klugkist for statistical advice.

Author contributions

KvS and IME developed the study concept. All authors contributed to the study design. KvS and YRH were responsible for data collection. KvS performed the analyses and drafted the manuscript. SCvV, IME, YRH, and MAvdH provided critical revisions.

Appendix A

IAPS numbers of pictures presented on Day 1, Day 8 and/or Day 9.

Set 1		Set 2		Set 3	
Neutral	Negative	Neutral	Negative	Neutral	Negative
7004	1200	2890	1274	9070	1201
7056	2981	5410	6570	2850	3064
7060	6313	2272	9265	5740	3051
7546	3100	7043	9042	7000	9592
7140	3130	7036	3071	7009	9220
7207	9423	7175	3225	7351	6312
2191	2703	7006	2205	7700	3530
2840	2710	7491	3500	9360	9330
2102	8485	7038	9435	7234	9910
7503	9911	5390	5973	7100	9421
7031	1270	2210	1220	2190	1275
7055	9561	7506	2811	1450	9420
7041	6315	2880	6940	5520	9405
7640	3068	7059	3110	7034	3230
7037	3140	7595	3080	7057	9415
2037	9425	7190	9410	7211	6530
7285	9040	7035	2717	7490	6250
2214	3180	5731	9428	7224	9611
8010	9120	9700	9520	7170	9921
7095	9830	5535	9320	7900	6821
7090	1301	2221	1525	2397	9560
7002	9800	8311	6210	7160	9433
7186	6560	1670	9600	7039	3400
7493	3550	7052	3150	7080	3160
7590	3170	7130	3016	7233	3300
2620	2683	7710	9570	7150	9041
5250	9253	7235	3350	7185	6212
2487	9432	5900	9810	7510	9301
2396	9926	7620	9000	7504	9181
7217	9001	7205	9290	7950	2751

Appendix B

Non-Transformed Total Performance on Day 9 in the recognition test. Means and standard deviations are presented.

Group	Day 9			
	Recognition Accuracy	Hits	False Alarms from Day 8	False Alarms from Day 9
RE+SNL	76.74 (17,14)	83.06 (14.23)	9.51 (12.53)	1.88 (4.61)
RE+WNL	74.65 (19.92)	80.42 (19.50)	9.44 (8.19)	1.25 (1.54)
RE+NNL	90.94 (9.55)	93.47 (9.29)	2.15 (3.79)	1.75 (2.54)
NRE+SNL	80.69 (14.81)	87.15 (8.38)	7.43 (7.95)	3.29 (7.32)
NRE+WNL	75.87 (12.76)	85.00 (10.63)	14.93 (10.51)	2.00 (2.21)
NRE+NNL	89.79 (9.88)	91.81 (10.04)	2.15 (3.56)	1.13 (1.87)

Note. The percentage of false alarms for never seen pictures is reported, but was not analyzed, because these false alarms are not indicative of the reconsolidation process. RE = Reactivation, NRE = No Reactivation, SNL = Strong New Learning., WNL = Weak New Learning, NNL = No New Learning.

Appendix C

Total Performance on Day 9 in the recognition test for negative pictures and neutral pictures. Means and standard deviations are presented.

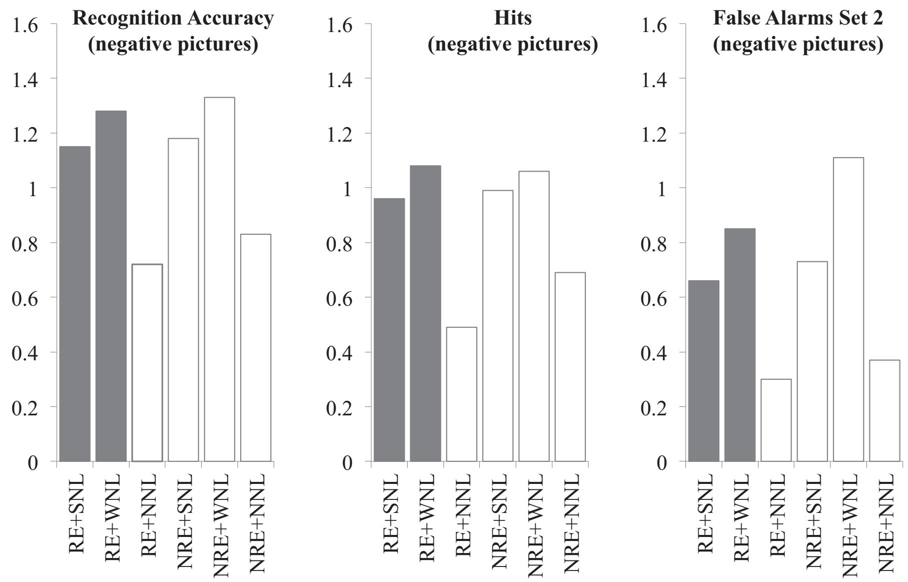
Group	Day 9					
	Negative			Neutral		
	Recognition Accuracy	Hits	False Alarms from Day 8	Recognition Accuracy	Hits	False Alarms from Day 8
RE+SNL	1.15 (0.42)	0.96 (0.48)	0.66 (0.56)	1.29 (0.44)	1.15 (0.48)	0.78 (0.53)
RE+WNL	1.28 (0.32)	1.08 (0.46)	0.85 (0.46)	1.36 (0.28)	1.19 (0.42)	0.78 (0.52)
RE+NNL	0.72 (0.49)	0.49 (0.51)	0.30 (0.43)	0.82 (0.48)	0.61 (0.58)	0.25 (0.37)
NRE+SNL	1.18 (0.31)	0.99 (0.25)	0.73 (0.53)	1.20 (0.35)	1.08 (0.41)	0.64 (0.48)
NRE+WNL	1.33 (0.23)	1.06 (0.33)	1.11 (0.28)	1.36 (0.20)	1.17 (0.21)	1.01 (0.49)
NRE+NNL	0.83 (0.44)	0.69 (0.43)	0.37 (0.44)	0.89 (0.47)	0.73 (0.56)	0.19 (0.35)

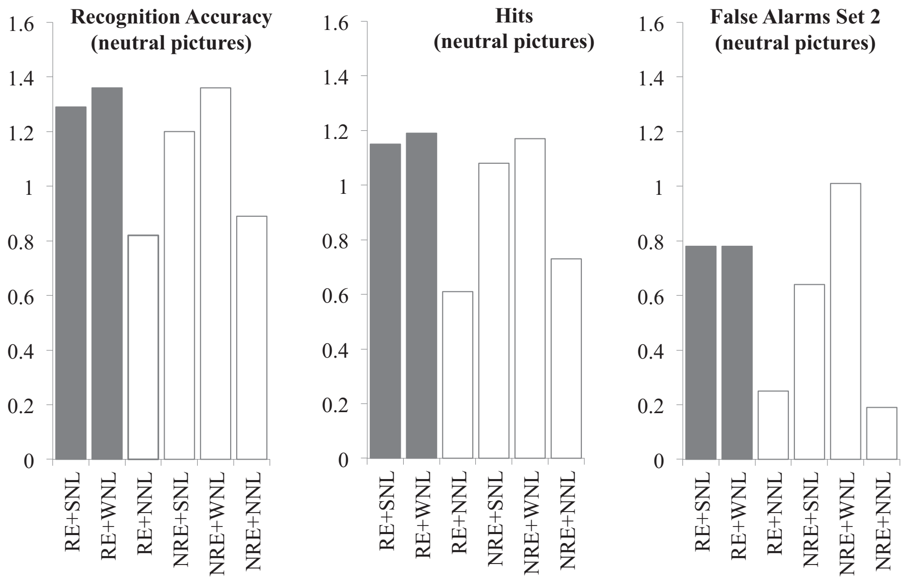
Note. Recognition accuracy and Hits are sign-reversed and log transformed scores. For these scores the interpretation of such a variable needs to be reversed as well; for instance, a higher score on recognition accuracy needs to be interpreted as *lowered* recognition accuracy. False alarms from Day 8 were log transformed only; a higher value reflects higher false alarms. Remember (%) reflects the percentage of Day 1 pictures correctly identified as being from Day 1 that were said to be remembered. RE = Reactivation, NRE = No Reactivation, SNL = Strong New Learning., WNL = Weak New Learning, NNL = No New Learning

8

Appendix D

Total Performance on Day 9 for recognition testing for negative pictures (top) and neutral pictures). The left panel depicts sign-reversed and transformed recognition accuracy scores (hits minus all false alarms). The center panel depicts the sign-reversed and transformed percentage of hits. Sign-reversed scores for recognition accuracy and hits should be reversed interpreted; a higher bar reflects a lower score. The right panel depicts the transformed percentage of false alarms from Set 2 pictures. Control groups are depicted in white. *Note*. RE = Reactivation, NRE = No Reactivation, SNL = Strong New Learning., WNL = Weak New Learning, NNL = No New Learning




Recognition Accuracy
(neutral pictures)
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
RE+SNL
RE+WNL
RE+NNL
NRE+SNL
NRE+WNL
NRE+NNL
Hits
(neutral pictures)
False Alarms Set 2
(neutral pictures)

SECTION 3

RETRIEVAL SUPPRESSION OF UNWANTED INTRUSIVE MEMORIES

CHAPTER 9

EMOTIONAL AND NON-EMOTIONAL MEMORIES ARE SUPPRESSIBLE UNDER DIRECT SUPPRESSION INSTRUCTIONS

K. VAN SCHIE
E. GERAERTS
M. C. ANDERSON

Cognition & Emotion, 27(6), 1122-1131 (2013)
doi:10.1080/02699931.2013.765387

Abstract

Research on retrieval suppression has produced varying results concerning whether negatively valenced memories are more or less suppressible than neutral memories. This variability may arise if, across studies, participants adopt different approaches to memory control. Cognitive and neurobiological research points to two mechanisms that achieve retrieval suppression: thought-substitution and direct suppression (Benoit & Anderson, 2012; Bergström, de Fockert, & Richardson-Klavehn, 2009). Using the Think/No-Think paradigm, this study examined whether participants can inhibit neutral and negatively valenced memories, using a uniform direct suppression strategy. Importantly, when strategy was controlled, negative and neutral items were comparably inhibited. Participants reported high compliance with direct suppression instructions, and success at controlling awareness predicted forgetting. These findings provide the first evidence that direct suppression can impair negatively valenced events, and suggest that variability in forgetting negative memories in prior studies is unlikely to arise from difficulty using direct suppression to control emotionally negative experiences.

Keywords: memory control, direct suppression, think/no-think paradigm

Introduction

For better or for worse, people regularly encounter reminders of things past. Objects, people, or situations can revive memories of a birthday celebration or a meeting with a friend, but can also remind us of the loss of a loved one. When negative memories intrude, they unsettle us, undermining our peace of mind, and people generally take mental action to limit the duration of such memories in awareness. Studying the mechanisms underlying this type of memory control and the limits on their operation is fundamental to understanding how people adapt the functioning of their memories in the aftermath of unpleasant life experiences.

The ability to control unwanted memories has been studied with the Think/No-Think (TNT) paradigm (Anderson & Green, 2001). The TNT paradigm investigates how suppressing retrieval in response to reminders affects a memory's retention. Participants first learn cue-target word pairs (e.g., ordeal-roach). They then perform a TNT task in which they receive trials presenting a cue from these pairs, and are asked to either recall the associated memory (i.e., think items), or to suppress its retrieval (no-think items). A third set of items (baseline items) is also studied, but is not shown in the TNT phase. Afterwards, memory for all pairs is assessed.

On the final test, a counterintuitive effect arises revealing the consequences of suppressing retrieval: no-think items are also recalled more poorly than baseline items that were neither retrieved nor suppressed. This difference, known as the *negative control effect,* shows that retrieval suppression causes more forgetting than would ordinarily occur due to the passage of time, and has been found in numerous studies (Anderson & Green, 2001; Anderson et al., 2004; Anderson, Reinholz, Kuhl, & Mayr, 2011; Paz-Alonso, Ghetti, Matlen, Anderson, & Bunge, 2009; see also Anderson & Huddleston, 2012; Levy & Anderson, 2008, for reviews). The negative control effect generalizes to non-verbal materials, such as faces (Depue, Banich, & Curran, 2006; Hanslmayr, Leipold, & Bäuml, 2010; Hanslmayr, Leipold, Pastötter, & Bäuml, 2009) and scenes (Depue et al., 2006; Depue, Curran, & Banich, 2007).

A key question, however, concerns how effectively suppression works for emotional memories. The literature contains diverging views on how emotion might influence memory suppression. On the one hand, suppressing emotional memories might be intrinsically more difficult because these memories are better encoded, consolidated and retrieved, than are non-emotional ones (Hamann, 2001; LaBar, & Cabeza, 2006; Levine & Pizarro, 2004). Memory enhancement for arousing emotional information also seems more automatic than it is for non-arousing information (Kensinger & Corkin, 2004). Moreover, the intrusive nature of traumatic memories in PTSD reinforces this view, suggesting that emotional memories are harder to suppress (Shipherd & Beck, 2005). On the other hand, the emotionality of memories could motivate people to engage in cognitive control over emotional memories. Disruption of those memory traces might be easier precisely because emotional memories are more accessible (Levy & Anderson, 2012).

Studies examining the ability to suppress retrieval of emotional memories produce evidence for both arguments. Some authors have found larger negative control effects for emotionally negative compared to neutrals or positive materials (e.g., Depue et al., 2006; Joorman, Hertel, Brozovich, & Gotlib, 2005; Lambert, Good, & Kirk, 2010), indicating people may be better able or more motivated to forget unpleasant stimuli. Others have found smaller effects for negative materials, (e.g., Marx, Marshall, & Castro, 2008; Nørby, Lange, & Larsen, 2010). Others have found no measurable effects for valence (Murray, Muscatell, & Kensinger, 2011). It thus remains unclear why the relative magnitude of negative control effects for neutral and negative memories has varied so much.

One possibility is that this variability stems from participants adopting different strategies for retrieval suppression across valence conditions and studies. One strategy that received considerable interest is thought-substitution, which often produces larger negative control effects than are typically observed without specific strategy instructions (e.g., Hertel & McDaniel, 2010; Joormann, Hertel, LeMoult, & Gotlib, 2009). For example, *Aided* participants are often provided with alternative words (thought-substitutes) to retrieve when no-think cues appear, whereas *Unaided* participants received no specific guidance. The *Aided* group often show a larger negative control effect compared to the *Unaided* group.

If thought-substitution increases the negative control effect, uncontrolled variation in this strategy may account for variability in how effectively negative memories are forgotten.By this hypothesis, forgetting emotional memories is more difficult because inhibitory control is inadequate to handle the putative intrusiveness of negative traces. Thus, unless participants resort to thought-substitution, negative control effects for negative materials will be smaller. Greater negative control effects for negative memories could arise if negative memories prompted thought-substitution more often than neutral memories, obscuring an underlying deficit in the ability to inhibit negative memories.

A second mechanism contributing to retrieval suppression is known as direct suppression (Levy & Anderson, 2008). Direct suppression is assayed by asking participants to avoid thinking of the target memory without replacing it with anything else; if the unwanted memory happens to come to mind, participants simply are asked to block it out. Using direct suppression instructions, several studies have shown impaired recall for no-think items (Benoit

& Anderson, 2012; Bergström, de Fockert, & Richardson-Klavehn, 2009; Hanslmayr, Leipold, Pastötter, & Bäuml, 2009). Importantly, both ERP and fMRI research comparing direct suppression with thought-substitution indicate that direct suppression is mediated by distinct control mechanisms, with the former suppressing neural processes that contribute to episodic retrieval (Benoit & Anderson, 2012; Bergstrom et al., 2009; Hanslmayr et al., 2010). For instance, using effective connectivity methods, Benoit and Anderson (2012) demonstrated that direct suppression engages right dorsolateral prefrontal cortex to down-regulate hippocampal activity, but thought-substitution does not; thought-substitution, by contrast, engages left ventrolateral prefrontal cortex to retrieve substitute memories, and is associated with increased retrieval-related activity in the hippocampus. Thus, these qualitatively different approaches to memory control are known to be dissociable at the neural level. Importantly, they indicate that inhibitory control can directly suppress retrieval of unwanted memories, impairing their retention. It remains unknown, however, whether direct suppression impairs negatively valenced memories, as all studies of direct suppression have used neutral items. Direct suppression may be especially ineffective in suppressing emotional content.

If inhibiting negative memories is difficult and thought-substitution is necessary to control such memories, then holding subjects' strategy constant by asking them to perform direct suppression should reveal less forgetting for negative compared to neutral memories. To examine this, we gave participants direct suppression instructions and encouraged them to use this approach, reducing variability in strategies. Additionally, we manipulated the emotional valence of the cue (neutral versus negative) and of the target (neutral versus negative). With this manipulation we could disentangle how direct suppression affects materials with different valences, and whether this effect is related to either the cue or target valence. The negative emotional valence of visible reminders may capture participants' attention, perhaps exaggerating any modulatory influence that affect may have on suppression (whether the influence is positive, or negative).

Method

Participants

Thirty-eight undergraduates (M = 20.63 years, SD = 2.03; 10 males) of the Erasmus University Rotterdam participated for course credit. Participants were excluded if they had a diagnosis of attention deficit disorder, did not have Dutch as a first language (learned prior to age 5) or if they were colour blind. Two participants were not used in analyses because they had a score higher than 3 on three post-experimental non-compliance questions for no-think instructions.

Materials and Design

The stimuli consisted of 72 weakly related Dutch cue-target word pairs (e.g., lane-meter, remove-cancer), 24 each in the think, no-think, and baseline conditions. Words and word pairs were in part selected from previous TNT studies (Anderson & Green, 2001; Anderson et al., 2004) and in part newly constructed. Care was taken to ensure that cue and target words were only relatable to each other and not to items from other pairs. First through careful independent inspection by the experimenter and the first author, and then by ensuring that use of our materials was justified by inspection of the Dutch association norms.

Of the 24 pairs in each condition, valence (negative versus neutral) of the cue and target words was manipulated, resulting in four valence groups: neutral-neutral, negative-neutral, neutral-negative, and negative-negative. Within each valence group six pairs each made up the think, no-think, and baseline conditions.

To assess valence and arousal, twenty-five students who did not participate in the experiment rated a large sample of words on valence ranging from 0 (*negative*) to 50 (*positive*) and arousal ranging from 0 (*low*) to 50 (*high*). Our 72 words were selected from this sample. Words scoring under 16.7 on valence qualified as negative; those between 16.7 and 33.3 as neutral. For each valence group, valence (v), arousal (a), and word length (wl) of cue and target is reported: cue for neutral-neutral pairs (Mv = 26.65; Ma = 15.34, Mwl = 5.44), target for neutral-neutral pairs (Mv = 27.88; Ma = 15.33, Mwl = 5.28), cue for neutral-negative pairs (Mv = 27.9; Ma = 17.33, Mwl = 5.61), target for neutral-negative pairs (Mv = 10.5; Ma = 24.41 Mwl = 7.28), cue for negative-neutral pairs (Mv = 10.98; Ma = 23.75, Mwl = 5.83), target for negative neutral pairs (Mv = 27.38; Ma = 17.64, Mwl = 5.28), cue for negative-negative pairs (Mv =12.04; Ma = 23.67, Mwl = 6.72), target for negative-negative pairs (Mv = 11.27; Ma = 23.61, Mwl = 6.28).

Within each valence group, critical pairs were counterbalanced so that each participated in every condition of the Think/No-Think task equally often.

Think/No-Think Procedure

The experiment was run with E-prime 2.0 (Psychology Software Tools, Pittsburgh, PA), using a procedure based on Anderson et al.'s (2004) experiment. During the phases, the experimenter sat behind the participant, scoring vocal responses, giving participants instructions and verbally encouraging them when necessary.

Learning Phase. Each pair appeared once individually in white font in the middle of a black screen for 5000 ms (400 ms ITI). Pseudo-randomized test-feedback cycles followed in which participants responded with the target into a microphone when a cue appeared. Cues disappeared after 3500 ms or when the participant responded. Regardless of the answer, the correct target appeared in green for 1000 ms (400 ms ITI). As is standard in the TNT task test-feedback cycles continued until a participant acquired a minimum of 50% of target answers for all conditions combined. Participants had up to seven cycles to achieve the criterion.

Think/No-Think phase. Participants were told that they would be receiving two types of trials in which a cue word from one of the pairs would appear for a short time. They were told that when the cues appeared in blue, they should immediately think of the associated target word and to keep it in mind while the cue was onscreen. When the cues appeared in yellow, however, they were asked to stop themselves from thinking of the associated word. For these no-think trials, participants received direct suppression instructions, asking them to continuously focus on the cue and suppress retrieval of the target by blocking thoughts about it, without replacing it with other thoughts. Each trial started with a fixation cross (400 ms) followed by the cue (3500 ms; 400 ms ITI). The colour representing each condition was counterbalanced over participants. The TNT phase took 40 minutes.

9

Participants performed 24 practice filler trials (12NT, 12T), followed by the experimental TNT phase. The TNT phase consisted of six blocks containing 96 think and no-think cues displayed in a pseudo-randomized order, with each think and no-think cue appearing twice; no more than three T or NT items appeared consecutively. Across all blocks, each think and no-think cue was repeated 12 times. Breaks of 30-45 seconds occurred between blocks.

Final Test Phase. All 72 pairs were tested with a "same probe" test (e.g., lane-_____). Each cue appeared once in white font in the middle of a black screen for 3500 ms (400 ITI). Participants were instructed to recall all words regardless of their colour and instruction in the TNT phase.

After the test, participants filled out a questionnaire in which they rated their success at controlling memory for each pair on a scale from 0 (*never able to avoid thinking about the target word*) to 4 (*always able to avoid thinking about the target word*). They also rated how often they used each of a collection of strategies on a scale from 0 (*never*) to 4 (*always*), and whether they intentionally were non-compliant with no-think instructions; 0 (*never*) to 4 (*very frequently*).

Results

Analysis was based only on pairs for which participants recalled the target on the final learning test (Anderson et al., 2004). This moderately high level of learning performance yielded data in every cell for every participant. Analyses of variance (ANOVAs) were conducted to examine the aftereffects of memory control in the final test. Counterbalancing condition was included as a between-subjects factor in all analyses to account for item effects, and non-significant results of this factor (or its interactions) are not reported. Instruction (think, no-think, and baseline) was analyzed as a within-subjects factor. A multivariate test (Pillai-Bartlett trace (V)), is reported when the assumption of sphericity was violated.

Learning Phase

No participant required more than three cycles to achieve the learning criterion of 50% ($M = 1.69$, $SD = 0.71$). Overall, recall on the final learning test was 75.35% ($SE = 1.80$), and was similar for words from different instruction conditions: baseline ($M = 73.61$, $SE = 2.15$), think ($M = 75.00$, $SE = 2.32$), no-think ($M = 77.43$, $SE = 1.82$), $F(2, 70) = 2.06$, $p > .05$, *partial* $\eta^2 = .06$. Recall for the valence conditions differed, $F(3, 90) = 18.41$, $p < .001$, *partial* $\eta^2 = .38$, with $M = 77.62$ ($SE = 3.05$), 78.70 ($SE = 1.99$), 80.25 ($SE = 1.82$), and 62.81 ($SE = 2.88$) for neutral-neutral, neutral-negative, negative-neutral, and negative-negative, respectively. Paired sample t-tests (against Bonferroni-corrected α of .0083) revealed that only the negative-negative condition differed from other groups, $ts(35) > 5.2$, $ps < .001$). For these reasons, we conditioned our analysis of participants' final recall on having correctly learned items on the final learning test.

Overall Effects of Memory Control on the Final Test

Instructions (baseline, no-think, think) affected recall, $V = 0.25$, $F(2, 29)$, $p < .05$, *partial* $\eta^2 = .25$. The effect of instruction was followed up with two contrasts, comparing baseline to either the think or the no-think condition. Recall was poorer for words in the no-think condition ($M = 0.82$, $SE = .04$), compared to baseline ($M = .91$, $SE = .02$), $F(1, 33) = 6.86$, $p < .05$, *partial* $\eta^2 = .17$. Conversely, recall for words in the think condition ($M = 0.94$, $SE = .01$) was better than baseline, $F(1, 33) = 5.80$, $p < .05$, *partial* $\eta^2 = .14$. Thus, the present findings replicate previous reports (Benoit & Anderson, 2012; Bergström et al., 2009; Hanslmayr et al., 2009, 2010) showing that direct suppression can impair recall of unwanted memories.

Effects of Cue Valence

To see whether cue valence influenced suppression success, cue valence (neutral, negative) was included as a within-subject factor. Here again, recall varied with instruction, $V = 0.26$, $F(2, 32)$, $p < .01$, *partial* $\eta^2 = .26$. Recall was lower for no-think words compared to baseline words, $F(1, 33) = 4.96$, $p < .05$, *partial* $\eta^2 = .13$, whereas recall was higher for think words than for baseline words, $F(1, 33) = 6.94$, $p < .05$, *partial* $\eta^2 = .17$. Overall recall did not vary with cue valence, $F<1$. Importantly, we observed no Cue Valence × Instruction interaction, $F < 1$. Consequently, the contrast comparing baseline and no-think performance did not interact with cue valence, $F < 1$, nor did the contrast of baseline and think performance, $F<1$ (See Figure 1, top row, for the effects of instruction as a function of cue valence). Thus, when using a direct suppression strategy, people's ability to suppress unwanted memories did not vary for memories cued by negative or neutral reminders.

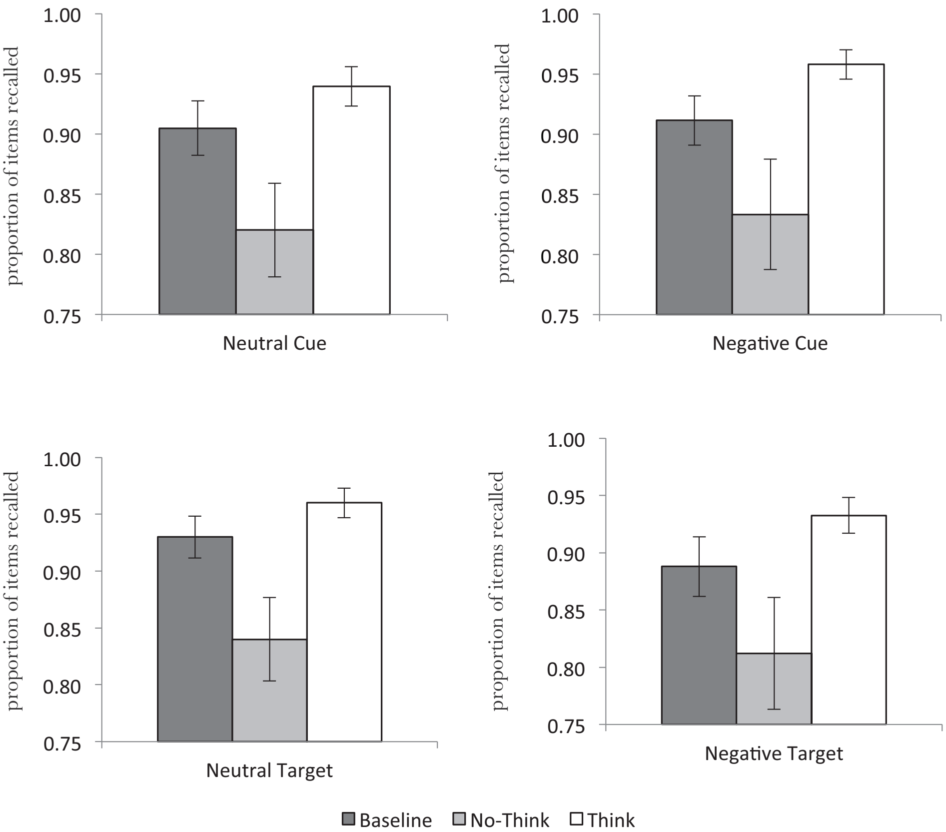


Figure 1. Mean proportion of recall for baseline (dark grey), no-think (light grey) and think (white) on the final test, displayed for valence of cue and target. Error bars represent standard errors of the mean.

Effects of Target Valence

To see whether the valence of the to-be-suppressed memories affected suppression success, target valence (neutral, negative) was added as a within-subjects factor. Again, recall varied with instruction, $V = .26$, $F(2, 32) = 5.65$, $p < .01$, *partial* $\eta^2 = .26$. Contrasts revealed that baseline words were better recalled than no-think words, $F(1, 33) = 4.96$, $p < .05$, *partial* $\eta^2 = .13$, but were worse recalled than think words, $F(1, 33) = 6.94$, $p < .05$, *partial* $\eta^2 = .17$. Overall recall varied according to target valence, $F(1, 33) = 4.95$, $p < .05$, *partial* $\eta^2 = .13$, with negatively valenced words being recalled less. Importantly, there was no interaction between Instruction × Target Valence, $F<1$ (see Figure 1, lower row, for the effects of instruction as a function of target valence). Thus, when using a direct suppression strategy, negative memories were neither more nor less suppressible than were neutral memories. These findings indicate that negatively valenced memories are capable of being forgotten by a direct suppression mechanism, and do not have to be forgotten by thought-substitution.

Questionnaire Analysis

Strategies for memory control. Most participants (88.89%) frequently employed strategies during no-think trials that are consistent with direct suppression instructions, such as staring intently at the hint word, repeating the hint word silently, or letting their mind go blank in response to the hint word. Thought-substitution strategies in which people used the cue to generate an alternative word, thought or sound, were infrequently used (13.89%). Thus, subjects largely followed our instructions to use direct suppression.

Self-reported success of memory control. Participants rated their success at controlling awareness for each of the suppression cues. We combined their ratings across items according to our distinctions: success at control for neutral versus negative cues, and success at control for neutral versus negative targets. Ratings for pairs were considered only if participants recalled the target word on the final trial in the learning phase, ensuring that differences in learning for different pair types did not contribute to success ratings.Participants reported greater success of not thinking of the memories associated to negative cues ($M = 2.46$, $SE = .11$) compared to those associated to neutral cues ($M = 2.30$, $SE = .12$), $F(1, 33) = 5.94$, $p < .05$, *partial* $\eta^2 = .15$. For target valence, participants reported being more successful at not thinking of negative targets ($M = 2.44$, $SE = .12$), compared to neutral targets ($M = 2.32$, $SE = .11$), although this relationship was only marginal, $F(1, 33) = 3.41$, $p = .073$, *partial* $\eta^2 = .09$. These findings suggest, intriguingly, that, contrary to the notion that negative memories are more difficult to suppress, our participants experienced greater success at controlling negative memories.

Success ratings and memory inhibition. Interestingly, we found a relation between inhibition scores (baseline recall − no-think recall) and self-reported success at memory control. There was a relationship between success at controlling awareness and inhibition scores for negative cues, $r = .42$, $p < .01$, and negative targets, $r = .37$, $p < .05$. This relationship was not present for neutral cues or targets, $p > .05$. Thus, at least for negative memories, forgetting resulting from direct suppression was related to the participants' experience of success in controlling conscious awareness.

Discussion

The present study examined whether unwanted memories could be forgotten with direct suppression, and compared the efficacy of this strategy for both neutral and negatively valenced memories. Consistent with recent work (Benoit & Anderson, 2012; Bergstrom et al., 2009; Hanslmayr et al., 2010), negative control effects were indeed observed, illustrating how memory control can be accomplished without thought-substitution, via inhibitory processes that expel unwanted memories from awareness. This conclusion is supported by participants' high compliance with our instructions, and by correlations between the negative control effect and self-reports of control over awareness. Thus, thought-substitution need not be engaged to forget unwanted memories.

More importantly, however, the current study establishes, for the first time, that emotionally negative memories can be forgotten by direct suppression. All previous studies demonstrating negative control effects for emotional memories, used no particular suppression instructions (e.g., Lambert et al., 2010; Marx et al., 2008; Murray et al., 2011), and so it cannot be known what processes caused these effects. Moreover, all studies of direct suppression have used neutral materials, leaving it unclear whether negative memories can be controlled in this fashion. These findings do not support the idea that negatively valenced materials necessarily undermine peoples' ability to control memory. Nor do they clearly address whether they might enhance memory suppression. However, the simple verbal materials in our study may not capture the emotional qualities people experience in unpleasant life events. Therefore, more work is necessary to determine whether these effects generalize to materials that are more naturalistic and autobiographical that might cause discomfort, shame, embarrassment, or sadness (see, Noreen & MacLeod, 2013).

Our findings speak to the potential causes of variability in the effect of valence on retrieval suppression (see Anderson & Huddleston, 2012). Of interest was whether variations in the spontaneous use of thought-substitution in prior studies masked a deficit in inhibiting negative memories. Crucially, when we controlled strategy with direct suppression instructions, no inhibition deficit was evident; neutral and negative memories were comparably forgotten. Because the instructions used here are known to trigger top-down inhibitory modulation of hippocampal activity by dorsolateral prefrontal cortex (Benoit & Anderson, 2012), our manipulation provided a theoretically refined test of the hypothesized difficulty in inhibiting negative memories. Indeed, to our surprise, participants, judged pairs including negative cues or targets to have been easier to keep out of awareness than pairs with neutral cues or targets. Interestingly, success ratings predicted forgetting for negative, but not neutral memories, suggesting that controlling awareness of negative memories may be more reliant on inhibition, consistent with the possibility that they are more intrusive. Even with this possibility, however, our results as a whole, do not suggest that inhibition is less able to contend with negative materials. Nevertheless, uncontrolled variations in thought-substitution may still explain variability in prior studies as to whether negative or neutral memories are more forgettable. Our findings simply show that this variability in thought-substitution, if it exists, does not mask a deficit in inhibiting negative memories.

Our experiment also manipulated both cue and target valence to examine which influences memory suppression. The majority of studies examining the negative control effect either show enhanced forgetting for neutral (Marx et al., 2008; Nørby et al., 2010) or for emotional materials (Depue et al., 2006; Joorman et al., 2005; Lambert et al., 2010). We observed neither effect, regardless of whether we considered cue or target valence. Nevertheless, these findings offer insights into several studies that show preferential suppression of emotional or non-emotional material. For instance, our study is the only one comparable to Lambert et al. (2010), who manipulated cue rather than target valence. Lambert et al. (2010) reported enhanced suppression for negative pairs, yet solely for neutral targets associated with a negative cue compared to neutral items paired with a positive cue. Because Lambert et al. (2010) lacked a condition with neutral cues, it was unclear whether these results reflected good forgetting of negative experiences, or difficulty suppressing positive items. Our finding of comparable forgetting for negative and neutral memories is compatible with theirs, but suggests that their finding may reflect diminished suppression of positive events and not enhanced suppression of negative ones.

Nørby et al. (2010) found greater forgetting for neutral than negative pairs, and suggested that this difference arose from the random intermixing of negative and neutral pairs. By this view, the difficulty in predicting emotional qualities of an upcoming item enhances the item's emotional salience, revealing how difficult it is to suppress negative memories. Studies in which predictability of emotional material is high, for instance through employing different blocks for emotional material (e.g., Depue et al., 2006, 2007), will show similar forgetting for negative and neutral items by this view, because the negative items are, in effect, less negative. Although we did not seek to address this issue, our findings suggest that random intermixing of valence types may not have the effects hypothesized by Nørby et al. (2010). In our study, negative and neutral items were intermixed and we nevertheless found emotional materials were comparably suppressible to non-emotional materials. Indeed, participants judged negative items to be easier to keep out of awareness. Clearly, our results are not in line with their account. Their finding of diminished negative control effects for negative materials thus remains unexplained.

One factor that may also contribute to variability in how suppression affects retention of negative and neutral memories concerns the semantic interrelatedness of the materials being suppressed. Because words of negative valence stem from a small set of categories (e.g., death, fear and violence) repeated retrieval of stimuli in the think condition (e.g., swamp-death) may affect later recall in the no-think condition (e.g., knife-murder) when there is high interrelatedness. Goodmon and Anderson (2011) have shown that high semantic relatedness in another inhibition phenomenon--retrieval induced forgetting--abolished forgetting effects. A similar principle may apply to retrieval suppression. The presence of uncontrolled inter-pair relationships may account for why Nørby et al. (2009) failed to observe negative control effects whereas we did, particularly given our efforts to control such relationships.

In conclusion, our results establish that negative memories may be forgotten via direct suppression and that thought-substitution is not essential (Benoit & Anderson, 2012; Bergström et al., 2009; Hanslmayr et al, 2009). These findings imply that reports of forgetting of both negative and neutral memories in the TNT paradigm are unlikely to be due to deficient inhibitory control for negative memories that is compensated for by thought-substitution. Though we found that negative memories were inhibited, it remains possible that variability in the spontaneous use of thought-substitution occurs in other studies, potentially underpinning differences in the size of negative control effects across valence conditions. For these reasons, we recommend holding participants' strategic approach to the task constant to better understand the effects of emotion on the ability to suppress unwanted memories.

Acknowledgements

We thank Eveline Stoker for her assistance in testing

Author contributions

KvS developed the study concept; MCA and EG contributed to the study design. KvS was responsible for data collection and analysis. KvS drafted the manuscript and MCA and EG provided critical revisions.

CHAPTER 10

SUCCESSFULLY CONTROLLING INTRUSIVE MEMORIES IS HARDER WHEN CONTROL MUST BE SUSTAINED

K. VAN SCHIE

M. C. ANDERSON

Memory, 25(9), 1201-1216 (2017)

doi:10.1080/09658211.2017.1282518

Abstract

After unpleasant events, people often experience intrusive memories that undermine their peace of mind. In response, they often suppress these unwanted memories from awareness. Such efforts may fail, however, when inhibitory control demands are high due to the need to sustain control, or when fatigue compromises inhibitory capacity. Here we examined how sustained inhibitory demand affected intrusive memories in the Think/No-Think paradigm. To isolate intrusions, participants reported, trial-by-trial, whether their preceding attempt to suppress retrieval had triggered retrieval of the memory they intended to suppress. Such counter-intentional retrievals provide a laboratory model of the sort of involuntary retrieval that may underlie intrusive memories. Using this method, we found that longer duration trials increased the probability of an intrusion. Moreover, on later No-Think trials, control over intrusions suddenly declined, with longer trial durations triggering more relapses of items that had been previously been purged. Thus, the challenges of controlling retrieval appear to cause a decline in control over time, due to a change in state, such as fatigue. These findings raise the possibility that characteristics often true of people with psychiatric disorders – such as compromised sleep, and increased demand on control – may contribute to difficulties in suppressing intrusive memories.

Keywords: inhibitory control, intrusive memories, involuntary retrieval, retrieval suppression, memory inhibition

Introduction

Memories of unpleasant life events sometimes intrude into awareness. When this happens, people often seek to limit awareness of the unwelcome reminding by stopping the retrieval process. Evidence indicates that when people consistently suppress retrieval in this manner, suppressed items grow increasingly difficult to recall on later occasions (for reviews, see Anderson & Hanslmayr, 2014, Anderson & Huddleston, 2012). Suppression not only reduces retention on measures of voluntary retrieval, but also on involuntary forms of retrieval such as the tendency for memories to intrude into awareness in response to reminders (Benoit, Hulbert, Huddleston, & Anderson, 2014; Levy & Anderson, 2012), free association (Hertel, Large, Stuck, & Levy, 2012), perceptual priming (Gagnepain, Henson, & Anderson, 2014; Kim & Yi, 2013), and implicit association tests (Hu, Bergström, Bodenhausen, & Rosenfeld, 2015). These findings suggest that people can recruit cognitive control mechanisms to adaptively regulate intrusive memories.

Whether this type of memory adaptation is possible, however, may hinge on other conditions that could compromise people's efforts to implement suppression, such as the total time over which the suppression of an unwanted thought must be sustained. Indeed, most people are familiar with how difficult it can be to sustain attention for long periods of time and to avoid distraction, particularly when the potential for distraction is significant. In the present article, we examined these putative difficulties in the context of memory control. In particular, we examined whether counter-intentional retrievals of the sort that could underlie naturally occurring intrusive memories grow more likely both with sustained challenge to inhibitory control mechanisms, and with fatigue.

Involuntary Memories and Their Control

The current investigation builds on two distinct, but complementary bodies of research concerning involuntary memory on the one hand, and memory control, on the other. Research on involuntary memory focuses on processes that trigger memories to come to mind automatically and reflexively, and whether they differ from intentional retrieval. Involuntary memories have usually been studied with diaries or questionnaires, in which participants write down how and when autobiographical memories were cued and what their content was (Berntsen, 1996; Mace, 2004; Bradley, Moulin, & Kvavilashvili, 2013, Kvavilashvili & Mandler, 2004). Recently, involuntary memories have been studied in controlled environments (Clark, Mackey, & Holmes, 2013). In one paradigm, participants perform an undemanding vigilance task, while they need to ignore short cue phrases (e.g., 'crossing the road') on a screen (Schlagman & Kvavilashvili, 2008). Whenever an involuntary autobiographical memory came to mind, participants paused the vigilance task and reported their memory and what triggered it. After completing the vigilance task participants reported other characteristics of the involuntary memories, such as emotional valence and whether that memory was general or specific. One week later, participants *voluntarily* recalled autobiographical memories after seeing cue phrases. In two experiments with this paradigm, Schlagman and Kvavilashvili (2008) showed that involuntary memories were retrieved almost twice as quickly as voluntary memories and that these memories were more likely to be specific. This suggests that involuntary memories spring to mind quickly and without effort. However, how people control involuntary memories when they arise has remained largely unstudied in this research (e.g., for exceptions, see Brewin & Smart, 2005; Geraerts, Merckelbach, Jelicic, & Smeets, 2006).

Complementing this work is a largely separate literature examining how people voluntarily control memory. Research on memory control tries to isolate mechanisms by which people might prevent or interrupt the reflexive retrieval of unwelcome memories when confronted with reminders, as well as the consequences of engaging those mechanisms for the retention of suppressed traces. This form of memory control is often studied with the Think/No-Think (TNT) paradigm (Anderson & Green, 2001), which examines people's ability to suppress the episodic retrieval process to control intrusive memories. In the TNT paradigm, people are repeatedly prompted with cues to previously studied associates and are asked to either retrieve the memory (hereinafter, Think trials), or to instead stop its retrieval and exclude the unwanted memory from awareness (hereinafter, No-Think trials). Participants study a third set of items that are neither recalled nor suppressed (hereinafter, baseline items). Afterwards, memory for all pairs is tested by an unexpected memory test to determine how retrieval suppression influenced later retention of the excluded memories.

A large body of research shows that suppressing retrieval impairs later recall of No-Think items compared to baseline items, an effect known as suppression-induced forgetting (Anderson & Green, 2001; Anderson & Hanslmayr, 2014; Anderson & Huddleston, 2012; Anderson et al., 2004). Suppression-induced forgetting arises with many types of stimuli, including word pairs, face–scene pairs (Depue, Curran, & Banich, 2007), face–word pairs (Hanslmayr, Leipold, Pastötter, & Bäuml. 2009), word–object pairs (Kim & Yi, 2013; Gagnepain et al., 2014), and pairs comprising words and nonsense shapes (Hart & Schooler, 2012). Suppression-induced forgetting has even been observed with autobiographical experiences especially for event details (Hu et al., 2015; Noreen & MacLeod, 2013; 2014; Stephens, Braid, & Hertel, 2013). Importantly, suppression-induced forgetting occurs for emotional materials, including negatively valenced words and scenes (e.g., Catarino, Küpper, Werner-Seidler, Dalgleish, & Anderson, 2015; Chen et al., 2012; Kim, Oh, Kim, Sim, & Lee. 2013; Depue et al., 2007; Joormann, Hertel, Brozovich, & Gotlib, 2005; Hertel & McDaniel, 2010; Küpper, Benoit, Dalgleish, & Anderson, 2014; Lambert, Good, & Kirk, 2010; LeMoult & Hertel, 2010; Marzi, Regina, & Righi, 2014; Murray, Muscatel, & Kensinger, 2011; van Schie, Geraerts, & Anderson, 2013; Murray, Anderson, & Kensinger, 2015). Evidence suggests that suppression-induced forgetting can be produced in two ways—either as an aftereffect of processes that directly suppress retrieval, or by processes that promote self-distraction, such as thought substitution (e.g., Benoit & Anderson, 2012; Bergström, de Fockert, & Richardson-Klavehn, 2009b; Hertel & Calcaterra, 2005; Hertel & McDaniel, 2010; Küpper et al., 2014). Overall, this body of work on suppression-induced forgetting indicates that suppressing unwanted thoughts about past events can be successful, in contrast to findings observed with Wegner's thought suppression procedure (Wenzlaff & Wegner, 2000), with the latter procedure suggesting that thought suppression can be counterproductive (a point to which we return in the Discussion).

Research on retrieval suppression assumes that presenting a cue triggers an automatic retrieval process that elicits associated memories, unless effort is made to proactively prevent or reactively interrupt this process. In essence, this work assumes the form of automatic retrieval process studied in research on involuntary memories, without which there would be no need for control. Moreover, because research on retrieval suppression focuses on how people suppress *unwanted* involuntary retrievals, its mission is to address intrusive memories per se, a subclass of involuntary memories that are both unwanted and perseverative (Kvavilashvili, 2014). Although the assumption of involuntary retrieval in the Think/No-Think task is plausible, evidence for it has, until recently, been indirect. Next we discuss direct evidence for intrusive retrievals, and the control processes needed to counter them.

Measuring Intrusions and Their Demands on Inhibitory Control

To examine the assumption that involuntary retrievals arise during retrieval suppression, Levy and Anderson introduced a trial-by-trial judgment task during the TNT phase of the TNT paradigm, asking participants' to report their phenomenal experience of whether the unwanted memory intruded into awareness (Benoit et al., 2014; Levy & Anderson, 2012; Hellerstedt, Johansson, & Anderson, 2016). This method was modeled after the trial-by-trial introspection method developed in research on selective attention (Sergent & Dehaene, 2004; Sergent Baillet, & Dehaene, 2005; Corallo et al., 2008). After every Think and No-Think trial, participants judged the extent to which the target came to mind by pressing one of three buttons associated with responses 'never', 'briefly', and 'often'. Using this procedure, Levy and Anderson (2012) found that involuntary retrievals of the unwanted memory occurred often; up to 60% of the time on the first suppression attempt. With repeated suppressions, however, participants reduced these retrievals substantially (on average, down to 30%). Levy and Anderson also found that a steeper reduction in involuntary retrievals over repeated suppressions (i.e., a steeper negative slope) predicted greater suppression-induced forgetting on the final test, suggesting that controlling involuntary retrieval engages processes that trigger forgetting (see also Hellerstedt, Johansson, & Anderson, 2016).

The foregoing findings suggest that online trial-by-trial reports may provide a useful new method to isolate automatic retrieval processes of key interest in research on involuntary retrieval (Berntsen, 2010). Indeed, the tendency for memories to enter awareness in direct opposition of participants' intensive efforts to prevent a retrieval from occurring provides a conceptually precise operational definition of involuntary retrieval, given that such retrievals are counter-intentional, a hallmark feature of automaticity. Unlike most diary studies of involuntary retrieval, which only establish a lack of clear intention to retrieve a memory (warranting the phrase "non-intentional"), counter-intentional retrievals justify the use of the phrase "involuntary". Counter-intentional retrievals may provide an informative laboratory model of the retrieval processes that underlie intrusive memories, which are involuntary retrievals that are both unwanted and potentially disruptive (Kvavilashvili, 2014). In line with their counter-intentional nature, we therefore refer to the involuntary retrievals reported by our participants as "intrusions" (Levy & Anderson, 2012) that may be relevant to clinical reports of intrusive memories (a point to which we will return in the Discussion).

Several studies have used this intrusion reporting procedure to study the cognitive control demands present during intrusive memories, and the brain systems that are engaged. In general, retrieval suppression increases activation in the (right) dorsolateral prefrontal cortex (hereinafter, DLPFC; Anderson et al., 2004; Benoit & Anderson, 2012; Depue et al., 2007; Depue, Banich, Burgess, Willcutt, & Ruzic, 2010; Gagnepain et al., 2014; Levy & Anderson, 2012), an area linked to inhibitory control over motor responses (see Levy & Wagner, 2011 for a meta-analysis). Attempting to stop retrieval in this manner reduces hippocampal activity, a finding often taken to reflect inhibitory processes that interrupt retrieval activity that might otherwise allow an intrusion to occur. Critically, Levy & Anderson (2012) found that during No-Think trials on which an intrusion was experienced, hippocampal down-regulation was especially pronounced, and this retrieval-related down-regulation predicted later forgetting far better (r = .7) than did hippocampal activity during

No-Think trials when participants did not experience an intrusion (*r* = -.07). Benoit et al. (2014) further found that although suppression generally engaged right DLPFC, trials with intrusions triggered elevated activity compared to trials without intrusions, consistent with the possibility that cognitive control is up-regulated. Recently, Hellerstedt et al. (2016) found that during retrieval suppression, No-Think trials that are accompanied by intrusions show evidence that the unwanted item is briefly retrieved into working memory and then is rapidly excluded (i.e., ERP indices of working memory appear and then are quickly truncated); in contrast No-Think trials that are not accompanied by intrusions do not show ERP evidence of the unwanted item entering working memory. Together, these findings demonstrate that intrusion reports can isolate involuntary retrieval in the TNT procedure, and that these intrusions can be linked to objective neural indices of elevated cognitive control.

If suppressing intrusions requires elevated cognitive control, this process should be resource demanding. This raises the possibility that diminished cognitive control resources may trigger difficulty dealing with intrusions. Consistent with this possibility, research on individual differences in inhibitory control ability supports the idea that compromised cognitive control leads to difficulties in suppressing intrusive memories. For example, individual differences in stop-signal reaction time (i.e., a measure of how quickly a person can inhibit a motor action, indexing inhibitory control) predict the ability to successfully forget emotionally negative pictures (Depue et al., 2010). Moreover, participants with major depressive disorder show reduced suppression-induced forgetting compared to control participants (Joormann, Hertel, Lemoult, & Gotlib, 2009), as do participants with anxiety disorder, post-traumatic stress disorder, and attention deficit disorder (Catarino et al., 2015; Depue et al. 2010; Marzi et al., 2014). Given this evidence, task or state-dependent strains on cognitive control might also be associated with increased incidence of intrusive memories. We discuss these possibilities next.

Task and State Dependent Challenges to Controlling Memory Intrusions

Task conditions and transitory states can place a burden on cognitive control and may influence how effectively retrieval suppression can be engaged to control intrusive memories. For instance, the effort required to control episodic retrieval may be especially pronounced when attention must be sustained on a reminder for a longer time. Sustained attention to a reminder provides more time for the cue to drive retrieval of associated memories, increasing the risk of involuntary retrieval. Moreover, sustained attention to a reminder would require continuous vigilance and engagement of inhibitory control mechanisms to counter the tendency for cues to elicit a memory. Thus, the need to sustain attention to reminders is a task state variable that arguably places a heavier burden on inhibitory control. Consistent with this possibility, Lee, Lee, and Tsai (2007) observed significantly less suppression-induced forgetting on long-duration No-Think trials (5s) than on short-duration trials (3s). Lee et al. (2007) argued that sustaining inhibition over longer durations increases the chances of cognitive control failure and, consequently, intrusions. With more intrusions, they argued, later recall could be enhanced, owing to reinforced encoding of intruding traces. However, Lee et al. had no way of measuring the frequency of intrusions to verify their assumption of increased inhibitory demand. In the present study, we therefore used Lee et al.'s manipulation of trial duration, which we adjusted to include short (2.5s) and long (5s) durations for both Think and No-Think trials, and combined this with the intrusion report method. Based on Lee et al.'s claim, we hypothesized that longer trial durations would trigger greater demands on inhibitory control, yielding a higher probability of intrusions on long trials compared to short trials. This greater intrusion frequency may arise from increased input to retrieval stemming from sustained viewing of the cue, or instead from lapses in the task set governing inhibitory control, due to fatigue, either of which would indicate greater demands on control.

To test this hypothesis, we used the intrusion report method during the TNT phase to determine whether participants report significantly more intrusions during longer trials than during shorter ones. If sustaining control over time is demanding, we should also find indications of decreased efficacy of control over a longer time scale spanning the entire 35-minute duration of the TNT phase. This putative waning efficacy may be detectable in the patterns of intrusions over the eight suppression attempts on an item. Prior work has established a monotonic decline in reported intrusions over repeated suppression attempts during the TNT phase, possibly reflecting the tendency for prior inhibitions to accumulate in their impact on suppressed traces (Levy & Anderson, 2012; Benoit et al., 2014; Hellerstedt et al., 2016). This accumulation of inhibition may mask, however, a waning efficacy of the control process owing to fatigue. To test this possibility, we examined fine-grained item-by-item, transition patterns in memory control success that arise across consecutive suppression attempts on the same item. For example, by looking at the intrusion histories of individual items we can examine how well a participant's efforts to suppress retrieval of *a particular item* on one trial then carry forward to the next trial with that same item. This enables us to look at the conditions that lead to successful suppression (i.e., when an intrusion on trial N is followed by successful control on trial N+1 with that item, hereinafter referred to as *new successes*), and also conditions that increase the chances of memory control *relapses* (i.e., when successful control on trial N is followed later by failed control on trial N+1 with that particular item). If sustaining control over a much longer time scale leads to fatigue, two patterns should emerge. First, although new successes should increase over early repetitions, they should, as fatigue sets in, ultimately decline, indicating a reduced ability to down-regulate intrusions from one repetition to the next. Second, an initial period in which participants can reduce control relapses over repetitions should be followed by an increase in relapses towards the end of the TNT phase, as participants grow fatigued. This pattern would reflect diminishing ability to suppress intrusions in a way that has a lasting impact on later trials. Finally, we would expect relapses to be generally more likely during longer duration trials, owing to the greater demand those trials place on inhibitory control.

Method

Participants

Participants were recruited from Erasmus University Rotterdam undergraduate psychology classes via the Erasmus Psychology Study Pool. To be eligible for the study, participants had to be between 18 and 35 years of age and fluent in Dutch. Additionally, they reported never to have been diagnosed with ADD or ADHD or to be color blind. Forty-four undergraduates participated for course credit. Data from three participants were excluded because of scoring errors made by the experimenter during the experimental procedure; one participant was excluded because of failing to reach learning criterion, leaving a final sample of 40 (M = 20.28 years, SD = 1.87; 25 females and 15 males).

Materials and Design

The stimuli consisted of 60 randomly combined neutral Dutch cue-target word pairs (e.g., BARREL – NUN), which were equally divided over five word groups. The cues and targets were drawn from three previous studies (Anderson & Green, 2001; Anderson et al., 2004; van Schie et al., 2013), and were also, in part, newly constructed. For newly constructed pairs, we selected emotionally neutral words from a large dataset of Dutch words by Moors et al. (2013). We used relatively short noun words with a maximum length of 10 letters and 3 syllables. Selected words typically are concrete noun words (e.g., Stomach, Desk). For as far as possible we tried to select words that did not have a strong association with other words in the stimulus set. Such an association was limited to the targets and independent probes only. For each pair's target we selected an associated word from the association database from the University of Leuven (www.kuleuven.be/semlab) to serve as an independent probe (see Table 1 in the Appendix for all pairs' cue, target, and independent probes). These independent probes provide secondary cues for the target words (e.g. Cloister N____ for Nun). These cues enable us, on the final test, to test all targets twice, once with the studied cue, once with an independent cue. Though not crucial to the purpose of the current study, testing targets in two ways provided converging evidence for the generality of suppression-induced forgetting across cues, supporting the involvement of inhibition in controlling retrieval (Anderson & Green, 2001). Importantly, there was no evidence for the five experimental word groups to differ statistically on ratings (taken from Moors et al., 2013) of valence, arousal, power, and age of acquisition for cue, target, and independent probe; nor did they differ on word length, word frequency, and association strength between target and independent probe. Eighteen additional cue-target pairs served as fillers. Filler pairs were used to train participants on the Think/No-Think task (see description of the Think/No-Think phase in The Think/No-Think Procedure below).

Word groups were rotated through conditions (Baseline, Think short, Think long, No-Think short, and No-Think long), so that each word group was presented equally often in each condition.

Think/No-Think Procedure

The experiment was run with E-prime 2.0 (Psychology Software Tools, Pittsburgh, PA). During the procedure, the experimenter was present to provide instructions for each phase and verbal encouragement when necessary. Vocal responses were scored out of the participant's sight.

General Instructions. At the outset of the experiment, participants were told that they were about to participate in an experiment on attention, and that their ability to ignore distraction would be assessed. They were told that they would learn pairs of words to be used in the attention test, and that they would need to ignore associations in memory. No reference was made to a final memory test for the words at any point in the procedure, to ensure that participants took the retrieval suppression task seriously and did not covertly rehearse suppression words during the Think/No-Think phase.

Learning Phase. The total set of 60 critical pairs was learned in three subsets of 20. Each subset of 20 critical items had an equal representation of items from each counterbalancing set. In the first set, 26 pairs (20 critical plus 6 filler) were presented one by one in white font on a black background in the middle of a screen for 4000ms with a 400ms intertrial interval (ITI). Pseudo-randomized test-feedback cycles followed in which participants were instructed to verbally recall each target when presented with its cue. Cues disappeared after 3500ms or upon verbal response. The correct target was displayed in blue for 1000ms followed by a 400ms ITI regardless of the participant's answer. If at least 50% of the experimental target words were recalled correctly, participants progressed immediately to the second subset of words; if not, the test was repeated once. Regardless of the percentage correct on set repetition, the participant continued to the second subset of 26 words, for which the same procedure ensued. After this, they continued to the third and final subset. The phase concluded with an integrated test (without any feedback) for pairs from all three sets. Participants repeated this phase if they did not reach the 50% cut-off. If participants did not reach the cut-off on their second test, they were excluded from further participation.

Think/No-Think Phase. On each trial, a fixation cross appeared for 400ms followed by a cue presented for either 2500ms for short trials or 5000ms for long trials. For Think cues, presented in green, participants were asked to retrieve its target silently; for No-Think trials, presented in red, they received direct suppression instructions, asking them to suppress retrieval of the target *without* substituting something else for the target. Cue words from *Baseline items*, though learned in the study phase, did not appear in this phase, and thus provided a baseline estimate of memory for pairs, given that neither suppression nor retrieval took place during the Think/No-Think phase. After each cue word, a rating scale appeared on the screen and participants had 1500ms to report the extent to which the target entered awareness during that trial; the scale included values of 'never', 'briefly', and 'often', and participants indicated their response by pressing a corresponding button with their dominant hand. Trials were separated by a 400ms ITI.

The Think/No-Think phase consisted of four blocks separated by 45-second breaks. Each block contained 96 cues, representing 12 cues in each Instruction × Duration combination, each repeated twice. No more than three cues with the same instruction or duration were displayed consecutively. Across all blocks, each cue was repeated eight times. Before the critical Think/No-Think phase, participants performed two practice blocks with 24 filler trials:12 No-Think and 12 Think trials. The first practice block was conducted without the rating scale for either Think or No-think items; the second practice block added the ratings for all trials.

Final Test Phase. After the Think/No-Think phase, participants' memory was assessed unexpectedly using two types of tests, the order of which was counterbalanced across participants. In the same probe test, participants were presented with the pair's cue (e.g. BARREL – _____) and in the independent probe test, with a word associated to the target along with a letter stem (CLOISTER – N____). All items were assessed in the same probe and in the independent probe test. Participants were given up to 10s to recall the target. Trials were separated by a 400ms ITI. A short practice test on fillers preceded these tests.

Post-Experimental Questionnaire. After the test, participants filled out a post-experimental questionnaire rating their use of different strategies during No-Think trials to verify that participants followed our direct suppression instructions. All strategies were rated on a 5-point scale ranging from 0 *never* to 4 *always* indicating use of each strategy and whether they intentionally did not follow experimental instructions, and how awake or sleepy they felt before they started the experiment. We further measured *intentional* non-compliance with No-Think instructions to ensure that all participants had genuinely tried to exclude memories from awareness (Hertel & Calcaterra, 2005). Intentional non-compliance was measured with three questions: (a) When I saw the red cue word, I quickly checked to see if I remembered the target word, (b) After a red cue word went off the screen, I checked to see if I still remembered the target word, and (c) When I saw a red cue word, I thought about the target word that went with it in an effort to improve my memory for that word pair. These questions were rated on a 5-point scale ranging from 0 *never* to 4 *very frequently*.

Results

All analyses are based only on pairs for which participants recalled the target on the final learning test (Anderson et al., 2004). Overall, learning performance on this test was sufficiently high: 72.44% (SE = 1.66). To retain power in all subsequent analyses, slight violations of sphericity were corrected with either Greenhouse-Geisser ($.70 \geq \varepsilon < .75$) or Huynh-Feldt corrections ($\varepsilon \geq .75$). In case of severe violations ($.70 < \varepsilon$) a multivariate test statistic (Pillai-Bartlett trace; V) is reported. In all analyses counterbalancing condition was included as a between-subjects factor to account for item effects, and nonsignificant results of this factor (or its interactions) are not reported.

Final Recall Performance

Though intrusion reports from the Think/No-Think phase were our primary interest, we replicated prior findings concerning suppression-induced forgetting on the final test. Replicating past work, suppression affected final recall, as evidenced by poorer recall of target words in the suppress conditions overall (No-Think-short M = .83, SE = .013; No-Think-long M = .841, SE = .013) compared to recall of target words from Baseline pairs (M = .861, SE = .011), $F(1, 35) = 4.337$, $p = .045$, $\eta_p^2 = .11$. There was no reliable difference in recall between short-duration and long-duration suppression conditions, $F(1, 35) = .378$, $p = .543$, nor were there interactions with test type (i.e., same probe or independent probe test), Fs < 1.3. None of the foregoing suppression-induced forgetting effects (or comparisons of short vs long No-Think) interacted with test-order counterbalancing (all ps > .25).

Replicating Key Patterns in Control over Intrusions During the Think/No-Think Task

Our main hypothesis was that longer trial durations would be accompanied by a greater number of intrusions than would shorter trial durations, as measured by our trial-by-trial intrusion reports collected during the Think/No-Think phase. However, we first assessed whether we replicated past work on reports of conscious awareness in the Think/No-Think phase (e.g., Levy & Anderson, 2012), by conducting a 2 (condition: Think/No-Think) × 8 (repetition) ANOVA, with short- and long-duration trials averaged per condition. If participants reported that a memory came to mind 'briefly', or 'often' during a No-Think trial, we considered this an intrusion; if they reported that a memory 'never' came to mind during the trial, we considered this a non-intrusion. Participants also made the same judgments after Think trials, in which case the awareness of the memory was the desired goal, so this would not be considered an "intrusion" per se. But the reports in the two conditions can be compared nonetheless, to quantify variation in mnemonic awareness as a function of task goals.

Directly comparing reports of mnemonic awareness across the Think and No-Think conditions, we observed an exceptionally robust main effect, $F(1, 35) = 212.839$, $p < .001$, $\eta_p^2 = .859$, demonstrating that people experienced target memories entering awareness far less often during No-Think trials ($M = .47$, $SE = .034$) than during Think trials ($M = .98$, $SE = .005$). Thus, according to participants' phenomenal reports, they showed a remarkable overall ability to voluntarily control episodic retrieval. Nevertheless, intrusions still were experienced on a significant proportion (approximately half of the time) of No-Think trials (see Figure 1, left panel).

10

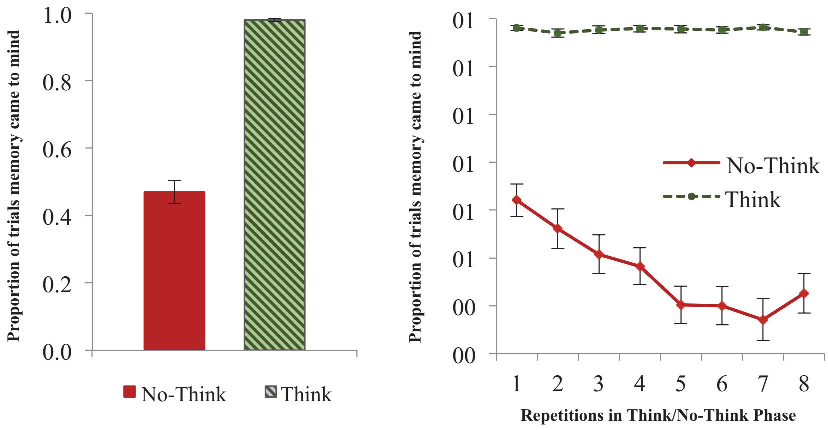


Figure 1. Proportion of trials during which the memory came to mind displayed for No-Think trials and Think trials (left panel), and the proportion of intrusions over repetitions in the Think/No-Think Phase (right panel) for No-Think and Think trials. Error bars represent standard errors of the mean.

Reports of intrusions declined substantially across suppression repetitions. This pattern is reflected in a robust Repetition × Condition interaction, V = .602, $F(7, 29) = 6.255$, $p <$ .001, $\eta_p^2 = .602$. This interaction was followed up by one ANOVA for Think trials and one for No-Think trials. These showed that reports of awareness did not vary over repetitions for Think trials V = .142, $F(7, 29) = .689$, $p = .681$, but were reduced substantially from the first No-Think repetition (M = .62 , SE = .034) to the eighth one (M = .43 , SE = .041), V = .599, $F(7, 29) = 6.196$, $p < .001$, $\eta_p^2 = .599$. Thus, participants grew increasingly effective at suppressing intrusions of No-Think items as repetitions progressed, and remained highly effective at bringing Think trials to mind across all repetitions (see Figure 1, right panel). These results strongly confirm findings from prior neuroimaging studies using this trial-by-trial intrusion scale (Levy & Anderson, 2012; Benoit, et al, 2014; Hellerstedt et al., 2016). Increasing success at controlling intrusions may reflect some combination of the accumulation of inhibition on suppressed traces over repetitions, improved skill at stopping retrieval in general, and a passive decline in the accessibility of No-Think items over time, though prior work clearly indicates a significant contribution of inhibitory control (e.g., Levy & Anderson, 2012; Benoit, Hulbert, Huddleston, & Anderson, 2014).

Sustaining Inhibitory Control Continuously is Demanding

To test whether confronting a reminder during No-Think trials for longer periods of time is more demanding than confronting a reminder for shorter periods of time, we included trial duration (short vs. long) as a within-subjects factor in our analysis of intrusion reports, and conducted a 2 (duration: Short/Long) × 8 (repetition) ANOVA. This analysis revealed a highly reliable difference in the frequency of intrusions between short (M = .44, SE = .035) and long duration (M = .50, SE = .035) No-Think trials, $F(1, 35) = 22.856$, $p < .001$, η_p^2 = .395. Thus, longer trial durations likely imposed a greater burden on inhibitory control, as reflected in the increased probability of intrusions, an effect that did not interact with repetition, F <1. This trial duration effect also was apparent when intrusion reports for No-Think trials were broken down into 'brief' and 'often' responses. Brief intrusions arose more often during long (M = .46, SE = .03) than during short trials (M = .42, SE = .03), $t(39)$ = 3.015, p = .005. Similarly, participants reported that an intrusion 'often' came to mind with greater frequency during long (M = .04, SE = .009) than during short trials (M = .02, SE = .004), $t(39)$ = 2.339, p = .025. Consistent with our hypothesis concerning trial durations, 'never' responses were mirrored to 'briefly' and 'often' responses; there were more 'never' responses on short trials (M = .56, SE = .03) compared to long trials (M = .50, SE = .03), $t(39)$ = 4.149, p < .001.

Further evidence for greater demand on control processes can be plainly seen in the cumulative intrusion functions for our duration conditions. In these functions, the number of intrusions on each repetition is the sum of intrusions reported during that repetition and all preceding repetitions. A 2 (duration: Short/Long) × 8 (repetition) ANOVA showed that on long duration trials there were more intrusions overall, $F(1, 35)$ =14.393, p = .001, η_p^2 = .291, and a significant Duration × Repetition interaction showed that the number of intrusions accumulated more quickly over No-Think repetitions for long-duration compared to short duration trials, V = .529, $F(7, 29)$ = 4.658, p = .001, η_p^2 = .529 (see Figure 2). Moreover, although people often were able to successfully control intrusions (i.e., make intrusions of an item stop entirely for all of its remaining No-Think trials), it took significantly more trials to achieve this goal for long duration No-Think trials. To quantify this, we computed, for each item, the average number of No-Think repetitions until the last intrusion was experienced. For longer duration trials, the average serial position of the final intrusion occurred later during the TNT phase (M = 5.95, SE = .24), than it did for short duration trials (M = 5.32, SE = .28), $t(39)$ = 3.754, p = .001. Thus, increasing trial duration made it harder to succeed quickly at suppression, likely reflecting the increased demand placed on inhibitory control mechanisms.

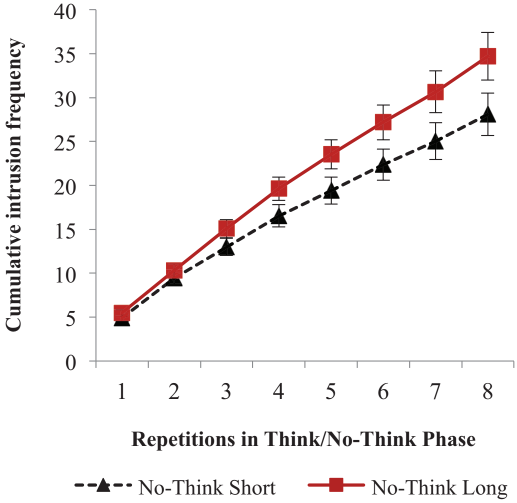


Figure 2. Cumulative intrusion frequency per repetition for short and long duration No-Think trials, which reflects the sum of reported intrusions on each repetition and all preceding repetitions. Error bars represent standard errors of the mean.

Next, we investigated the nature of the relationship between time and the probability of an intrusion occurring. It is conceivable that if one intrusion is experienced within a short duration No-Think trial (i.e., 2.5s), then 2 intrusions should be experienced within a long duration No-Think trial (i.e., 5s). This would be classified as an additive relationship, wherein intrusions increase uniformly with the number of seconds (see Figure 3, line with circular markers). Alternatively, the relationship may be superadditive; in this case, the number of intrusions experienced over time is higher than what would be expected based on an additive relationship. Long duration No-Think trials would show more than twice the number of intrusions compared to short duration trials, perhaps reflecting increasing within-trial fatigue. A final possibility is an underadditive relationship, wherein with increasing trial duration, intrusions may still increase, but less than would be expected based on a simple additive relationship. Figure 3 illustrates an exploratory analysis we conducted illustrating that an underadditive relationship best characterizes our data. The proportion of intrusions for long duration trials is much less than twice the amount observed during short trials, and the cumulative recall function increases for long trials, but not at a rate that is double the slope of the short trials.

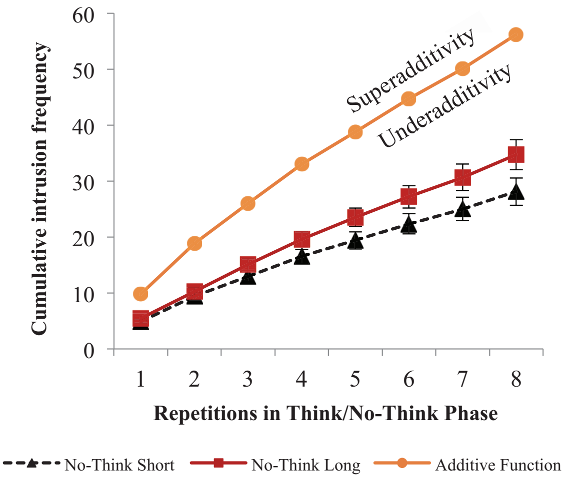


Figure 3. Additive function for cumulative intrusion frequency, which is plotted based on the average number of cumulative intrusions in short-duration No-Think trials times two. It represents a uniform increase of intrusions with time. The position of the long duration No-Think cumulative intrusions relative to the additive function, shows that the relationship between time and the probability of an intrusion occurring is underadditive (i.e., the No-Think duration line lies below the additive function). Error bars represent standard errors of the mean.

To further characterize the relationship between time and an intrusion occurring, we looked at the number of intrusions per second for short and long duration trials. We estimated the number of intrusions per second by totaling the number of intrusions reported for each item across its 8 repetitions, and then dividing by the total number of seconds the participant spent suppressing that item (e.g. 8 suppression attempts × 2.5 seconds per attempt = 20s; or × 5 seconds per attempt = 40s). Remarkably, the number of intrusions per second was much lower in the long (M = .10, SE = .007) than in the short trials (M = .17, SE = .01), $t(39)$ = 9.067, p < .001. To verify this difference was not the result of an overall difference in the total time suppressing unwanted memories in the short (20s of total suppression time) and long trials (40s of total suppression time), we restricted our comparison to the first 20s within the short and long conditions. Thus, for the short condition, the first 20 seconds would cover all 8 repetitions (2.5 seconds × 8 repetitions), whereas for the long condition, it would cover the first 4 repetitions (5 seconds × 4 repetitions). In addition, to illustrate how the number of intrusions *per unit time* changed with repetitions, we compared the development of intrusion rates across the first 20 seconds in the long-duration condition and the first 20 seconds in the short duration condition. To do this, we divided the first 20 seconds into 4 equal size bins of 5 seconds each. For the long-duration condition, these bins were simply the first four 5 suppression trials, each of which was 5 seconds long. For the short-duration condition, these 5-second bins were each composed by aggregating over two consecutive 2.5-second long trials on a given item. For each of these 5-second bins, we then summed the intrusions observed within it. This enables us to plot intrusion rate (intrusions per 5-second bin) across 4 time bins matched for duration in the short and long conditions. A 2 (duration: Short/Long) × 4 (repetition interval) ANOVA showed that even with this more precise matching of total time, the average number of reported intrusions per 5 second bin in long condition (M = .57, SE = .07 or .11 intrusions per second) remained lower than in the short condition (M = .87, SE = .034 or .17 intrusions per second), $F(1, 35) = 44.561$, $p < .001$, $\eta_p^2 = .56$. Interestingly, a significant Duration × Repetition interaction showed that though intrusions per each 5s intervals in the short condition decreased with repetition, they remained higher at the 4th repetition, compared to long condition intervals, $F(3, 105) = 10.46$, $p < .001$, $\eta_p^2 = .23$ (see Figure 4).

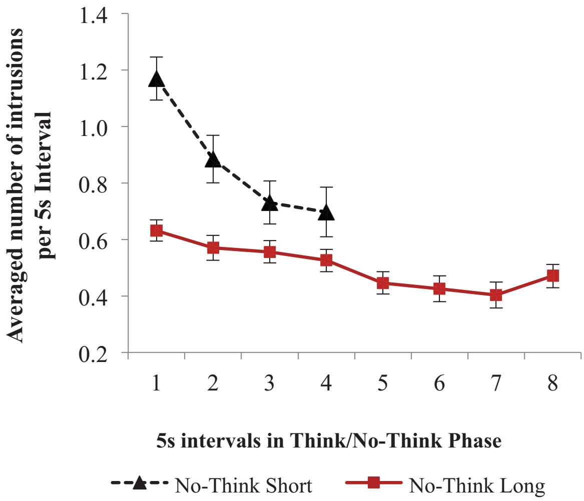


Figure 4. Average number of intrusions per 5s interval for the short and long duration No-Think conditions. Each 5s interval for the short duration No-Think condition is the sum of two repetition intervals that were 2.5s originally. Error bars represent standard errors of the mean.

Finally, we considered the possibility that the foregoing conclusions about underadditivity and intrusion rate might have arisen because we failed to adequately consider trials that contained multiple intrusions, which occurred slightly more often on long-duration than short duration trials (4% of the trials instead of 2%). Specifically, the foregoing analyses do not distinguish between trials where participants report a single brief intrusion (the subject responded "briefly" on the intrusion scale) and trials where participants report having experienced more than one intrusion (the subject responded "frequently"), because, in either case, a participant was simply classified as having an intrusion on that trial. It is possible that if we gave credit for trials with more than one intrusion, the conclusions might differ. To address this, we recomputed all analyses using a graded scale (0 for non-intrusions, 1, for a brief intrusion, 2 for frequent intrusions). All of the foregoing analyses turned out the same, with differences between conditions in fact growing slightly larger and more reliable, not less (intrusions per second in the long versus short conditions, M = .11 and .18 respectively, p < .001; in the analysis using only the first 20s, .124 versus .184, respectively (.62 vs .92 intrusions per 5-second bin), $p < .001$, $\eta_p^2 = .58$). The similarity in the analyses reflects the fact that multi-intrusion trials were rare (only 3% of reports) and thus could not exert much influence on the data. Thus, our conclusions about under-additivity and intrusion rates seem unlikely to derive from this analysis choice. Nevertheless, although "frequent" responses were a very small part of the data, it would be desirable, in future studies, to quantify the number of intrusions on such trials more precisely to yield a more accurate measurement.

Repeated Inhibitory Control is Demanding

If applying inhibitory control is effortful, the quality of that effort may decline as blocks progress, particularly in the later blocks and under more strenuous conditions. To test this we calculated a transition score between every two sequential repetitions of each item (e.g., the 2nd suppression attempt of item X, and the 3rd attempt on X, irrespective of other items that may have intervened between those repetitions of X), and determined in what proportion of transitions an intrusion was followed by a non-intrusion (i.e., new successes), and a non-intrusion was followed an intrusion (i.e., relapses). Successes give us a measure of how well one's efforts to suppress retrieval of a particular item carry forward to later trials with that item; this may indicate persisting aftereffects of suppression on the item in question. Relapses, on the other hand, suggest that one's prior successes at controlling awareness did not produce lasting effects on the suppressed item, and possibly signal a lapse in applying control mechanisms to that item.[6]

For successes, there was an effect for Repetition which shows that the probability of a new success (i.e., transitions from intrusions to non-intrusions) increased reliably over repetitions at first, but then the pattern reverses in later blocks displaying a clear quadratic relationship, $\chi2$ (35) = 89.191, $p < .001$ (see Figure 5, left panel). There was no reliable difference in the probability of new successes between short- and long-duration conditions, $\chi2$ (35) = -4.719, $p > .05$, nor was there an interaction of No-Think trial duration with repetition, $\chi2$ (210) = -6.787, $p > .05$. Taken together, these findings indicate that the probability of achieving a new successful control declined suddenly towards the end of the TNT task, consistent with the possibility that participants were growing fatigued during the final repetitions. Alternatively, the drop in new successes may reflect greater difficulty in suppressing the remaining intrusions, which may be the most persistent and demanding items.

In contrast to new Successes, Relapses did show sensitivity to trial duration; relapses were significantly more frequent in the long duration condition (M = .376, SE = .029) compared to the short duration condition (M = .309, SE = .031), $\chi2$ (35) = 63.55, $p = .002$ (see Figure 5, right panel). There was no main effect of repetition, $\chi2$ (108) = -4.137, $p > .05$ or an

[6] We used multiple imputation in SPSS (Version 22) for transition scores in successes and relapses, because 5.63% of the data were missing. Little's MCAR test shows no identifiable pattern exists for the missing data, χ^2 (583) = 596.58, $p = .339$. Next, 25 datasets were imputed (automatic imputation method). Regular ANOVAs for main or interaction effects were performed on the imputed datasets and all 25 F-values were pooled using the R package miceadds, which produces a χ^2 statistic per effect to test for significance (Robitzsch, Grund, Henke, & Robitzsch, 2016).

interaction of duration with repetition, $\chi2$ (210) = -19.126, $p > .05$. This suggests that when trial duration was longer, it was more difficult for suppression to induce a persisting impact that carried forward to the next suppression trial for a given item.

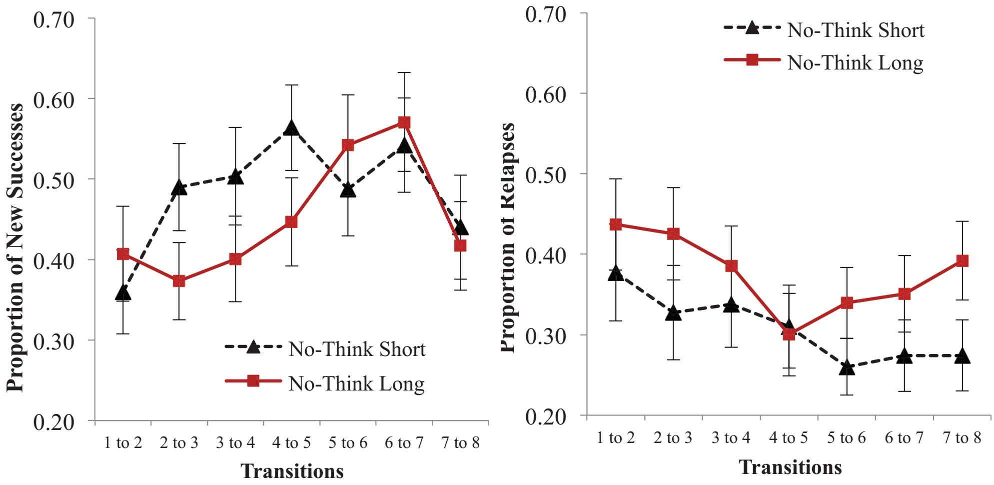


Figure 5. Proportion of new successes (left panel) and relapses (right panel) per repetition-to-repetition transition for short and long duration No-Think trials. New successes reflect the transition from an intrusion to non-intrusion, while relapses represent the transition from a non-intrusion to an intrusion. Pooled means and standard errors of the mean are presented.

We also sought to relate the foregoing patterns to participants' subjective reports of alertness. Post-experimentally, we asked participants to rate, on a 5-point scale, their alertness at the outset of the experiment. Ratings on this scale were uniformly high (M = 4.18), so there unfortunately was little variability in initial alertness that could be related to subsequent intrusions. Thus, participants having lower alertness at the outset of the study did not reliably predict increased relapse probability (when examined overall, r = -.16; when focusing on the second half of TNT phase, r = -.26), nor did it predict decreased new successes (overall, r = .01, second half, r = -.05). However, the 3 participants who reported unusually compromised alertness (scores of 1 or 2) did have a far higher relapse probability (M = .63) than the remaining subjects with relatively normal alertness (M = .37), especially when considering the second half of the TNT phase (Low = .68, Remainder = .31).

Compliance.

All (100%) of participants reported that they *often* or *always* used strategies consistent with direct suppression instructions, such as staring intently at the cue word, repeating the cue word silently, or letting their mind go blank in response to the cue word. Strategies consistent with thought-substitution, such as generating a word, thought or sound in response to the No-Think cue, were never (0%) reported as being used *often* or *always*. A minority of participants (45%) reported they *rarely* or *sometimes* used a thought substitution strategy. Thus, participants largely complied with our instructions to directly suppress retrieval of No-Think targets. Moreover, non-compliance with No-Think instructions was low; the majority of participants indicated *never* or *rarely* trying to intentionally violate No-Think instructions, respectively 87.5%, 95%, and 97.5% for our three compliance questions (see Methods).

Discussion

The findings of the current study support the conclusion that control over involuntary retrieval is more challenging if it must be sustained for longer periods of time. To characterize how challenging it was for participants to control retrieval, we used a trial-by-trial intrusion report method developed in prior work to document participants' phenomenal experience of intrusions during retrieval suppression. These reports have been used successfully to identify distinctive hemodynamic and electrophysiological signatures of inhibitory control over memory (Levy & Anderson, 2012; Benoit et al., 2014; Hellerstedt et al., 2016). Using intrusion reports, we tested two main hypotheses. First, if needing to continuously suppress episodic retrieval in response to a prepotent reminder for longer intervals taxes inhibitory control processes to a greater extent, more intrusions should be observed during long-duration trials (5s) than short duration trials (2.5s). Increased intrusion frequency during longer trials may be produced either by more sustained input to the retrieval process owing to attention to the cue, or to lapses in task set maintenance with longer intervals, either of which would make stopping retrieval more challenging. The results provide clear support for this duration effect: participants reported experiencing intrusions with a greater probability during longer suppression trials than during shorter ones. This pattern arises primarily from an increase in the number of trials on which a single intrusion occurred, rather than an increase in the number of intrusions per trial, though a relatively modest increase in multi-intrusion trials also was evident in the data. The increase in the probability of an intrusion occurring was present across all suppression repetitions. These findings demonstrate that being exposed to reminders for longer durations increases the challenge posed to controlling retrieval. This increased challenge is also reflected in the fact that, on average, it took participants 5.95 trials to fully control all intrusions of an item for long trials, but only 5.32 repetitions for short duration trials.

Beyond the effects of trial duration, we also hypothesized that there would be a gradual deterioration of the ability to control intrusions when control must be maintained over very long intervals. In a typical Think/No-Think task, participants engage in multiple blocks of memory control that, taken together, may last between 30 and 60 minutes, and we have previously noted that participants can become fatigued after such intervals (Anderson & Huddleston, 2011). If sustaining control in this manner is challenging, evidence for increasing burden may be evident in intrusion reports, which may reveal participants' declining ability to control retrieval towards the end of our Think/No-Think task. We predicted that although this pattern would be seen in changes in overall intrusion frequency, it would be particularly striking in our sequential dependency analyses of the intrusion data. The data partially support these predictions. On the one hand, we found the predicted deterioration in new successes (cases in which an intrusion on trial N is followed by a non-intrusion on trial N+1) towards the end of the TNT phase, (see Figure 5); on the other hand, the predicted quadratic trend in relapses (cases in which a non-intrusion on trial N is followed by an intrusion on trial N+1), while numerically present for long-duration trials (see Figure 5), was not reliable. Long duration trials were, nevertheless associated with more relapses than were short-duration trials, consistent with the notion that they impose greater demands on inhibitory control.

One interpretation of these cross-block changes in the success of inhibitory control is that participants suffered from increasing fatigue after 35 minutes of performing a demanding retrieval suppression task. This possibility is broadly consistent with the sensitivity of executive functioning to sleep deprivation, which shows that sleep deprivation increases the difficulty in inhibiting inappropriate responses (Chuah, Venkatraman, Dinges, & Chee, 2006; Drummond, Paulus, & Tapert, 2006). For example, a recent study using a Go-No Go task showed that mental fatigue indeed gives rise to delayed motor inhibition (Kato, Endo, & Kizuka, 2009). Because there is overlap between the systems involved in motor and memory inhibition (Anderson & Weaver, 2009; Depue, Orr, Smolker, & Banich, 2015), mental fatigue may affect memory inhibition in a similar fashion. However, although fatigue provides a plausible explanation for these cross-block patterns, the present data cannot distinguish fatigue from other factors that could also have changed over blocks, any of which may account for poor task performance, including changes in mood, and, perceived novelty of the task (i.e., boredom). There is evidence indicating, for example, that the sort of cognitive conflict that triggers cognitive control is an aversive state (Inzlicht, Bartholow, & Hirsh, 2015) that people seek to avoid or minimize such cognitive effort (Kool, McGuire, Rosen, & Botvinick, 2010). Thus, it is possible that engaging in inhibitory control over a sustained interval may induce negative mood. Future studies should seek to distinguish these accounts, perhaps through a detailed characterization of fatigue, affect, and interest throughout the task. Our current measurement of alertness was restricted to alertness at the outset of the task, and (fortunately or unfortunately) revealed nearly all participants to feel fully alert.

Despite the clear increase in reported probability of intrusions for longer duration trials, a separate question remains whether control is more *efficiently* exercised if it is done over shorter or longer intervals. One way to quantify the efficiency of control is to compare the rate of intrusions per second across the two conditions. Interestingly, by this measure, longer duration trials were more efficient, in that they are associated with significantly fewer intrusions per second (.10) than were shorter duration trials (.17). This conclusion held when we focused on the first 20 seconds of suppression for each condition (note that there the short duration condition involved a total of only 20s — i.e., 2.5s across 8 trials) and compared "time bins" of the same size (i.e., 5 second bins—which is a single trial in the long duration condition, and two trials in the short duration condition). Thus, the true probability of intrusions as a function of time suppressing was much lower for long duration trials (.11 intrusions per second, or .57 per 5 seconds) than it was for shorter duration trials (.17 intrusions per second, or .87 intrusions per 5 seconds), suggesting that participants got better results from their efforts. Indeed, as can be seen in Figure 4, the increased intrusion rate for shorter duration trials is particularly pronounced during early bins, and shrinks rapidly with repeated effort. In contrast, for longer duration trials, the intrusion rate is uniformly lower, and appears relatively constant over time bins, with little evidence for improvement in rate over repetitions. This increased efficiency with longer trial durations is likely to explain the pattern of underadditivity we observed in our data: we found that 5-second trials were characterized by far fewer intrusions than would be expected if these trials were equivalent to two 2.5-second trials, and inhibition efficacy was strictly a function of the time devoted to the task.

Why would control over intrusions be more efficient if it is applied for longer intervals? One likely explanation is that when performing two 2.5-second trials versus one 5-second trial, participants need to reinstate the control operation twice in the former instance. Transitioning into a No-Think trial surely involves switching of task sets (Allport, Styles, & Hsieh, 1994), and, until the suppression task set is successfully implemented by the participant on a given trial, they run the risk of being reminded of the associated memory. Put differently, the race between two parallel processes – the involuntary retrieval process on the one hand, and the controlled implementation of the inhibition task set on the other – may be lost by the inhibition process some fraction of the time, yielding an intrusion. Thus, when comparing the same total amount of time (5s), there are two occasions on which the task set must be implemented in the short duration condition and only one in the longer duration condition. Long duration trials enable participants to simply sustain an already implemented control operation and devote more time to application of control and less to its initial implementation. If this interpretation is correct, then the steeper decline in the rate of intrusions per second during short duration trials over consecutive time bins suggests (relative to that observed during long-duration trials; see Figure 4), that short-duration trials, while less efficient, are more effective in driving some form of durable improvement. This could, in principle reflect more effective buildup of inhibition on suppressed traces during short duration trials.

The current study demonstrates the importance of trial duration in experiments on motivated forgetting. In previous experiments there has been considerable between-experiment variation in No-Think trial duration; anywhere between 2s (Bergström, deFockert, & Richardson-Klavehn, 2009a) and 6s (Marx, Marshall, & Castro, 2008), or even 8s (Wang, Cao, Zhu, Cai, & Wu, 2015). This may have inadvertently affected the number of experienced intrusions and, with it, the size of suppression-induced forgetting. Indeed our effects of trial duration are consistent with the interpretations offered by Lee et al. (2007), who speculated that long trial durations during the Think/No-Think task are associated with a greater burden on inhibitory control. In support of this idea, Lee et al. (2007) observed less suppression-induced forgetting on long-duration (5 seconds) trials compared to short-duration (3 seconds) trials. Because the current study added a trial-by-trial intrusion report measurement, we were able to show that long duration trials were indeed more likely to elicit the need for control; these trials were accompanied by more intrusions, a faster accumulation of intrusions over repetitions, and a larger number of trials until control over intrusions fully succeeded. Unlike Lee et al. (2007), however, we did not find a clear distinction between short and long duration No-Think trials in suppression-induced forgetting, as assessed with the final recall test. This finding suggests that participants' reports of intrusions provide a more sensitive measure for assessing the effects of the suppression process, compared to a voluntary final recall test. It is possible, for example, that control processes engaged during explicit recall on the final test may undo the effects of suppression (especially on partially suppressed items) in a way that does not arise when accessibility of No-Think items is assessed by the propensity for an item to involuntarily intrude.

Our evidence that "relapses" in controlling intrusive memories are more likely with longer trial durations may shed light on another frequently used paradigm investigating the nature and controllability of intrusive thoughts. Studies using the "white bear" paradigm have found that suppressing thoughts about a white bear continuously for 5 minutes often paradoxically increases the presence of the suppressed thoughts as measured by a similar self-report process to the one used here (Wegner, Schneider, Carter, & White, 1987). Work with this paradigm has suggested that suppression is not only an ineffective method of mental control, but may also be a counterproductive one (Wenzlaff & Wegner, 2000). Although there are a variety of important ways in which the Think/No-Think and White Bear paradigms differ (see Anderson & Huddleston, 2012 for a discussion with hypotheses), the current experiment suggests that a key difference may lie in the total amount of time participants must continuously suppress. Our data suggest that purging unwanted thoughts of white bears may be more successful when purging attempts are implemented in short spikes of activity. "White bear" experiments may yield a counterproductive pattern partially as a result of a heavy burden placed on inhibitory control arising from the need to suppress an unwanted thought for a very long period of time. The greater effectiveness of suppression attempts on shorter distributed trials may be related to the benefits of spaced practice over massed practice (Donovan & Radosevich, 1999).

To the extent that our findings provide evidence that sustained control (whether on a trial or session level) can be fatiguing, they may have implications for psychopathology. It has been put forward that inhibitory deficits may be the reason that intrusive memories or thoughts can be observed in a wide-range of psychopathologies, such as PTSD (flashbacks), obsessive–compulsive disorder (obsessions), or depression (rumination) (e.g., Anderson & Hanslmayr, 2014; Catarino et al., 2015; Fawcett et al. 2015; Hertel, & Gerstle, 2003; Joorman et al., 2009; Marzi et al., 2014). Given the prevalence of sleep disorders in psychiatric conditions (Wulff et al., 2010), fatigue is especially important to identify, because fatigued patients may be less likely to engage inhibitory control effectively. As a result they will experience more intrusive memories than those who are well rested, perhaps further deteriorating their psychiatric condition.

The intrusion report method used here identifies and quantifies the frequency of counter-intentional retrievals that are triggered by strong reminders to an unwanted memory. Because these retrievals occur despite participants' efforts to stop them, and because they recur whenever cues are presented, we have argued that the processes triggered may be related to those underlying intrusive memories in psychopathology. Consistent with this view, recent discussions of how intrusive memories differ from involuntary retrievals more broadly have emphasized qualities such as their perseverative nature and how they arise, despite being unwanted (Kvavilashvili, 2014). However, clinically relevant intrusive memories also are emotionally intense and stressful, and can be highly disruptive to the individual experiencing them, unlike the intrusions studied here. These differences are clearly important and may signify fundamentally different retrieval mechanisms from the ones studied here. Our view, however, is that the emotional nature of naturally occurring intrusive memories and their personal significance likely function to substantially increase the frequency with which people "self-cue" the internal triggers that lead to the intrusions, via their emotional state or thoughts, but that, once cued, the involuntary retrieval processes are fundamentally similar to the ones at work here. However, regardless of one's theoretical assumptions, the current paradigm provides a useful tool for comparing the intrusive qualities of a wide variety of memories--whether neutrally-, negatively-, or positively-valenced, or whether autobiographical or laboratory-based--and testing the extent to which the cognitive and neural mechanisms triggered are similar or distinct.

In conclusion, the current findings provide evidence supporting the view that intrusive memories can be controlled, but that it is harder to do so when control must be sustained. These findings suggest that the need to sustain attention to reminders indeed places a heavier burden on inhibitory control. Our novel approach of analyzing sequential dependencies between intrusion repetitions confirmed this and additionally showed that needing to sustain control over very long intervals decreases successful suppression of unwanted memories. The finding that sustained engagement compromises inhibitory control is also especially relevant for understanding intrusive thoughts and memories in psychopathologies, particularly to the extent that such deficits arise from fatigue. Indeed, fatigue could, in principle, be a key factor contributing to failed memory control in psychiatric conditions. In future research, phenomenal reports of intrusions of the sort used here may provide a helpful method for studying, in-depth, how intrusive memories are controlled in different psychiatric conditions. Indeed, the particular sequential dependency patterns operationalized here (e.g., relapse probability) may provide new measures that could predict vulnerability to intrusive symptomatology, above and beyond individual variation in suppression-induced forgetting (Levy & Anderson, 2008).

Acknowledgements

This work was supported by the Medical Research Council of the UK (MC-A060-5PR00 to MCA). We thank Lea Jabbarian for her assistance in testing, Christiaan Tieman and Marcel Boom for technical support, and Gerko Vink for helpful advice on performing multiple imputation.

Author contributions

KvS and MCA designed the research, analyzed the data, and wrote the paper. KvS did the data collection.

Appendix

Table 1. Cue, target, and independent probe for the words in the five experimental groups and words in the filler group. The original Dutch words are presented accompanied by English translations.

Group	Cue	Target	Independent Probe
1	Gelei (Jelly)	Ruw (Crude)	Grof (Coarse)
	Jurk (Dress)	Vaandel (Banner)	Vlag (Flag)
	Vork (Fork)	Paard (Horse)	Ruiter (Horseman)
	Molen (Mill)	Vin (Fin)	Vis (Fish)
	Zuilen (Columns)	Viool (Violin)	Snaar (Snare)
	Azië (Asia)	Rivier (River)	Stroom (Stream)
	Sproet (Freckle)	Boodschap (Message)	SMS (Text message)
	Koning (King)	Heuvel (Hill)	Berg (Mountain)
	Hooi (Hay)	Perron (Platform)	Trein (Train)
	Gordijn (Curtain)	Poster (Poster)	Muur (Wall)
	Radio (Radio)	Sneeuw (Snow)	Winter (Winter)
	Stoel (Chair)	Vuur (Fire)	Vlam (Flame)
2	Spook (Ghost)	Neef (Nephew)	Tante (Aunt)
	Bont (Fur)	Fornuis (Stove)	Koken (To cook)
	Helm (Helmet)	Cassette (Cassette)	Video (Video)
	Standbeeld (Statue)	Verf (Paint)	Kwast (Brush)
	Kin (Chin)	Gras (Gras)	Tuin (Garden)
	Vat (Barrel)	Non (Nun)	Klooster (Cloister)
	Tomaat (Tomato)	Vest (Vest)	Knopen (Buttons)
	Uur (Hour)	Cent (Cent)	Munt (Coin)
	Blouse (Blouse)	Zwaard (Sword)	Ridder (Knight)

	Smaragd (Emerald)	Ballet (Ballet)	Dans (Dance)
	Museum (Museum)	Kluis (Safe)	Code (Code)
	Korst (Crust)	Leraar (Teacher)	Leerling (Pupil)
3	Bezem (Broom)	Pop (Doll)	Speelgoed (Toys)
	Kaak (Jaw)	Noord (North)	Zuid (South)
	Tolk (Interpreter)	Cirkel (Circle)	Kring (Circle)
	Nonchalant (Nonchalant)	Boek (Book)	Lezen (To read)
	Raam (Window)	Maag (Stomach)	Orgaan (Organ)
	Uil (Owl)	Kantoor (Office)	Bureau (Desk)
	Patent (Patent)	Ei (Egg)	Dooier (Yolk)
	Krant (Newspaper)	Planeet (Planet)	Aarde (Earth)
	Clown (Clown)	Lens (Lens)	Bril (Glasses)
	Gewoonte (Habit)	Toren (Tower)	Kerk (Church)
	Gang (Corridor)	Ketel (Kettle)	Stoom (Steam)
	Machine (Machine)	Deugd (Virtue)	Geduld (Patience)
4	Strik (Bow)	Laars (Boot)	Schoen (Shoe)
	Dorp (Village)	Olijf (Olive)	Olie (Oil)
	Hoek (Corner)	Zalf (Ointment)	Tube (Tube)
	Spieren (Muscles)	Dwerg (Dwarf)	Klein (Small)
	Dun (Thin)	Meter (Meter)	Lengtemaat (Measure of Length)
	Item (Item)	Slapen (To sleep)	Droom (Dream)
	Vlecht (Braid)	Valk (Falcon)	Roofvogel (Bird of prey)
	Kwart (Quart)	Arm (Arm)	Elleboog (Elbow)
	Nederig (Humble)	Lamp (Lamp)	Licht (Light)
	Voetbal (Soccer)	Tiener (Teenager)	Jong (Young)
	Stad (City)	Fluit (Flute)	Orkest (Orchestra)
	Walnoot (Wallnut)	Blauw (Blue)	Kleur (Color)
5	Ivoor (Ivory)	Gedicht (Poem)	Rijm (Rhyme)
	Room (Cream)	Citroen (Lemon)	Zuur (Sour)
	Laan (Lane)	Kast (Cabinet)	Lade (Drawer)
	Hardloper (Runner)	Nacht (Night)	Donker (Dark)
	Rum (Rum)	Reparatie (Repair)	Garage (Garage)
	Hek (Fence)	Mos (Moss)	Bos (Forest)
	Sprong (Leap)	Huis (House)	Baksteen (Brick)
	School (School)	Papier (Paper)	Pen (Pen)
	Schaduw (Shadow)	Wortel (Carrot)	Konijn (Rabbit)

10

	Schort (Apron)	Vakantie (Holiday)	Zon (Sun)
	Bus (Bus)	Kelder (Cellar)	Zolder (Attic)
	Priester (Priest)	Schip (Ship)	Kapitein (Captain)
Filler	Kinderkamer (Nursery)	Mand (Basket)	Riet (Reed)
	Globe (Sphere)	Klok (Clock)	Tijd (Time)
	Werktuig (Utensil)	Vinger (Finger)	Duim (Thumb)
	Gereedschap (Tool)	Schrijver (Writer)	Auteur (Author)
	Paraplu (Umbrella)	Schaar (Scissors)	Knippen (To cut)
	Golfer (Golfer)	Taxi (Taxi)	Chauffeur (Driver)
	Lipgloss (Lipgloss)	Deur (Door)	
	Gebruik (Custom)	Vierkant (Square)	
	Laken (Sheets)	Dokter (Doctor)	
	Onderdeel (Part)	Kom (Bowl)	
	Apparaat (Appliance)	Koeien (Cows)	
	Spray (Spray)	Snoer (Cord)	
	Formulier (Form)	Tijm (Thyme)	
	Verrekijker (Binoculars)	Kauwgom (Chewing gum)	
	Salade (Salad)	Koffie (Coffee)	
	Wolk (Cloud)	Straat (Street)	
	Reclame (Commercial)	Brug (Bridge)	
	Badkamer (Bathroom)	IJzer (Iron)	

CHAPTER 11

GENERAL DISCUSSION

General Discussion

The aim of this thesis was to gain more insights in how unwanted memories can be changed. We used an experimental psychopathology approach to examine three different interventions that target unwanted memories: (1) taxing working memory by dual-tasks, (2) memory disruption and updating during reconsolidation via new learning, and (3) retrieval suppression. The results will be summarized below and will be followed up with a discussion per intervention (including future directions). This chapter broadens out and presents an integrative discussion (of methods) and clinical implications.

Taxing Working Memory by Dual-Task Interventions

A substantial body of evidence shows that performing a dual-task intervention (e.g., making eye movements, playing Tetris or counting) simultaneously with recall reduces self-reported ratings of vividness and/or emotionality compared to a recall only control condition (e.g., Andrade et al., 1997; van den Hout, Muris, Salemink, & Kindt, 2001; for a meta-analysis see, Lee & Cuijpers, 2013). This dual-task intervention is clinically exploited in Eye Movement Desensitization and Reprocessing (Shapiro, 1989). The studies in **chapter 3**, **4**, and **6** indeed show reductions in vividness and/or emotionality following dual-task interventions, and this strengthens the idea that competition in working memory leads to these effects. **Chapters 2** and **5** do not show these results for vividness or emotionality, but lack of effects for vividness and emotionality may have been the result of how these studies were set up: cued recall for multiple memories (**chapter 2**), or the absence of a standard recall only control condition (**chapter 5**).

Chapter 3 and **chapter 4** provide further evidence for a working memory explanation for the effect of dual-task interventions by independently showing that a higher speed of eye movements during recall of an aversive memory is more effective in reducing ratings of vividness and/or emotionality. An earlier study previously manipulated speed of eye movements and also showed this effect (Maxfield, Melnyk, & Hayman, 2008). Moreover, the results from these studies are in line with the finding that passively listening to beeps during recall of an aversive memory is less taxing and therefore less effective, than actively making eye movements during memory recall (van den Hout et al., 2011). Together, these studies consistently show that performing a more cognitively demanding dual-task during recall enlarges changes in vividness and emotionality of the recalled memory. At first sight, this seems contradictory with the idea of an inverted U-curve relationship between working memory taxation and changes in memory ratings (Engelhard, van den Hout, & Smeets, 2011;

Gunter & Bodner, 2008). According to working memory theory, which predicts an inverted U-curve, decreases in vividness and emotionality are largest when taxation is just right: not too much, but also not too little taxation. Our results seem to provide evidence for a linear relationship instead. However, this is true only on the premise that the highest speed of eye movements taxed working memory optimally (leaving some space for memory recall). Perhaps our most taxing condition hit peak value, leading to the largest decreases in vividness and emotionality, but even more taxing conditions may exist. For instance working memory could have been taxed more with additional tasks (in a different modality) on top of eye movements, for instance counting (e.g., Kemps & Tiggeman, 2008), which might then have resulted in smaller decreases in vividness and emotionality compared to our most taxing condition. Currently, the results from the studies in **chapter 3** and **4** are inconclusive about a linear or inverted U-curve relationship. Further research is necessary to elucidate the exact nature of the relationship between dual-tasks and changes in vividness/emotionality.

Dual-tasks and intrusive re-experiencing

The majority of research on the underlying mechanisms of EMDR has been focused on changes in self-reported vividness and/or emotionality, which have served as valuable variables that over the years have elucidated the process of working memory taxation (for an overview, see van den Hout & Engelhard, 2012). However, little work has been devoted to the search for treatment mechanisms and specifically how changes in vividness and/or emotionality cascade into changes in the manifestation of posttraumatic stress disorder (PTSD; Gunter & Bodner, 2009). One question we posed was if dual-task interventions affect (lab analogues of) intrusive re-experiencing of a traumatic memory. Research on patients with PTSD has shown that most intrusive memories are related to the memory hotspot (Holmes, Grey, & Young, 2005; Grey & Holmes, 2008), and since dual-tasks target the hotspot and reduce its vividness and/or emotionality (i.e., truncating the hotspot's excitable nature), this may subsequently reduce the number of intrusive memories.

In **chapter 5** we addressed whether dual-task interventions affect intrusive re-experiencing with two experiments, and our preliminary results showed that prolonged dual-task interventions (of approximately 10 minutes) may be able to reduce the number of intrusions, regardless of the modality of that intervention (visuospatial vs. verbal). Dual-task interventions of standard duration (approximately 3 minutes) were not able to modulate the number of intrusions. However, a – still to be performed – third experiment needs to provide a definitive answer on whether actual dual-tasking (i.e., recalling a memory while simultaneously performing a task) causes a reduction in intrusive re-experiencing or whether the memory recall component drives this effect. For now, the effect prolonged dual-task interventions had is in line with typical 10-minute single-task interventions within the trauma film paradigm, such as playing Tetris or finger tapping (e.g., Deeprose, Zhang, DeJong, Dalgleish, & Holmes, 2012; Holmes, James, Coode-Bate, & Deeprose, 2009; Holmes, James, Kilford, & Deeprose, 2010; James, Lau-Zhu, Tickle, Horsch, & Holmes, 2016b). If a third experiment confirms our results thus far, it would suggest that single or dual-task interventions following the trauma film need to be of a sufficient duration to modulate intrusions. Perhaps not surprising, because some research indeed suggests that dual-task interventions seem to have long-term (24h) effects when the intervention duration is lengthened (Leer, Engelhard, & van den Hout, 2014); though long-term (1 week) effects have been found with short duration interventions too (e.g., Gunter & Bodner, 2008). Interestingly, the results from the two experiments in this study, and the hypothesized results from third (to-be-performed) experiment, might then also suggest that simultaneous memory recall may be a superfluous element of the dual-task intervention. After all, interventions without and with simultaneous memory recall would then be able to reduce intrusions (e.g., Holmes et al., 2009; 2010, see **chapter 5**).

In **chapter 6** we tested whether continuous memory recall is necessary for a dual-task intervention to effectively modulate intrusions. We used an experiment that was similar in design to **chapter 5**. We compared playing Tetris with or without continuous memory recall, but failed to find effects on intrusive re-experiencing. This was unexpected especially for the Tetris *without* continuous memory recall intervention, which – supposedly – is rather effective in reducing intrusions (e.g., Hagenaars, Holmes, Klaassen, & Elzinga, 2017; Holmes et al., 2009; 2010; Horsch et al., 2017; Iyadurai et al., 2017; James et al., 2015). It was also unexpected for the Tetris *with* memory recall intervention that was similar to the prolonged dual-task interventions in **chapter 5** that reduced intrusions compared to no-task control. Do these findings represent unfortunate, but erroneous, null results, or are the effects of these intrusion-modulating interventions more volatile than we believe? A comprehensive review (yet not a meta-analytic review) shows that engaging in visuospatial tasks at or after encoding the trauma generally blocks intrusions (Brewin, 2014). Therefore, our results may represent an incorrect null finding. Note, however, that the majority of this work comes from the same research group. Alternatively, if the effects of these interventions are volatile, it requires replication and investigation into the exact parameters of the effect.

If dual-tasks truly modulate intrusions, a question remains via what mechanism they do. In **chapter 5** we observed drops in vividness or emotionality following the interventions, but these drops did not predict the number of intrusions in the following week. Interestingly, the number of intrusions was lower after a visuospatial and verbal dual-task compared to no-task control. This casts some doubt on the idea that there is a relationship between changes in vividness and/or emotionality and intrusions, though work on dual-tasks and intrusions is very preliminary at this point. Changes in vividness and emotionality (or subjective units of distress in clinical practice) are convenient and useful indices, but may not necessarily reflect therapeutic change. An alternative path is to look at more objective, and perhaps also more sensitive, indices of memory change, such as memory accessibility (van den Hout, Bartelski, & Engelhard, 2013), startle eye-blink (Engelhard, van Uijen, & van den Hout, 2010b), and heart rate (Kearns & Engelhard, 2015), which may better predict therapeutic outcomes.

Modality specific effects of dual-task interventions

One ongoing debate in the literature is whether an intervention needs to be taxing visuospatially (Andrade et al., 1997) or whether it should be taxing in general, regardless of its modality (Gunter & Bodner, 2008). The majority of studies in this thesis employed eye movement dual-task interventions and it is presumed that this primarily is a visuospatial intervention. The studies in **chapters 2, 3** and **4** exclusively used this intervention during memory recall and studies presented in the latter two chapters, indeed, found that it was effective in reducing ratings of vividness and/or emotionality. The work in **chapter 6** used a Tetris dual-task intervention and also found reductions in ratings of vividness. Reasonably, making eye movements or playing Tetris tax visuospatial working memory resources, however, it does not preclude possessing general taxation properties. A study that perhaps better, though not yet comprehensively, addresses this question is discussed in **chapter 5**. This study tentatively showed that prolonged dual-task interventions (compared to no-task control) reduce the number of intrusions regardless of whether it was visuospatial (eye movements) or a verbal dual-task (counting backwards).That either dual-task (regardless of modality) is effective in reducing intrusions compared to no-task control shows that intrusion modulation is dependent on general taxation. However, the verbal task was more taxing on working memory compared to a visuospatial task suggesting that modality specific taxation of the visuospatial task is also involved and that modality congruency may have improved the visuospatial task's efficiency[7]. Thus, this study shows that modality specific effects and general dual-task load can co-occur, but it does not provide an unequivocal answer to whether

[7] It has been argued that counting might require visual imagery by forming a mental number line (Restle, 1970), which according to theory would then constitute an intervention that is visuospatial in nature.

modality specific taxation is necessary, because we do not know whether the effect was due to the dual-task or exposure component. At this point, both accounts are supported with sufficient evidence (e.g., Andrade, Pears, May, & Kavanagh, 2012; May, Andrade, Panabokke, & Kavanagh, 2010; Engelhard et al., 2011; Tadmor, McNally, and Engelhard, 2017; van den Hout et al., 2010), and there are also studies showing evidence that general taxation and modality-specific taxation can co-occur (Baddeley & Andrade, 2000; Kemps & Tiggeman, 2007; Matthijssen, van Schie, & van den Hout, 2017). Possibly, certain types of tasks work better for certain types of memories. Craving for instance seems to be effectively reduced when using visuospatial dual-tasks only (e.g., Andrade et al., 2012; May et al., 2010). Though, in passing, it is arguable whether craving constitutes 'a memory'.

Memory Disruption and Updating During Reconsolidation

Next to dual-task interventions that tax working memory, unwanted memories may also be targeted by a different intervention. Behavioral new learning interventions during reconsolidation disrupt or update unwanted memories. When consolidated unwanted memories are retrieved they may, under certain conditions, enter a phase in which these memories are temporarily sensitive to disruption or updating before that are re-stored in memory (e.g., Ågren, 2014; Besnard, 2012; Elsey & Kindt, 2017; Schwabe, Nader, & Pruessner, 2014). By definition, two key components to memory change during reconsolidation are (1) retrieval that successfully reactivates/destabilizes the memory, and (2) an effective intervention that disrupts or updates the memory (Lee, Nader, & Schiller, 2017). In other words, memory disruption or updating should only occur in experimental conditions that satisfy these two conditions. It is important to note that memories can only be changed in the order of hours after destabilization, and not after that period.

In the studies described in **chapter 7** and in **chapter 8** we did not find that changes during reconsolidation were specific to conditions in which new learning followed reactivation. More specifically, we found that new learning, regardless of preceding reactivation, resulted in memory disruption and/or memory updating. In unison, these results point to a different mechanism of action: retroactive interference (e.g., McGeoch, 1942). Initially, we speculated that similarity in memory change between new learning with and without reactivation was the result of unintended reactivation in the no-reactivation + new learning conditions; opening the reconsolidation window in both these conditions. Though, this was an option, it seemed unlikely since the design of the studies in these chapters closely mirrored other experiments on reconsolidation that were successful in providing evidence for the reconsolidation hypothesis (e.g., Hupbach, Hardt, Gomez, & Nadel, 2008; Wichert, Wolf, & Schwabe, 2011, 2013a, 2013b).

Alternatively, lack of effects specific to the reactivation + new learning conditions might be the result of insufficient reactivation in these conditions; thereby not fulfilling one of two essential conditions to memory change during reconsolidation (Lee et al., 2017). It is possible that direct reactivation, which we used in our studies, produces smaller changes in memories compared to indirect reactivation, because direct one-time retrieval (in **chapter 7**) and retrieval practice (in **chapter 8**) may protect against interference (e.g., Potts, & Shanks, 2012; Roediger & Karpicke, 2006; though see Hupbach, 2015). However, others have argued that direct reactivation is necessary, and that indirect reactivation does not produce favorable results (Dębiec, Doyère, Nader, & LeDoux 2006). A recent meta-analysis concluded that changes during reconsolidation were not moderated by direct or indirect reactivation (Scully, Napper, & Hupbach, 2016). It is, of course, still possible that strength of reactivation is not solely determined by the indirect or direct dichotomy, but by other factors such as the materials used or the number of retrievals.

The number of studies that show evidence for reconsolidation largely outweigh those that do not; though an all-encompassing meta-analysis that would better substantiate that claim has not yet been performed (for reviews see Ågren, 2014; Beckers & Kindt, 2017; Besnard et al., 2012; Elsey & Kindt, 2017; Riccio, Millin, & Bogart, 2006; Schwabe et al., 2014; Treanor, Brown, Rissman, & Craske, 2016). However, taken together, the studies in **chapter 7** and **8** are difficult to reconcile with the idea of reconsolidation. They join other studies that have been unsuccessful in finding evidence in favor of the process of reconsolidation (e.g., Bos, Beckers, & Kindt, 2014; Golkar, Bellander, Olsson, & Öhman, 2012; Hardwicke, Taqi, & Shanks, 2016; Kindt & Soeter, 2013; Klucken et al., 2016; Nader & Einarsson, 2010; Soeter & Kindt, 2011; Schroyens, Beckers, & Kindt, 2017; Thome et al., 2016; Wichert et al., 2011;Wood et al., 2015). This suggests that reconsolidation is a delicate process that does not always take place. Indeed, a number of boundary conditions have been discovered, such as the memory's age or strength (e.g., Eisenberg, Kobilo, Berman, & Dudai, 2003; Wichert et al., 2011), or the type of reactivation (Dębiec et al., 2006). Other variables such as trait anxiety seem to moderate memory change during reconsolidation (Soeter & Kindt, 2013). It is currently unclear if the presence of a boundary condition completely restricts initiation of (memory change during) reconsolidation or simply makes it more difficult (Nader & Einarsson, 2010). Evidence suggest the latter since certain boundary conditions can be overcome given the right approach, such as delay in time (Wang, Alvares, & Nader, 2009), presenting new information (Winters, Tucci, & DaCosta-Furtado, 2009), or prolonged reactivation (Suzuki et al., 2004).

Failures to replicate cast doubt on (the robustness of) the process of reconsolidation, but are also fruitful learning opportunities allowing us to discover the boundaries and limitations of the process (Dunbar & Taylor; 2017; Lee et al., 2017). Indeed, extended replications of null-result studies, in which parameters are varied, may provide us with more information about the reconsolidation process, specifically when (memory change during) reconsolidation takes place. To best answer this question requires quantifiable indices (preferably those that can be used in clinical practice) that can help us in determining whether a memory is successfully (de)stabilized and when an intervention is supposed to be successful. Regarding the former, it has been argued that memory destabilization requires a situation in which there is a mismatch between what is expected and what actually happens (i.e., prediction error or expectancy violation; Lee, 2009). This mismatch signals that what was predicted, based on information stored in a consolidated memory, did not happen. Therefore, the prediction following from the consolidated memory needs updating with newly encountered information. Only then will change during reconsolidation take place. A number of studies, indeed, show that retrieval alone is not sufficient to induce destabilization, but that prediction error plays a role in updating memories during reconsolidation (Sevenster, Beckers, & Kindt, 2012, 2013, 2014; for overviews see Exton-McGuinness, Lee, & Reichelt, 2015; Fernández, Rodrigo, Boccia, & Pedreira, 2016).

The exact degree of prediction error may be quite subtle. If the degree of expectancy violation is too big a new memory trace will be formed; if there is no or too little violation the trace will not be destabilized and will simply be retrieved. Only moderate violation will result in destabilization of the memory (Beckers & Kind, 2017; Sevenster et al., 2014; Exton-McGuinness et al., 2015). Moreover, evidence for prediction error primarily comes from well-controlled lab experiments, in which participants acquire relatively simple, and personally irrelevant, threat memories (e.g., picture-shock) that are subsequently targeted with pharmacological agents (Sevenster et al., 2012, 2013). Envisioning prediction error in more complicated contexts, or other paradigms testing other types of memory, such as episodic memory in **chapters 7** and **8**, already is a puzzling endeavor. Translating the (preliminary) implications of this work on (the exact degree of) prediction error and reconsolidation to clinical practice is even more challenging. Therapists are faced with strongly consolidated and highly personal traumatic memories that would each require a personalized degree of prediction error (Elsey & Kindt, 2017; Krawczyk, Fernández, Pedreira, & Boccia, 2017).

What may provide a further challenge in translating the findings of reconsolidation research to clinical practice is the exact type of intervention(s) psychologists should use (and this thus requires further investigation). The majority of work on animal and human reconsolidation uses pharmacological interventions targeting the noradrenergic system (see Kroes, Schiller, LeDoux, & Phelps, 2015). Though these interventions are useful, the vast majority of psychological treatments of PTSD employ behavioral interventions (e.g., exposure; imagery rescripting; Foa & Kozak, 1986; Holmes, Arntz, & Smucker, 2007). In **chapters 7** and **8** we wanted to fundamentally test whether reconsolidation can be targeted with behavioral interventions, but there is no direct translation of our employed intervention to the clinic. A more suitable lab intervention, one that mimics exposure-based therapy, is a reactivation-extinction intervention (e.g., Schiller et al., 2010). This intervention consists of a brief reactivation of the consolidated memory (destabilization) followed by repeated exposure to the feared stimulus. This intervention has proven to be effective in changing threat memory during reconsolidation, but there are problems with replication (see Ågren, 2014). Moreover, the role of prediction error in combination with reactivation-extinction has never been examined.

Future studies on reconsolidation, specifically those employing behavioral interventions, should further unravel when (memory change during) reconsolidation is initiated, and when it is not (i.e., boundary conditions), and what can be done to overcome these conditions of non-change. Fundamental studies are necessary for gaining insights in the mechanisms, while translational studies are required to start bridging the gap between fundamental lab science and clinical practice.

Retrieval Suppression of Unwanted Intrusive Memories

The third intervention we examined is retrieval suppression or retrieval stopping. This ability allows individuals to limit the duration of unwelcome memories in mental awareness by preventing or interrupting the reflexive retrieval of these memories. Retrieval of unwelcome memories is specifically studied in situations when people are confronted with cues that remind them of these memories. This is especially relevant in the context of PTSD given that according to cognitive theories of PTSD cues that have become associated with the traumatic event (e.g., spatial cues, light patterns, shapes) regularly trigger retrieval of traumatic memories (e.g., Brewin, 2001; Dalgleish, 2004; Ehlers & Clark, 2000; Ehlers et al., 2002).

It has been argued that successful retrieval suppression might be dependent on the valence of what people try to suppress (e.g., Depue, Banich, & Curran, 2006; Marx, Marshall,

& Castro, 2008; Nørby, Lange, & Larsen, 2010) and on strategy that people employ (e.g., Hertel & Calcaterra, 2005; Hotta & Kawaguchi, 2009). There is considerable variability in how successful people are at suppressing negative materials. One possibility is that this variability depends on the strategy (direct suppression vs. thought-substitution) participants adopt for retrieval suppression. Thought-substitution strategies generally result in larger suppression-induced forgetting effects compared to direct suppression strategies. Arguably, direct suppression strategies to retrieval suppression may not be powerful enough to induce suppression-induced forgetting especially when negative unwanted memories need to be suppressed. In **chapter 9** we critically tested this idea and showed that with direct suppression negative memories can be suppressed without resorting to thought-substitution strategies. Moreover, suppression-induced forgetting was observed regardless of the memory's valence and the cue's valence (neutral vs. negative). In line with other studies (e.g., Bergström, de Fockert, & Richardson-Klavehn, 2009b; Benoit & Anderson, 2012), this study strengthens the idea that direct suppression can be used a strategy to avoid thinking of unwanted memories.

One could argue that suppression of 'unwanted' words lacks certain ecological validity and that directly suppressing these materials simply is possible, because words are easier to suppress then more ecologically valid materials, such as aversive pictures or film clips. A recent study by Küpper, Benoit, Dalgleish, and Anderson, (2014) addressed this issue and used aversive pictures from the International Affective Picture System (IAPS; Lang, Bradley, & Cuthbert, 2008) combined with naturalistic reminders. For instance, a picture of a coat hanger as cue for the target picture of soldiers raiding a room while a family hides in a closet. Küpper et al. (2014) showed that adopting direct suppression strategies also leads to forgetting of more ecologically valid materials. A study by Catarino, Küpper, Werner-Seidler, Dalgleish, and Anderson (2015) took this one step further and used these materials in patients with PTSD, and showed that patients with the greatest deficits in suppression-induced forgetting, were those with the highest severity of PTSD symptoms. Taken together, these studies on direct suppression support the idea that employing this strategy leads to forgetting regardless of the material one is trying to forget.

But why is suppression via direct suppression strategies important? The primary reason is that direct suppression (compared to thought-substitution) directly engages the core mechanism of inhibitory control that leads to forgetting of these unwanted memories (e.g., Anderson, 2003, 2005; Benoit & Anderson, 2012; Wang, Cao, Zhu, Cai, & Wu, 2015). Behaviorally, this can be observed by testing the target memory with a novel cue (e.g., Anderson & Green, 2001; Anderson et al., 2004; Bergström et al., 2009b; Wang et al., 2015). For instance, if a participant needed to suppress the target 'NUN' when seeing the cue 'BARREL', the novel, yet semantically (but not experimentally) associated, test cue could be 'CLOISTER'. If the target is indeed inhibited in memory, this novel test cue should also produce forgetting (i.e., cue-independence; Anderson & Spellman, 1995). However, sometimes these two strategies produce seemingly comparable patterns of voluntary memory recall; on a neural level they engage different brain structures to achieve this. When employing thought substitution a person uses retrieval processes and engages the hippocampus, while during direct suppression the hippocampus is disengaged, because it is inhibited by prefrontal areas (e.g., Anderson & Hanslmayr, 2014; Benoit & Anderson, 2012).

Because direct suppression engages inhibitory mechanisms, participants were instructed to also use this strategy in the study we described in **chapter 10**. This study extents earlier research on suppression-induced forgetting (and also the work described in **chapter 9**) by adding a more clinically relevant measure of intrusive memories (e.g., Levy & Anderson, 2012). The addition of this measure allowed us to examine whether retrieval suppression was able to downregulate the number of intrusive memories, and whether this depended on how long retrieval suppression needed to be applied. In line with earlier studies, we observed that retrieval suppression robustly reduces the number of intrusions with repeated suppression attempts (e.g., Benoit, Hulbert, Huddleston, & Anderson, 2014; Hellerstedt, Johansson, & Anderson, 2016; Levy & Anderson, 2012). Moreover, participants experienced more intrusions when they needed to suppress retrieval longer, and with progression of the total time suppressing participants start to show relapses in inhibitory control. These results support an idea proposed by Lee, Lee, and Tsai (2007) that suppression-induced forgetting is reduced when trials are longer, because of an increase in the probability of intrusions occurring. When suppression is employed in relatively short bursts, for instance in the study described in **Chapter 9**, suppression-induced forgetting effects may be bigger. Nevertheless, suppression-induced forgetting seems to be a rather robust phenomenon overall regardless of whether trials are relatively short (2s; Bergström et al., 2009a) or relatively long (8s; Wang et al., 2015, for a meta-analytic overview see Anderson & Huddleston, 2012, p62). It is, however, still possible that the valence of to-be-suppressed memories moderates the number of experienced intrusions on short and long trials.

Intrusions seem to play a pivotal role (in the think/no-think paradigm) in engaging inhibitory control. Evidence for this comes from neuroimaging and electrophysiological studies that showed that suppression-induced forgetting is more strongly associated with No-

Think trials in which the unwanted memory *briefly* intrudes compared to trials in which it does not intrude (e.g., Benoit et al., 2014; Levy & Anderson, 2012). Intrusions that enter awareness trigger inhibitory control processes that then rapidly remove these memories from awareness (Hellerstedt et al., 2016). This supports the idea that retrieval suppression interrupts the reflexive retrieval of unwanted memories. It is also worth noting that it seems that an intruding memory that is active in awareness is therefore more vulnerable to inhibitory processes (e.g., Detre, Natarajan, Gershman, & Norman, 2013), and that this process mirrors the reconsolidation process where only reactivated memories are sensitive to memory changing interventions (e.g., Nader, Schafe, & LeDoux, 2000).

The introduction of a method to measure intrusions makes the TNT paradigm more suitable to study intrusion-related clinical phenomena. However, key questions remain especially in the light of (future) clinical applicability. For instance, is retrieval suppression something that can be trained, and if so, is it a durable effect that can be measured in the long-term? A number of studies suggest that it can indeed be trained; with more suppression attempts the suppression-induced forgetting effect increases (e.g., Anderson & Green, 2001; Lambert, Good, & Kirk, 2010; Lee et al., 2007), and the number of intrusions decreases (e.g., Gagnepain, Hulbert, & Anderson, 2017; Levy & Anderson, 2012; see also **chapter 10**). However, studies that examined the durability of the effect have produced mixed results. Some found that suppression-induced forgetting is observed after 24 hours (Hotta & Kawaguchi, 2009), but not after 1 week (Nørby et al., 2010), while others found rebound effects after 1 week (Meier, Köning, Parak, & Henke, 2010), or even after 1 year (Noreen & MacLeod, 2014). Rebound effects show that participants actually remembered to-be-suppressed items better than baseline items (i.e., items that are neither recalled, nor suppressed), while it was expected that recall of to-be-suppressed items should be worse compared to baseline items. Interpretation of these results is complicated and genuine suppression-induced forgetting effects may be obscured by repeated testing (immediately after retrieval suppression and after a delayed interval), which slows the forgetting rate over longer intervals (e.g., Karpicke & Roediger, 2008). Clearly, a purer test of long-term effects of retrieval suppression is needed. Also, it is important to note that suppression-induced forgetting effects are created by a relatively short intervention (between 8 and 16 repetitions of 4s), and participants probably stop exerting inhibition after an immediate test, while in real cases retrieval suppression is most likely exerted repeatedly over prolonged periods of time, making generalization from lab to clinic difficult at this point.

Another question is whether (the training of) retrieval stopping, and thus forgetting and the downregulation of intrusions, can be successfully employed by anyone. Levy and Anderson (2008) suggest that there is high variability in the ability to exert inhibitory control. Indeed, deficits in retrieval suppression are associated with high trait anxiety and high brooding (Dieler, Hermann, & Fallgatter, 2013; Marzi, Regina, & Righi, 2014), and PTSD diagnosis (Catarino et al., 2015). This suggests impairments of inhibitory control in anxiety and trauma and stress related disorders (Nørby, 2017). It is, however, still unclear whether impaired inhibitory control is a risk factor for the development of these disorders, or whether the latter contributes to the former. Prospective longitudinal design studies with multiple measures of suppression in populations with high risk for trauma (e.g., soldiers, emergency personnel) might elucidate this question. Moreover, perhaps adjustments in retrieval suppression strategies may overcome initially observed deficits. A study by Murray, Anderson, and Kensinger (2015) showed that the elderly, who were hypothesized to have deficits in retrieval suppression, performed similarly to people in their twenties when given the right suppression strategies. Until more is known about the causal status of impaired inhibitory control and diagnosis, and whether adopting certain strategies can circumvent deficits in control, retrieval suppression may be best employed by those without deficits in ability to engage inhibitory control.

A Link Between Methods

In this thesis three independent interventions were examined that target unwanted memories. At this point, the studied method have yielded bits of information that have largely been studied in isolation from one another using different independent and dependent variables and different explanatory constructs[8]. Still they touch upon strongly related phenomena and concepts, such as unwanted memories, but also memory change, emotion regulation, and forgetting. The differences between these interventions may indeed give the impression that they are unique and independent from each other. Perhaps they are, but it is also possible that there is overlap, for instance, in the mechanisms these interventions tap into, or that there is interplay between these mechanisms. The best-case scenario would be to provide an integration of these different methods into a complete whole. However, a complete

[8] Barclay (1973) mentions that in psychology scientific progress has rapidly advanced. He argues that this process, known as the 'knowledge explosion', can be better described as an explosion in a confetti factory. Everything is covered with little bits of paper, but these are not whole sheets of paper any more. The whole was lost and researchers became interested in smaller and smaller bits of behavior and cognition (that are difficult to relate to each other).

integration may be difficult, because each method is more than an intervention; it is a separate paradigm, not only with accompanying theories, but also dependent variables, experimental designs, working mechanisms, etcetera. In a preliminary attempt to connect these methods, we suggest linking pins showing where methods overlap. These common areas may be fertile grounds for new research.

Regarding the different working mechanisms of the methods studied there is some overlap. For instance, van den Hout and Engelhard (2012) hypothesized that long-term effects of dual-task interventions involve reconsolidation mechanisms. How would this work exactly? Memory recall destabilizes the unwanted memory, and performing a dual-task (e.g., making eye movements, playing Tetris, counting) while simultaneously recalling a memory causes competition in working memory. This process impedes memory recall and as a consequence memory vividness and emotionality are reduced. After this process the (still) destabilized memory is reconsolidated into long-term memory. Preliminary evidence suggests that this indeed may be the case (Leer et al., 2014). However, in dual-task paradigms it is difficult to separately test the involvement and necessity of the two key components to memory change during reconsolidation, because retrieval (destabilization) and intervention always occur simultaneously.

Reconsolidation may also be involved in retrieval suppression (or at least echoes research on reconsolidation). Retrieval suppression seems to be more effectively exercised when unwanted memories briefly intruded into awareness (e.g., Benoit et al., 2014; Detre et al., 2013); awareness may be equated with working memory (Hellerstedt et al., 2016). This pattern suggests that (re)activation of an unwanted memory in working memory makes it more vulnerable to change, which mimics reconsolidation research where reactivation (initiating destabilization) frees the way for applying interventions (e.g., Nader et al., 2000). Drawing parallels between these mechanisms has to be done with caution, because typically in retrieval suppression research (including studies in the current thesis) study of (novel) materials, intervention, and test take place on the same day (e.g., Anderson & Green, 2001; Anderson et al., 2004). Therefore, most research likely affects (synaptic) consolidation (Dudai, 2004; Dudai & Morris, 2000), instead of reconsolidation processes (though for some preliminary work, see Noreen & MacLeod, 2014).

Inhibitory processes involved in retrieval suppression may also be related to working memory. In fact, the extent to which inhibitory control can be employed depends on an individual´s working memory capacity; individuals with high capacity display better inhibition compared to individuals with low capacity (e.g., Aslan & Bäuml, 2011; Redick, Calvo, Gay, &

Engle 2011; for an overview, see Redick, Heitz, & Engle, 2007). A key question concerns if inhibition is also involved in taxing working memory when performing dual-tasks. As we speculated in **chapter 3**, taxing working memory may be explained as continuously and rapidly switching task sets (e.g., Allport, Styles, & Hsieh, 1994; for an overview see Kiesel et al., 2010). It is possible that memory recall and making horizontal eye movements are not performed simultaneously, but are rapidly alternated. In literature on task switching it has been argued that a mechanism needs to be present that reduces activation of one task (e.g., memory recall), in order to successfully switch to another task (e.g. making eye movements), or vice versa. Inhibition has been put forward as a potential mechanism of action (for an overview see Koch, Gade, Schuch, & Philipp, 2010). When inhibition of the recalled memory persists over time it has become less accessible and less vivid, which would fit with current results on dual-task interventions (e.g., van den Hout et al., 2013; van den Hout & Engelhard, 2012).

The change all these methods aim to achieve also relates to another theoretical discussion (on mechanisms) that is not often addressed (see Tulving & Pearlstone, 1966). Is change achieved by altering the memory's *availability* or its *accessibility*? An alteration in availability reflects that (a part of) previously stored information is lost. Accessibility reflects how well one can retrieve information that is stored in the memory system. There is evidence that dual-task interventions not only reduce vividness and/or emotionality, but also result in changes in accessibility (e.g., van den Hout, Bartelski, & Engelhard, 2013; Leer et al., 2017). It is however unclear whether an apparent reduction in accessibility unambiguously points to changes in accessibility or to changes in availability (see Leer et al., 2017). Because reconsolidation processes are hypothesized to also be involved long-term memory change following dual-task interventions (van den Hout &Engelhard, 2012), the change may reflect reduced availability. Reconsolidation is assumed to alter consolidated memories and thus affects availability. In fact, one criterion of reconsolidation specifically states that memory change should not be attributable to temporarily reduced accessibility, which recovers with time (Lattal & Abel, 2004). However, it is currently unclear whether reconsolidation is an all-or-nothing process or whether memories can be attenuated instead of overwritten completely. In case of attenuation, memory recovery would still be possible (Kroes, Schiller, LeDoux, & Phelps, 2015). Retrieval suppression is argued to affect accessibility of memories (e.g., Anderson, 2003; Anderson & Hanslmayr, 2014), which is also evident from an effect that is known as 'release of inhibition', where suppressed, and thus forgotten, items become available again later (Bjork, 1989). Overall, memory change as reflected in reduced accessibility seems a common dominator, perhaps partly because demonstrating reduced availability remains a challenging endeavor. After all, absence of evidence is not evidence of absence.

Besides possibly interconnecting mechanisms, there also seem to be some convergence in how different methods assess memory change. Change is primarily assessed via self-report and behavioral indices. In dual-task paradigms participants report vividness and emotionality of autobiographical memories (e.g., Lee & Cuijpers, 2013). In reconsolidation paradigms (with fear or declarative memories) change is usually assessed by the number of remembered and forgotten items (e.g., Bos, Schuijer, Lodestijn, Beckers, & Kindt, 2014) or change psychophysiological variables such as skin conductance or startle (e.g., Sevenster et al., 2012). Akin to reconsolidation paradigms, change is frequently assessed by the number of forgotten items in retrieval suppression studies (see Anderson & Huddleston, 2012). Most work in reconsolidation and retrieval suppression studies is based on non-autobiographical memories, though there are some exceptions (e.g., Noreen & MacLeod, 2014; Schwabe & Wolf, 2009). Interestingly, change is assessed in these studies by the number of forgotten event details, while this has not been assessed (yet) in memories in dual-task paradigms. Lastly, use of measures that mimic clinical symptoms are becoming increasingly more common, specifically intrusive re-experiencing, in dual-task studies (see **chapter 5** and **6**), in reconsolidation studies (e.g., Hupbach, Gomez, Hardt, Nadel, 2007; Hupbach et al., 2008), and retrieval suppression studies (e.g., Benoit et al., 2014; Levy & Anderson, 2012; see **chapter 10**).

The link between the three methods may also be apparent in brain structures that are involved in the different interventions. Neuroimaging studies on retrieval suppression show that direct suppression engages the prefrontal cortex, which disengages the hippocampus limiting episodic retrieval (Benoit & Anderson, 2012). A recent study showed that a reduction of intrusive aversive memories is accompanied by a downregulation of the amygdala via prefrontal upregulation (e.g., Gagnepain et al., 2017). The amygdala has also been a main target site for interventions during the reconsolidation process (e.g., Nader et al., 2001). Attenuation of the amygdala during reconsolidation is linked to disrupting fear memories and therefore to decreased return of fear (Ågren et al., 2012). Preliminary evidence from dual-task studies suggests that (specifically deactivation of) the amygdala is the result of making eye movement as a dual-task intervention (Thomaes, Engelhard, Sijbrandij, Cath, & van den Heuvel, 2016; De Voogd, 2017).

At the very least, the common ground between the three methods examined and discussed in this thesis should show that they might be more interrelated than is evident at first sight. There is overlap in the working mechanisms, the theoretical concepts, and – in an increasing fashion – how memory change is measured. Homogenization of the (in)dependent variables and explanatory constructs could be further improved to enable direct comparisons of the effectiveness of these interventions.

Potential Clinical Implications

The current studies present relatively fundamental work aiming to contribute to the understanding of changing unwanted memories. Nevertheless, the topic studied is clinically highly relevant and some clinical implications can be formulated, some implications being straightforward, others being less so and some implications merely being suggestions for further study. There are two straightforward clinical implications that follow from the studies on taxing working memory by dual-task interventions: increasing speed of eye movements increases the intervention's effectiveness (**chapter 3** and **4**). These effects occur regardless of individual differences in vividness of the memory and working memory capacity of the individual. Hence, the intervention's effectiveness increases with dual-task interventions that are more taxing on working memory. This implication is in line with earlier findings that show that passively listening to beeps – a then frequently employed dual-task intervention in EMDR – is less effective and less taxing. Presumably, it is less effective, because it is less taxing than actively making eye movements (van den Hout et al., 2011). Secondly, the study in **chapter 5** tentatively showed that dual-tasks might also reduce intrusive re-experiencing (though an exposure control condition was missing). This, of course, by no means necessitates therapeutic changes, but it does show that the effects of dual-task interventions may extent from variables that are a reflection of the improvement process (such as vividness, emotionality or distress) to variables that actually reflect clinical improvement, such as less intrusive re-experiencing.

The idea of memory disruption and updating during reconsolidation has fostered great excitement and great expectations about its clinical applicability (e.g., Lane, Ryan, Nadel, & Greenberg, 2015; Parsons & Ressler, 2013; Pitman, 2011; Schwabe et al., 2014). Admittedly some have been successful in translating (mainly) pharmacological interventions to the clinic (e.g., Soeter & Kindt, 2015; Kindt & van Emmerik, 2016), but questions remain regarding the (short term) viability of disrupting reconsolidation in clinical (PTSD) patients, especially when using behavioral interventions (e.g., Beckers & Kindt, 2017; Treanor et al., 2017). The studies in this thesis suggest that we may need to take a step back and examine, possibly in lab environments, when memory change during reconsolidation occurs. We need to parametrically vary features that may constrain memory change, and specifically conditions that are relevant to the clinical populations we intend to treat, such as memory strength (Suzuki et al., 2004), trait anxiety (Soeter & Kindt, 2013), and stress (Hoffman et al., 2015).

Less straightforward are the implications following the results on retrieval suppression. The studies in this thesis showed that by using direct suppression, as an approach to engage in retrieval suppression, unwanted memories can be forgotten and intrusions can be downregulated. Though these results may seem applicable to clinical practice, it seems unrecommendable to translate this into clinical interventions due to deficits in inhibitory control in patients with PTSD and in highly anxious individuals (e.g., Catarino et al., 2015; Marzi et al., 2014). Moreover, we observed in **chapter 10** that participants eventually become fatigued at engaging in inhibitory control and therefore they became less successful in suppressing intrusions. That fatigue affects inhibition is in accordance with literature that shows that sleep-deprived individuals experience difficulty with inhibition in general (e.g., Drummond, Paulus, & Tapert, 2006; Kato, Endo, & Kizuka, 2009). In fact, sleep disorders are prevalent in psychiatric conditions (Wulff, Gatti, Wettstein, & Foster, 2010), and disturbed sleep is related to PTSD (Germain, 2013). This may indicate that engaging in retrieval suppression by patients, especially those who are fatigued, should not be recommended.

Conclusion

To conclude, this thesis shows that unwanted memories can be changed, and that different interventions can be employed to achieve this goal. The strongest evidence comes from subjective reports of memory change after dual-task interventions that tax working memory and retrieval suppression interventions. The results of dual-task interventions are especially promising and support clear clinical implications (e.g., higher speeds of eye movements during EMDR). Regarding behavioral interventions during reconsolidation, our work is contradictory with a large body of evidence that does support memory change during reconsolidation. Though, behavioral interventions during reconsolidation may hold implications for clinical practice, a fundamental understanding of when (memory change during) reconsolidation is initiated, when it is not, and how non-change can be overcome is now crucial. Clinical applicability of retrieval suppression interventions may be limited as well, because of questions about trainability, durability and whether this is an effective intervention for everyone. Nevertheless, these interventions hold clear implications for future research, which may in the end eventually lead to new suggestions for improving treatment effectiveness.

CHAPTER 12

REFERENCES

References

Ågren, T. (2014). Human reconsolidation: A reactivation and update. *Brain Research Bulletin, 105*, 70-82. doi:10.1016/j.brainresbull.2013.12.010

Ågren, T., Engman, J., Frick, A., Björkstrand, J., Larsson, E. M., Furmark, T., & Fredrikson, M. (2012). Disruption of reconsolidation erases a fear memory trace in the human amygdala. *Science, 337*(6101), 1550-1552.

Aldao, A., Nolen-Hoeksema, S., & Schweizer, S. (2010). Emotion-regulation strategies across psychopathology: A meta-analytic review. *Clinical psychology review, 30*(2), 217-237.

Allport, D. A., Styles, E. A., & Hsieh, S. (1994). Shifting intentional set: Exploring the dynamic control tasks. In C. Umilta & M. Moscovitch (Eds.), *Attention and performance XV* (pp. 421–452). Cambridge, MA: MIT Press.

American Psychiatric Association. (2000). Diagnostic and statistical manual, 4th edn, Text Revision (DSM-IV-TR). *American Psychiatric Association, Washington*.

American Psychiatric Association. (2004). Practice guideline for the treatment of patients with acute stress disorder and post traumatic stress disorder. Arlington, VA: American Psychiatric Association Practice Guidelines.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (DSM-5®). American Psychiatric Pub.

Anderson, M. C. (2003). Rethinking interference theory: Executive control and the mechanisms of forgetting. *Journal of memory and language*, *49*(4), 415-445.

Anderson, M. C., & Green, C. (2001). Suppressing unwanted memories by executive control. *Nature, 410*, 366-369. doi:10.1038/35066572

Anderson, M. C., & Hanslmayr, S. (2014). Neural mechanisms of motivated forgetting. *Trends in cognitive sciences, 18*(6), 279-292. doi:10.1016/j.tics.2014.03.002

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: Toward a Reconciliation of the Debate. Vol. 58: Nebraska Symposium on Motivation*, doi: 10.1007/978-1-4614-1195-6_8, (pp. 53-120). New York: Springer.

Anderson, M. C., & Levy, B. J. (2009). Suppressing unwanted memories. *Current Directions in Psychological Science*, *18*(4), 189-194.

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*(5655), 232-235.doi:10.1126/science.108950

Anderson, M. C., Reinholz, J., Kuhl, B. A., & Mayr, U. (2011). Intentional suppression of unwanted memories grows more difficult as we age. *Psychology and aging*, *26*(2), 397.

Anderson, M. C., & Spellman, B. A. (1995). On the status of inhibitory mechanisms in cognition: memory retrieval as a model case. *Psychological review*, *102*(1), 68.

Anderson, M. C., & Weaver, C. (2009). Inhibitory control over action and memory. *The new encyclopedia of neuroscience*, 153-163.

Andrade, J., Kavanagh, D., & Baddeley, A. (1997). Eye-movements and visual imagery: A working memory approach to the treatment of post-traumatic stress disorder. *British Journal of Clinical Psychology, 36*, 209-223. doi:10.1111/j.2044-8260.1997.tb01408.x

Andrade, J., Pears, S., May, J., & Kavanagh, D. J. (2012). Use of a clay modeling task to reduce chocolate craving. *Appetite, 58*(3), 955-963. doi:10.1016/j.appet.2012.02.044

Arch, J. J., & Craske, M. G. (2009). First-line treatment: A critical appraisal of cognitive behavioral therapy developments and alternatives. *Psychiatric Clinics of North America, 32*(3), 525-547.

Arnaudova, I., & Hagenaars, M. A. (2017). Lights… action: Comparison of trauma films for use in the trauma film paradigm. *Behaviour Research and Therapy*. doi: 10.1016/j.brat.2017.02.007

Arntz, A., Tiesema, M., & Kindt, M. (2007). Treatment of PTSD: A comparison of imaginal exposure with and without imagery rescripting. *Journal of behavior therapy and experimental psychiatry, 38*(4), 345-370.

Asendorpf, J. B., Conner, M., De Fruyt, F., De Houwer, J., Denissen, J. J., Fiedler, K., ... & Perugini, M. (2013). Recommendations for increasing replicability in psychology. *European Journal of Personality, 27*(2), 108-119. doi: 10.1002/per.1919

Aslan, A., & Bäuml, K. H. T. (2011). Individual differences in working memory capacity predict retrieval-induced forgetting. Journal of Experimental Psychology: Learning, Memory, and Cognition, 37(1), 264.

Asselbergs, J., Sijbrandij, M., E. Hoogendoorn., P. Cuijpers., L. Olie., K. Oved., J. Merkies., T. Plooijer., S. Eltink., & Riper, H. (2017). Development and Testing of TraumaGameplay: An Iterative Experimental Approach Using the Trauma Film Paradigm. Manuscript submitted for publication

Baddeley, A. (2012). Working memory: theories, models, and controversies. *Annual review of psychology, 63,* 1-29.

Baddeley, A. D., & Andrade, J. (2000). Working memory and the vividness of imagery. *Journal of Experimental Psychology-General, 129*(1), 126-145.

Baddeley, A. D., & Hitch, G. (1974). Working memory. *Psychology of learning and motivation*, *8*, 47-89.

Barclay, A. M. (1973). Death and rebirth in psychology. *Psyccritiques*, *18*(7), 333-334.

Barrowcliff, A. L., Gray, N. S., Freeman, T. C., & MacCulloch, M. J. (2004). Eye-movements reduce the vividness, emotional valence and electrodermal arousal associated with negative autobiographical memories. *Journal of Forensic Psychiatry & Psychology*, *15*(2), 325-345. doi:10.1080/14789940410001673042

Bechara, A., Tranel, D., Damasio, H., Adolphs, R., Rockland, C., & Damasio, A. R. (1995). Double dissociation of conditioning and declarative knowledge relative to the amygdala and hippocampus in humans. *Science 269*, 1115–1118. doi:10.1126/science.7652558

Beck, A. T., Steer, R. A., & Brown, G. K. (1996). Beck depression inventory-II. *San Antonio*, *78*(2), 490-8.

Beckers, T., & Kindt, M. (2017). Memory Reconsolidation Interference as an Emerging Treatment for Emotional Disorders: Strengths, Limitations, Challenges and Opportunities. *Annual Review of Clinical Psychology, 13*(1).

Benoit, R. G., & Anderson, M. C. (2012). Opposing Mechanisms Support the Voluntary Forgetting of Unwanted Memories. *Neuron 76*, 450-460. doi:10.1016/j.neuron.2012.07.025.

Benoit, R. G., Hulbert, J. C., Huddleston, E., & Anderson, M. C. (2014). Adaptive top-down suppression of hippocampal activity and the purging of intrusive memories from consciousness. *Journal of Cognitive Neuroscience, in press*. doi:10.1162/jocn_a_00696

Berntsen, D. (1996). Involuntary autobiographical memories. *Applied Cognitive Psychology, 10*, 435-454. doi: 10.1002/(SICI)1099-0720(199610)10:5<435::AID-ACP408>3.0.CO;2-L

Berntsen, D. (2010). The unbidden past: involuntary autobiographical memories as a basic mode of remembering. *Current Directions in Psychological Science*, *19*(3), 138-142. doi: 10.1177/0963721410370301

Bergström, Z. M., de Fockert, J. W., & Richardson-Klavehn, A. (2009*a*). Event-related potential evidence that automatic recollection can be voluntarily avoided. *Journal of Cognitive Neuroscience, 21*(7), 1280-1301. doi:10.1162/jocn.2009.21075

Bergström, Z. M., de Fockert, J. W., & Richardson-Klavehn, A. (2009b). ERP and behavioural evidence for direct suppression of unwanted memories. *NeuroImage, 48*(4), 726-737.doi:10.1016/j.neuroimage.2009.06.051

Berntsen, D. (1998). Voluntary and involuntary access to autobiographical memory. *Memory,* 6(2), 113-141.

Besnard, A., Caboche, J., & Laroche, S. (2012). Reconsolidation of memory: a decade of debate. *Progress in Neurobiology, 99*(1), 61-80. doi:10.1016/j.pneurobio.2012.07.002

Bisson, J.I., Ehlers, A., Mathews, A., Pilling, S., Richards, D, & Turner, S (2007). Psychological treatments for chronic post-traumatic stress disorder: systematic review and meta-analysis. *British Journal of Psychiatry, 190*, 97-104. doi:10.1192/bjp.bp.106.021402

Bjork, R. A. (1989). Retrieval inhibition as an adaptive mechanism in human memory. *Varieties of memory & consciousness*, 309-330.

Boeschoten, M.A., Bakker, A., Jongedijk, R.A. & Olff, M. (2014). PTSD Checklist for DSM-5– Nederlandstalige versie. Uitgave: Arq Psychotrauma Expert Groep, Diemen.

Bos, M. G. N., Beckers, T., & Kindt, M. (2014). Noradrenergic blockade of memory reconsolidation: a failure to reduce conditioned fear responding. *Frontiers in behavioral neuroscience, 8*, 412. doi:10.3389/fnbeh.2014.00412

Bos, M. G., Schuijer, J., Lodestijn, F., Beckers, T., & Kindt, M. (2014). Stress enhances reconsolidation of declarative memory. *Psychoneuroendocrinology, 46*, 102-113.

Bourne, C., Frasquilho, F., Roth, A. D., & Holmes, E. A. (2010). Is it mere distraction? Peri-traumatic verbal tasks can increase analogue flashbacks but reduce voluntary memory performance. *Journal of behavior therapy and experimental psychiatry, 41*(3), 316-324. doi: 10.1016/j.jbtep.2010.03.001

Bower, G. H., & Clapper, J. P. (1989). Experimental methods in cognitive science. In M. I. Posner (Ed.), Foundations of cognitive science (pp. 245–300). Cambridge: MIT Press.

Bradley, R., Greene, J., Russ, E., Dutra, L., & Westen, D. (2005). A multidimensional meta-analysis of psychotherapy for PTSD. *American Journal of Psychiatry, 162*, 214-227

Bradley, R. J., Moulin, C. J. A., & Kvavilashvli, L. (2013). Involuntary autobiographical memories. *The Psychologist, 26*, 190–193.

Breslau, N. (2009). The epidemiology of trauma, PTSD, and other posttrauma disorders. *Trauma, Violence, & Abuse, 10*(3), 198-210. doi:10.1177/1524838009334448

Breslau, N., Kessler, R. C., Chilcoat, H. D., Schultz, L. R., Davis, G. C., & Andreski, P. (1998). Trauma and posttraumatic stress disorder in the community: the 1996 Detroit Area Survey of Trauma. *Archives of general psychiatry, 55*(7), 626-632. doi:10.1001/archpsyc.55.7.626

Brewin, C. R. (2001). A cognitive neuroscience account of posttraumatic stress disorder and its treatment. *Behavior Research and Therapy, 39*, 373–393. doi: 10.1016/S0005-7967(00)00087-5

Brewin, C. R. (2014). Episodic memory, perceptual memory, and their interaction: Foundations for a theory of posttraumatic stress disorder. *Psychological Bulletin, 140*(1), 69. doi: 10.1037/a0033722

Brewin, C. R., & Burgess, N. (2014). Contextualisation in the revised dual representation theory of PTSD: A response to Pearson and colleagues. *Journal of behavior therapy and experimental psychiatry, 45*(1), 217-219. doi:10.1016/j.jbtep.2013.07.011

Brewin, C. R., Dalgleish, T., & Joseph, S. (1996). A dual representation theory of posttraumatic stress disorder. *Psychological review, 103*(4), 670. doi:10.1037/0033-295X.103.4.670

Brewin, C. R., & Holmes, E. A. (2003). Psychological theories of posttraumatic stress disorder. *Clinical psychology review, 23*(3), 339-376. doi:10.1016/S0272-7358(03)00033-3

Brewin, C. R., & Smart, L. (2005). Working memory capacity and suppression of intrusive thoughts. *Journal of behavior therapy and experimental psychiatry, 36*(1), 61-68. doi:10.1016/j.jbtep.2004.11.006

Brunet, A., Orr, S. P., Tremblay, J., Robertson, K., Nader, K., & Pitman, R. K. (2008). Effect of post-retrieval propranolol on psychophysiologic responding during subsequent script-driven traumatic imagery in post-traumatic stress disorder. *Journal of Psychiatric Research, 42*(6), 503-506. doi:10.1016/j.jpsychires.2007.05.006

Catarino, A., Küpper, C. S., Werner-Seidler, A., Dalgleish, T., & Anderson, M. C. (2015). Failing to Forget: Inhibitory-Control Deficits Compromise Memory Suppression in Posttraumatic Stress Disorder. *Psychological Science*, 1-13. doi:10.1177/0956797615569889

Chen, C., Liu, C., Huang, R., Cheng, D., Wu, H., Xu, P., ... & Luo, Y. J. (2012). Suppression of aversive memories associates with changes in early and late stages of neurocognitive processing. *Neuropsychologia, 50*(12), 2839-2848. doi:10.1016/j.neuropsychologia.2012.08.004

Chuah, Y. M., Venkatraman, V., Dinges, D. F., & Chee, M. W. (2006). The neural basis of interindividual variability in inhibitory efficiency after sleep deprivation. *Journal of Neuroscience 26*, 7156–7162. doi:10.1523/JNEUROSCI.0906-06.2006

Clark, I. A., Mackay, C. E., & Holmes, E. A. (2013). Positive involuntary autobiographical memories: You first have to live them. *Consciousness and cognition, 22*(2), 402-406. doi:10.1016/j.concog.2013.01.008

Conway, A. R. A., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005).Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin and Review, 12*, 769–786. doi:10.3758/BF03196772

Corallo, G., Sackur, J., Dehaene, S., & Sigman, M. (2008). Limits on Introspection: Distorted Subjective Time During the Dual-Task Bottleneck. *Psychological Science, 19*(11), 1110-1117. doi:10.1111/j.1467-9280.2008.02211.x

Craske, M. G., Treanor, M., Conway, C. C., Zbozinek, T., & Vervliet, B. (2014). Maximizing exposure therapy: an inhibitory learning approach. *Behaviour research and therapy*, *58*, 10-23.

Cuperus, A. A., Laken, M., van den Hout, M. A., & Engelhard, I. M. (2016). Degrading emotional memories induced by a virtual reality paradigm. *Journal of behavior therapy and experimental psychiatry, 52*, 45-50. doi:10.1016/j.jbtep.2016.03.004

Dalgleish T. (2004). Cognitive approaches to posttraumatic stress disorder: the evolution of multirepresentational theorizing. *Psychological. Bulletin, 130*(2), 228–260. doi: 10.1037/0033-2909.130.2.228

Daneman, M., & Carpenter, P.A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior, 19,* 450–466. doi:10.1016/S0022-5371(80)90312-6

Davidson, P. R., & Parker, K. C. (2001). Eye movement desensitization and reprocessing (EMDR): a meta-analysis. *Journal of Consulting and Clinical Psychology, 69*, 305-316.

Deacon, B.J., & Abramowitz, J.S. (2004). Cognitive and behavioral treatments for anxiety disorders: A review of meta-analytic findings. *Journal of Clinical Psychology, 60*, 429

Dębiec, J. (2012). Memory reconsolidation processes and posttraumatic stress disorder: promises and challenges of translational research. *Biological Psychiatry, 71*(4), 284-285.

Dębiec, J., Doyère, V., Nader, K., & LeDoux, J. E. (2006). Directly reactivated, but not indirectly reactivated, memories undergo reconsolidation in the amygdala. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(9), 3428-3433.

Deeprose, C., Zhang, S., DeJong, H., Dalgleish, T., & Holmes, E. A. (2012). Imagery in the aftermath of viewing a traumatic film: Using cognitive tasks to modulate the development of involuntary memory. *Journal of behavior therapy and experimental psychiatry, 43*(2), 758-764. doi:10.1016/j.jbtep.2011.10.008

de Jongh, A., & ten Broeke, E. (2012). Handboek EMDR: Een geprotocolleerde behandelmethode voor de gevolgen van psychotrauma [EMDR handbook: A protocol treatment for the effects of psychological trauma]. Amsterdam: Pearson.

De Voogd, L. D. (2017). *Do you remember what you did last summer? On the formation and alteration of memories for stressful experiences* (Doctoral thesis, Radboud University Nijmegen, Nijmegen, the Netherlands). Retrieved from http://repository.ubn.ru.nl/handle/2066/169274 de Vries, G. J., & Olff, M. (2009). The lifetime prevalence of traumatic events and posttraumatic stress disorder in the Netherlands. *Journal of traumatic stress*, *22*(4), 259-267. doi:10.1002/jts.20429

Dek, E. C., van den Hout, M. A., Giele, C. L., & Engelhard, I. M. (2010). Repeated checking causes distrust in memory but not in attention and perception. *Behaviour research and therapy, 48*(7), 580-587. doi:10.1016/j.brat.2010.03.009

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(5),* 441-447. doi:10.1111/j.1467-9280.2006.01725.x

Depue, B. E., Burgess, G. C., Willcutt, E. G., Ruzic, L., & Banich, M. T. (2010). Inhibitory control of memory retrieval and motor processing associated with the right lateral prefrontal cortex: evidence from deficits in individuals with ADHD. *Neuropsychologia*, *48*(13), 3909-3917.

Depue, B. E., Curran, T., & Banich, M. T. (2007). Prefrontal regions orchestrate suppression of emotional memories via a two-phase process. *Science, 317*(5835), 215-219. doi:10.1126/science.1139560

Depue, B. E., Orr, J. M., Smolker, H. R., Naaz, F., & Banich, M. T. (2015). The organization of right prefrontal networks reveals common mechanisms of inhibitory regulation across cognitive, emotional, and motor processes. *Cerebral Cortex*, doi: 10.1093/cercor/bhu324

deRoon-Cassini, T. A., Mancini, A. D., Rusch, M. D., & Bonanno, G. A. (2010) Psychopathology and resilience following traumatic injury: a latent growth mixture model analysis. *Rehabilitation psychology, 55*(1), 1. doi:10.1037/a0018601

DeStefano, D., & LeFevre, J. A. (2004). The role of working memory in mental arithmetic. *European Journal of Cognitive Psychology, 16*(3), 353-386. doi:10.1080/09541440244000328

Detre, G. J., Natarajan, A., Gershman, S. J., & Norman, K. A. (2013). Moderate levels of activation lead to forgetting in the think/no-think paradigm. *Neuropsychologia*, *51*(12), 2371-2388.

de Vries, G. J., & Olff, M. (2009). The lifetime prevalence of traumatic events and posttraumatic stress disorder in the Netherlands. *Journal of traumatic stress, 22*(4), 259-267.

Dieler, A. C., Herrmann, M. J., & Fallgatter, A. J. (2014). Voluntary suppression of thoughts is influenced by anxious and ruminative tendencies in healthy volunteers. *Memory*, *22*(3), 184-193.

Dienes, Z. (2016). How Bayes factors change scientific practice. *Journal of Mathematical Psychology, 72*, 78-89. doi: 0.1016/j.jmp.2015.10.003

Dolan, R. J. (2002). Emotion, cognition, and behavior. S*cience, 298*(5596), 1191-1194.

Donovan, J. J., & Radosevich, D. J. (1999). A meta-analytic review of the distribution of practice effect: Now you see it, now you don't. *Journal of Applied Psychology, 84*(5), 795. doi:10.1037/0021-9010.84.5.795

Drummond, S. P., Paulus, M. P., Tapert, S. F. (2006) Effects of two nights sleep deprivation and two nights recovery sleep on response inhibition. *Journal of Sleep Research, 15*, 261–265. doi:10.1111/j.1365-2869.2006.00535.x

Dudai, Y. (2004). The neurobiology of consolidations, or, how stable is the engram?. *Annual Review of Psychology, 55*, 51-86. doi: 10.1146/annurev.psych.55.090902.142050

Dudai, Y. (2006). Reconsolidation: the advantage of being refocused. *Current opinion in neurobiology*, *16*(2), 174-178.

Dudai, Y., & Morris, R. G. (2000). To consolidate or not to consolidate: what are the questions. *Brain, perception, memory. Advances in cognitive sciences*, 149-162.

Dunbar, A. B., & Taylor, J. R. (2017). Reconsolidation and psychopathology: Moving towards reconsolidation-based treatments. *Neurobiology of learning and memory*, *142*, 162-171.

Durham, R. C., Higgins, C., Chambers, J. A., Swan, J. S., & Dow, M. G. (2012). Long-term outcome of eight clinical trials of CBT for anxiety disorders: symptom profile of sustained recovery and treatment-resistant groups. *Journal of affective disorders, 136*(3), 875-881.

Dutch National Steering Committee Guidelines Mental Health Care (2003).Multidisciplinary Guideline Anxiety Disorders. Quality Institute Heath Care CBO/Trimbos Institute. Utrecht, Netherlands.

Duvarci, S., Nader, K., & LeDoux, J. E. (2008). De novo mRNA synthesis is required for both consolidation and reconsolidation of fear memories in the amygdala. *Learning & Memory, 15*(10), 747-755. doi:10.1101/lm.1027208

Ehlers, A., & Clark, D. M. (2000). A cognitive model of posttraumatic stress disorder. *Behaviour Research and Therapy, 38*, 319-345. doi: 10.1016/S0005-7967(99)00123-0

Ehlers, A., Hackmann, A., Steil, R., Clohessy, S., Wenninger, K., & Winter, H. (2002). The nature of intrusive memories after trauma: the warning signal hypothesis. *Behaviour Research and Therapy, 40*, 995-1002. doi: 10.1016/S0005-7967(01)00077-8

Eisenberg, M., Kobilo, T., Berman, D. E., & Dudai, Y. (2003). Stability of retrieved memory: inverse correlation with trace dominance. *Science, 301*(5636), 1102-1104. doi: 10.1126/science.1086881

Eldridge, L. L., Sarfatti, S., & Knowlton, B. J. (2002). The effect of testing procedure on remember-know judgments. *Psychonomic Bulletin & Review, 9*(1), 139-145. doi: 10.3758/BF03196270

Elsey, J. W., & Kindt, M. (2017*a*). Tackling maladaptive memories through reconsolidation: From neural to clinical science. *Neurobiology of Learning and Memory*, *142*, 108-117.

Elsey, J. W., & Kindt, M. (2017b). Breaking boundaries: optimizing reconsolidation-based interventions for strong and old memories. Learning & Memory, 24(9), 472-479.

Engelhard, I. M. (2012). Making science work in mental health care. *European journal of psychotraumatology, 3:1,* 18740.

Engelhard, I.M., Sijbrandij, M., van den Hout, M.A., Rutherford, N.M., Rahim, H.F., & Koçak, F. (2012). Choking under pressure: Degrading flashforwards related to performance anxiety. *Journal of Experimental Psychopathology, 3*, 158-67. doi:10.5127/jep.024111

Engelhard, I. M., van den Hout, M. A., Arntz, A., & McNally, R. J. (2002). A longitudinal study of "intrusion-based reasoning" and posttraumatic stress disorder after exposure to a train disaster. *Behaviour Research and Therapy, 40*(12), 1415-1424. doi:10.1016/S0005-7967(02)00018-9

Engelhard, I. M., van den Hout, M. A., Dek, E. C. P., Giele, C. L., van der Wielen, J. W., Reijnen, M., & van Roij, B. (2011). Reducing vividness and emotional intensity of recurrent "flashforwards" by taxing working memory: An analogue study. *Journal of Anxiety Disorders, 25*, 599-603. doi:10.1016/j.janxdis.2011.01.009

Engelhard, I. M., van den Hout, M. A., Janssen, W. C., & van der Beek, J. (2010*a*). Eye movements reduce vividness and emotionality of "flashforwards". *Behaviour Research and Therapy*, *48*(5), 442-447. doi:10.1016/j.brat.2010.01.003

Engelhard, I. M., van den Hout, M. A., Smeets, M. A. M. (2011*a*). Taxing working memory reduces vividness and emotional intensity of images about the queen's day tragedy. *Journal of Behavior Therapy and Experimental Psychiatry, 42*, 32–37. doi:10.1016/j.jbtep.2010.09.004

Engelhard, I. M., van den Hout, M. A., & Smeets, M. A. (2011*b*). Taxing working memory reduces vividness and emotional intensity of images about the Queen's Day tragedy. *Journal of Behavior Therapy and Experimental Psychiatry, 42*(1), 32-37. doi:10.1016/j.jbtep.2010.09.004

Engelhard, I. M., van Uijen, S. L., & van den Hout, M. A. (2010*b*). The impact of taxing working memory on negative and positive memories. *European Journal of Psychotraumatology, 1*: 5623, 1-8. doi: 10.3402/ejpt.v1i0.5623

Exton-McGuinness, M. T., Lee, J. L., & Reichelt, A. C. (2015). Updating memories—the role of prediction errors in memory reconsolidation. *Behavioural brain research*, *278*, 375-384.

Fawcett, J. M., Benoit, R. G., Gagnepain, P., Salman, A., Bartholdy, S., Bradley, C., ... & Anderson, M. C. (2015). The origins of repetitive thought in rumination: Separating cognitive style from deficits in inhibitory control over memory. *Journal of behavior therapy and experimental psychiatry*, *47*, 1-8. doi:10.1016/j.jbtep.2014.10.009

Fernández, R. S., Boccia, M. M., & Pedreira, M. E. (2016). The fate of memory: reconsolidation and the case of prediction error. *Neuroscience & Biobehavioral Reviews*, *68*, 423-441.

Fisher, R. A. (1935). The logic of inductive inference. *Journal of the Royal Statistical Society, 98*(1), *39–82*.

Foa, E.B., *The posttraumatic diagnostic scale (PDS)* manual 1995, Minneapolis, MN: National Computer Systems.

Foa, E. B., & Kozak, M. J. (1986). Emotional processing of fear: exposure to corrective information. *Psychological Bulletin*, *99*(1), 20.

Foa, E. B., Steketee, G., & Rothbaum, B. O. (1989). Behavioral/cognitive conceptualization of post-traumatic stress disorder. *Behavior Therapy, 20*, 155–176. doi: 10.1016/S0005-7894(89)80067-X

Forcato, C., Burgos, V. L., Argibay, P. F., Molina, V. A., Pedreira, M. E., & Maldonado, H. (2007). Reconsolidation of declarative memory in humans. *Learning & Memory, 14*(4), 295-303. doi:10.1101/lm.486107

Frans, Ö., Rimmö, P. A., Åberg, L., & Fredrikson, M. (2005). Trauma exposure and post-traumatic stress disorder in the general population. *Acta Psychiatrica Scandinavica, 111*(4), 291-290.

Gagnepain, P., Henson, R. N., & Anderson, M. C. (2014). Suppressing unwanted memories reduces their unconscious influence via targeted cortical inhibition. *Proceedings of the National Academy of Sciences*, *111*(13), E1310-E1319. doi:10.1073/pnas.1311468111

Gagnepain, P., Hulbert, J., & Anderson, M. C. (2017). Parallel Regulation of Memory and Emotion Supports the Suppression of Intrusive Memories. *Journal of Neuroscience*, 2732-16.

Galea, S., Uddin, M., & Koenen, K. (2011). The urban environment and mental disorders: epigenetic links. *Epigenetics, 6*(4), 400-404.

12

Geraerts, E., Merckelbach, H., Jelicic, M., & Smeets, E. (2006). Long term consequences of suppression of intrusive anxious thoughts and repressive coping. *Behaviour Research and Therapy, 44*(10), 1451-1460. doi: 10.1016/j.brat.2005.11.001

Germain, A. (2013). Sleep disturbances as the hallmark of PTSD: where are we now?. *American Journal of Psychiatry*, *170*(4), 372-382.

Golkar, A., Bellander, M., Olsson, A., & Öhman, A. (2012). Are fear memories erasable?– reconsolidation of learned fear with fear-relevant and fear-irrelevant stimuli. *Frontiers in behavioral neuroscience*, *6*, 80.doi: 10.3389/fnbeh.2012.00080

Goodmon, L. B., & Anderson, M. C. (2011). Semantic integration as a boundary condition on inhibitory processes in episodic retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37*(2), 416-436. Doi:10.1037/a0021963.

Grey, N., & Holmes, E. A. (2008). "Hotspots" in trauma memories in the treatment of post-traumatic stress disorder: A replication. *Memory, 16*(7), 788-796. doi:10.1080/09658210802266446

Gunter, R. W., & Bodner, G. E. (2008). How eye movements affect unpleasant memories: Support for a working-memory account. *Behaviour Research and Therapy, 46*, 913-931. doi:10.1016/j.brat.2008.04.006

Gunter, R. W., & Bodner, G. E. (2009). EMDR works... but how? Recent progress in the search for treatment mechanisms. *Journal of EMDR Practice and Research, 3*(3), 161-168. doi: 10.1891/1933-3196.3.3.161

Hackmann, A., Ehlers, A., Speckens, A., & Clark, D. M. (2004). Characteristics and Content of Intrusive Memories in PTSD and Their Changes With Treatment. *Journal of Traumatic Stress, 17*(3), 231–240. doi: 10.1023/B:JOTS.0000029266.88369.fd

Hackmann, A., & Holmes, E. (2004). Reflecting on imagery: a clinical perspective and overview of the special issue of memory on mental imagery and memory in psychopathology. *Memory, 12*, 389–402. doi:10.1080/09658210444000133

Hagenaars, M. A., van Minnen, A., Holmes, E. A., Brewin, C. R., & Hoogduin, K. A. (2008). The effect of hypnotically induced somatoform dissociation on the development of intrusions after an aversive film. *Cognition and Emotion*, *22*(5), 944-963. doi: 10.1080/02699930701575151

Hagenaars, M. A., Holmes, E. A., Klaassen, F., & Elzinga, B. (2017). Tetris and Word games affect intrusive memories when applied after memory reactivation 4 days post analogue trauma.

Hagger, M. S., Chatzisarantis, N. L., Alberts, H., Anggono, C. O., Batailler, C., Birt, A. R., ... & Calvillo, D. P. (2016). A multilab preregistered replication of the ego-depletion effect. *Perspectives on Psychological Science, 11*(4), 546-573. doi:10.1177/1745691616652873

Hamann, S. (2001). Cognitive and neural mechanisms of emotional memory. *Trends in Cognitive Sciences, 5*, 394–400.

Hanslmayr, S., Leipold, P., & Bäuml, K. H. (2010). Anticipation boosts forgetting of voluntarily suppressed memories. *Memory, 18*(3), 252-257. doi:10.1080/09658210903476548

Hanslmayr, S., Leipold, P., Pastötter, B., & Bäuml, K. H. (2009). Anticipatory signatures of voluntary memory suppression. *The Journal of Neuroscience, 29*(9), 2742-2747. doi:10.1523/JNEUROSCI.4703-08.2009

Hardwicke, T. E., Taqi, M., & Shanks, D. R. (2016). Postretrieval new learning does not reliably induce human memory updating via reconsolidation. *Proceedings of the National Academy of Sciences*, 201601440. doi:10.1073/pnas.1601440113

Hart, R. E., & Schooler, J. W. (2012). Suppression of novel stimuli: Changes in accessibility of suppressed nonverbalizable shapes. *Consciousness and Cognition, 21*, 1541-1546. doi:10.1016/j.concog.2012.06.005

Hellerstedt, R., Johansson, M., & Anderson, M. C. (2016). Tracking the intrusion of unwanted memories into awareness with event-related potentials. *Neuropsychologia*, *89*, 510-523. doi:10.1016/j.neuropsychologia.2016.07.008

Herbert, J. D., Lilienfeld, S. O., Lohr, J. M., Montgomery, R. W., O'Donohue, W. T., Rosen, G. M., & Tolin, D. F. (2000). Science and pseudoscience in the development of eye movement desensitization and reprocessing: Implications for clinical psychology. *Clinical Psychology Review, 20*(8), 945-971. doi:10.1016/S0272-7358(99)00017-3

Hertel, P. T., & Calcaterra, G. (2005). Intentional forgetting benefits from thought substitution. *Psychonomic Bulletin and Review, 12*(3), 484-489. doi:10.3758/BF03193792

Hertel, P. T., & Gerstle, M. (2003). Depressive deficits in forgetting. *Psychological Science, 14*(6), 573-578. doi:10.1046/j.0956-7976.2003.psci_1467.x

Hertel, P. T., & Hayes, J. A. (2015). Distracted by cues for suppressed memories. *Psychological science*, *26*(6), 775-783.

Hertel, P. T., Large, D., Stück, E. D., & Levy, A. (2012). Suppression-induced forgetting on a free-association test. *Memory, 20*(2), 100-109. doi:10.1080/09658211.2011.647036

Hertel, P., & McDaniel, L. (2010). The suppressive power of positive thinking: Aiding suppression-induced forgetting in repressive coping. *Cognition and Emotion, 24*(7), 1239-1249. doi:10.1080/02699930903172377

Hoffman, A. N., Parga, A., Paode, P. R., Watterson, L. R., Nikulina, E. M., Hammer, R. P., & Conrad, C. D. (2015). Chronic stress enhanced fear memories are associated with increased amygdala *zif268* mRNA expression and are resistant to reconsolidation. *Neurobiology of Learning and Memory*, *120*, 61–68. doi:10.1016/j.nlm.2015.02.004308. --- effects on go/nogo

Hofmann, S. G., & Smits, J. A. (2008). Cognitive-behavioral therapy for adult anxiety disorders: a meta-analysis of randomized placebo-controlled trials. *The Journal of clinical psychiatry, 69*(4), 621.

Holmes, E. A., Arntz, A., & Smucker, M. R. (2007). Imagery rescripting in cognitive behaviour therapy: Images, treatment techniques and outcomes. *Journal of behavior therapy and experimental psychiatry, 38*(4), 297-305.

Holmes, E. A., & Bourne, C. (2008). Inducing and modulating intrusive emotional memories: A review of the trauma film paradigm. *Acta psychologica, 127*(3), 553-566. doi:10.1016/j.actpsy.2007.11.002

Holmes, E. A., Brewin, C. R., & Hennessy, R. G. (2004). Trauma films, information processing, and intrusive memory development. *Journal of Experimental Psychology: General*, *133*(1), 3. doi:10.1037/0096-3445.133.1.3

Holmes, E. A., Craske, M. G., & Graybiel, A. M. (2014). A call for mental-health science. *Nature, 511*(7509), 287.

Holmes, E. A., Grey, N., & Young, K. A. (2005). Intrusive images and "hotspots" of trauma memories in posttraumatic stress disorder: An exploratory investigation of emotions and cognitive themes. *Journal of behavior therapy and experimental psychiatry, 36*(1), 3-17. doi:10.1016/j.jbtep.2004.11.002

Holmes, E. A., James, E. L., Coode-Bate, T., & Deeprose, C. (2009). Can playing the computer game "Tetris" reduce the build-up of flashbacks for trauma? A proposal from cognitive science. *PloS one, 4*(1), e4153. doi: 10.1371/journal.pone.0004153

Holmes, E. A., James, E. L., Kilford, E. J., & Deeprose, C. (2010). Key steps in developing a cognitive vaccine against traumatic flashbacks: Visuospatial Tetris versus verbal Pub Quiz. *PloS one, 5*(11), e13706. doi:10.1371/journal.pone.0013706

Horowitz, M. J. (1969). Psychic trauma. Return of images after a stressful film. *Archives of General Psychiatry, 20*(5), 552–559. doi:10.1001/archpsyc.1969.01740170056008

Horowitz, M., Wilner, N., & Alvarez, W. (1979). Impact of Event Scale: a measure of subjective stress. *Psychosomatic medicine, 41*(3), 209-218.

Horsch, A., Vial, Y., Favrod, C., Harari, M. M., Blackwell, S. E., Watson, P., ... & Holmes, E. A. (2017). Reducing intrusive traumatic memories after emergency caesarean section: A proof-of-principle randomized controlled study. *Behaviour Research and Therapy, 94,* 36-47. doi:10.1016/j.brat.2017.03.018

Hotta, C., & Kawaguchi, J. (2009). Self-initiated use of thought substitution can lead to long term forgetting. *Psychologia*, *52*(1), 41-49.

Hu, X., Bergström, Z. M., Bodenhausen, G. V., & Rosenfeld, J. P. (2015). Suppressing Unwanted Autobiographical Memories Reduces Their Automatic Influences Evidence From Electrophysiology and an Implicit Autobiographical Memory Test. *Psychological science*, 0956797615575734. doi:10.1177/0956797615575734

Hupbach, A. (2015). Retrieval practice does not safeguard memories from interference-based forgetting. *Learning and Motivation*, *49*, 23-30.

Hupbach, A., Gomez, R., Hardt, O., & Nadel, L. (2007). Reconsolidation of episodic memories: A subtle reminder triggers integration of new information. *Learning & Memory, 14*, 47-53. doi:10.1101/lm.365707

Hupbach, A., Gomez, R., & Nadel, L. (2009). Episodic memory reconsolidation: updating or source confusion?. *Memory, 17*(5), 502-510. doi:10.1080/09658210902882399

Hupbach, A., Hardt, O., Gomez, R., & Nadel, L. (2008). The dynamics of memory: Context dependent updating. *Learning & Memory, 15*, 574-579. doi: 10.1101/lm.1022308

Inzlicht, M., Bartholow, B.D., & Hirsh, J.B. (2015). Emotional foundations of cognitive control. *Trends in Cognitive Sciences*, *19*, 3, 126-132. doi: 10.1016/j.tics.2015.01.004

Ioannidis, J. P. (2005). Why most published research findings are false. *PLoS Med, 2*(8), e124. doi:10.1371/journal.pmed.0020124

Iyadurai, L., Blackwell, S. E., Meiser-Stedman, R., Watson, P. C., Bonsall, M. B., Geddes, J. R., ... & Holmes, E. A. (2017). Preventing intrusive memories after trauma via a brief intervention involving Tetris computer game play in the emergency department: a proof-of-concept randomized controlled trial. *Molecular Psychiatry*. doi: 10.1038/mp.2017.23

James, E. L., Bonsall, M. B., Hoppitt, L., Tunbridge, E. M., Geddes, J. R., Milton, A. L., & Holmes, E. A. (2015). Computer game play reduces intrusive memories of experimental trauma via reconsolidation-update mechanisms. *Psychological science*, *26*(8), 1201-1215.

James, E. L., Lau-Zhu, A., Clark, I. A., Visser, R. M., Hagenaars, M. A., & Holmes, E. A. (2016a). The trauma film paradigm as an experimental psychopathology model of

12

psychological trauma: intrusive memories and beyond. *Clinical psychology review, 47*, 106-142. doi: 10.1016/j.cpr.2016.04.010

James, E. L., Lau-Zhu, A., Tickle, H., Horsch, A., & Holmes, E. A. (2016*b*). Playing the computer game Tetris prior to viewing traumatic film material and subsequent intrusive memories: examining proactive interference. *Journal of Behavior Therapy and Experimental Psychiatry, 53*, 25-33. Doi:10.1016/j.jbtep.2015.11.004

JASP Team (2017). JASP (Version 0.8.2) [Computer software].

JASP Team (2016). JASP (Version 0.7.5) [Computer software].

Jeffreys, H. (1961). *Theory of Probability*. Oxford: UK Oxford University Press.

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological bulletin*, *114*(1), 3. doi:10.1037/0033-2909.114.1.3

Joormann, J., Hertel, P. T., Brozovich, F., & Gotlib, I. H. (2005). Remembering the good, forgetting the bad: Intentional forgetting of emotional material in depression. *Journal of Abnormal Psychology, 114(4),* 640-648. doi:10.1037/0021-843X.114.4.640

Joormann, J., Hertel, P. T., Lemoult, J., & Gotlib, I. H. (2009). Training forgetting of negative material in depression. *Journal of Abnormal Psychology, 118(1),* 34- 43. doi:10.1037/a0013794

Hagger, M. S., Chatzisarantis, N. L., Alberts, H., Anggono, C. O., Batailler, C., Birt, A. R., ... & Calvillo, D. P. (2016). A multilab preregistered replication of the ego-depletion effect. *Perspectives on Psychological Science, 11*(4), 546-573. doi:10.1177/1745691616652873

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

Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association, 90*, 773-795. doi:10.1080/01621459.1995.10476572

Kato, Y., Endo, H., & Kizuka, T. (2009). Mental fatigue and impaired response processes: Event-related brain potentials in a Go/NoGo task. International Journal of Psychophysiology, 72(2), 204–211. doi:10.1016/j.ijpsycho.2008.12.008

Kavanagh, D. J., Freese, S., Andrade, J., & May, J. (2001). Effects of visuospatial tasks on desensitization to emotive memories. *British Journal of Clinical Psychology, 40*, 267-280. doi:10.1348/014466501163689

Kazdin, A. E. (2001). Progression of therapy research and clinical application of treatment require better understanding of the change process. *Clinical Psychology: Science and Practice, 8*(2), 143–151.

Kearns, M., & Engelhard, I. M. (2015). Psychophysiological responsivity to script-driven imagery: An exploratory study of the effects of eye movements on public speaking flashforwards. *Frontiers in Psychiatry*, *6*.

Kemps, E., & Tiggemann, M. (2007). Reducing the vividness and emotional impact of distressing autobiographical memories: The importance of modality-specific interference. *Memory, 15*(4), 412-422. doi:10.1080/09658210701262017

Kensinger, E. A., & Corkin, S. (2004). Two routes to emotional memory: Distinct neural processes for valence and arousal. *Proceedings of the National Academy of Sciences, USA*, *101*, 3310-3315. doi:10.1073/pnas.0306408101

Kessler, R. C. (2000). Posttraumatic stress disorder: the burden to the individual and to society. *The Journal of clinical psychiatry, 61*, 4.

Kessler, R. C., Berglund, P., Demler, O., Jin, R., Merikangas, K. R., & Walters, E. E. (2005). Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. *Archives of general psychiatry, 62*(6), 593-602.

Kessler, R. C., Sonnega, A., Bromet, E., Hughes, M., & Nelson, C. B. (1995). Posttraumatic stress disorder in the National Comorbidity Survey. *Archives of general psychiatry*, *52*(12), 1048-1060. doi:10.1001/archpsyc.1995.03950240066012

Kessler, R.C., Tat Chiu, W., Demler, O., & Walters, E.E. (2005). Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry, 62*, 617-627. doi:10.1001/archpsyc.62.6.617

Kiesel, A., Steinhauser, M., Wendt, M., Falkenstein, M., Jost, K., Philipp, A. M., & Koch, I. (2010). Control and interference in task switching—A review. *Psychological bulletin*, *136*(5), 849.

Kim, D. Y., Oh, D. H., Kim, S. H., Sim, K. B., & Lee, J. H. (2013). Effects of intentional suppression of recall of unwanted images in repressors and nonrepressors. *Social Behavior and Personality: an international journal, 41*(2), 319-326. doi:10.2224/sbp.2013.41.2.319

Kim, K., & Yi, D.-J. (2013). Out of mind, out of sight: perceptual consequences of memory suppression. *Psychological Science, 24*, 569–574. doi: 10.1177/0956797612457577

Kindt, M., & van Emmerik, A. (2016). New avenues for treating emotional memory disorders: towards a reconsolidation intervention for posttraumatic stress disorder. *Therapeutic advances in psychopharmacology*, *6*(4), 283-295. doi:10.1177/ 2045125316644541

Kindt, M., & Soeter, M. (2013). Reconsolidation in a human fear conditioning study: a test of extinction as updating mechanism. *Biological psychology, 92*(1), 43-50. doi:10.1016/j.biopsycho.2011.09.016

Kindt, M., Soeter, M., & Vervliet, B. (2009). Beyond extinction: erasing human fear responses and preventing the return of fear. *Nature Neuroscience, 12*(3), 256-258.doi:10.1038/nn.2271

Kindt, M., van den Hout, M., Arntz, A., & Drost, J. (2008). The influence of data-driven versus conceptually-driven processing on the development of PTSD-like symptoms. *Journal of Behavior Therapy and Experimental Psychiatry, 39*(4), 546-557. doi:10.1016/j.jbtep.2007.12.003

Kleim, B., Graham, B., Bryant, R. A., & Ehlers, A. (2013). Capturing intrusive re-experiencing in trauma survivors' daily lives using ecological momentary assessment. *Journal of abnormal psychology, 122*(4), 998. doi:10.1037/a0034957

Klucken, T., Kruse, O., Schweckendiek, J., Kuepper, Y., Mueller, E. M., Hennig, J., & Stark, R. (2016). No evidence for blocking the return of fear by disrupting reconsolidation prior to extinction learning. *Cortex, 79*, 112-122. doi: 10.1016/j.cortex.2016.03.015

Koch, I., Gade, M., Schuch, S., & Philipp, A. M. (2010). The role of inhibition in task switching: A review. *Psychonomic bulletin & review*, *17*(1), 1-14.

Kool, W., McGuire, J.T., Rosen, Z.B., & Botvinick, M.M. (2010). Decision making and the avoidance of cognitive demand. *Journal of Experimental Psychology: General. 139,* 665-682. doi: 10.1037/a0020198

Koole, S. L., & Lakens, D. (2012). Rewarding replications a sure and simple way to improve psychological science. *Perspectives on Psychological Science, 7*(6), 608-614. doi:10.1177/1745691612462586

Kosslyn, S. M., Ganis, G., & Thompson, W. L. (2001). Neural foundations of imagery. *Nature Reviews Neuroscience*, *2*(9), 635-642. doi:10.1038/35090055

Krans, J., Näring, G., & Becker, E. S. (2009). Count out your intrusions: effects of verbal encoding on intrusive memories. Memory, 17(8), 809-815. doi:10.1080/09658210903130780

Krawczyk, M. C., Fernández, R. S., Pedreira, M. E., & Boccia, M. M. (2017). Toward a better understanding on the role of prediction error on memory processes: From bench to clinic. *Neurobiology of learning and memory*, *142*, 13-20.

Kredlow, M. A., & Otto, M. W. (2015). Interference with the reconsolidation of trauma-related memories in adults. *Depression and Anxiety, 32*(1), 32-37. doi: 10.1002/da.22343

Kroes, M. C., Schiller, D., LeDoux, J. E., & Phelps, E. A. (2015). Translational approaches targeting reconsolidation. In *Translational Neuropsychopharmacology* (pp. 197-230). Springer International Publishing.

Krypotos, A. M., Blanken, T. F., Arnaudova, I., Matzke, D., & Beckers, T. (2016). A primer on Bayesian analysis for experimental psychopathologists. *Journal of Experimental Psychopathology*, *8*(2), 140. doi:10.5127/jep.057316

Krypotos, A. M., Klugkist, I., & Engelhard, I. M. (2017). Bayesian hypothesis testing for human threat conditioning research: an introduction and the condir R package. European Journal of Psychotraumatology, 8(sup1), 1314782. doi: 10.1080/20008198.2017.1314782

Küpper, C. S., Benoit, R. G., Dalgleish, T., & Anderson, M. C. (2014). Direct suppression as a mechanism of controlling unpleasant memories in daily life. *Journal of Experimental Psychology: General*, *143*(4), 1443-1449. doi:10.1037/a0036518

Kvavilashvili, L. (2014). Solving the mystery of intrusive flashbacks in posttraumatic stress disorder: Comment on Brewin (2014). *Psychological Bulletin*, *140*, 98-104. doi: 10.1037/a0034677

Kvavilashvili, L., & Mandler, G. (2004). Out of one's mind: A study of involuntary semantic memories. *Cognitive Psychology, 48*(1), 47-94. doi:10.1016/S0010-0285(03)00115-4

LaBar, K. S., & Cabeza R. (2006). Cognitive neuroscience of emotional memory. *Nature Reviews Neuroscience 7*(4), 54–64.

Lambert, A. J., Good, K. S., & Kirk, I. J. (2010). Testing the repression hypothesis: Effects of emotional valence on memory suppression in the think–no think task. *Consciousness and cognition*, *19*(1), 281-293.

Lane, R. D., Nadel, L., Greenberg, L., & Ryan, L. (2015). The integrated memory model: A new framework for understanding the mechanisms of change in psychotherapy. *Behavioral and Brain Sciences*, 38, e30. doi:10.1017/S0140525X14000041, e1

Lang, P.J., Bradley, M.M., & Cuthbert, B.N. (2005). *International affective picture system (IAPS): Digitized photographs, instruction manual and affective ratings. Technical Report A-6.* University of Florida, Gainesville, FL.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (2008). International affective picture system (IAPS): Affective ratings of pictures and instruction manual. *Technical report A-8*.

Lattal, K. M., & Abel, T. (2004). Behavioral impairments caused by injections of the protein synthesis inhibitor anisomycin after contextual retrieval reverse with time. *Proceedings of the National Academy of Sciences of the United States of America, 101*(13), 4667-4672.

LeDoux, J. E., Romanski, L., & Xagoraris, A. (1989). Indelibility of subcortical emotional memories. *Journal of Cognitive Neuroscience, 1*(3), 238-243. doi:10.1162/jocn.1989.1.3.238

Lee, J. L. (2009). Reconsolidation: maintaining memory relevance. *Trends in neurosciences*, *32*(8), 413-420.

Lee, C. W., & Cuijpers, P. (2013). A meta-analysis of the contribution of eye movements in processing emotional memories. *Journal of Behavior Therapy and Experimental Psychiatry*, *44*(2), 231-239.

Lee, C. W., & Drummond, P. D. (2008). Effects of eye movement versus therapist instructions on the processing of distressing memories. *Journal of Anxiety Disorders, 22*(5), 801-808.

Lee, Y., Lee, H., & Tsai, S. (2007). Effects of post-cue interval on intentional forgetting. *British Journal of Psychology, 98*, 257-272. doi:10.1348/000712606X120410

Lee, J. L., Nader, K., & Schiller, D. (2017). An Update on Memory Reconsolidation Updating. *Trends in Cognitive Sciences*.

Leer, A., Engelhard, I. M., Altink, A., & van den Hout, M. A. (2013). Eye movements during recall of aversive memory decreases conditioned fear. *Behaviour research and therapy, 51*(10), 633-640. doi:10.1016/j.brat.2013.07.004

Leer, A., Engelhard, I. M., Dibbets, P., & Van den Hout, M. A. (2013). Dual-tasking attenuates the return of fear after extinction. *Journal of Experimental Psychopathology, 4*(4), 325-340.

Leer, A., Engelhard, I. M., Lenaert, B., Struyf, D., Vervliet, B., & Hermans, D. (2017). Eye movement during recall reduces objective memory performance: An extended replication. *Behaviour Research and Therapy, 92*, 94-105. doi: 10.1016/j.brat.2017.03.002

Leer, A., Engelhard, I. M., & van den Hout, M. A. (2014). How eye movements in EMDR work: Changes in memory vividness and emotionality. *Journal of behavior therapy and experimental psychiatry*, *45*(3), 396-401. doi: 10.1016/j.jbtep.2014.04.004

LeMoult, J., Hertel, P. T., & Joormann, J. (2010). Training the forgetting of negative words: The role of direct suppression and the relation to stress reactivity. *Applied Cognitive Psychology, 24*(3), 365-375. doi:10.1002/acp.1682

Levine, L. J., & Pizarro, D. A. (2004). Emotion and memory research: A grumpy overview. *Social Cognition, 22*(5), 530–554.

Levy, B. J., & Anderson, M. C. (2008). Individual differences in the suppression of unwanted memories: The executive deficit hypothesis. *Acta Psychologica, 127(3),* 623-635. doi:10.1016/j.actpsy.2007.12.004

Levy, B. J., & Anderson, M. C. (2012). Purging of memories from conscious awareness tracked in the human brain. *Journal of Neuroscience*, *32*(47), 16785-16794. doi:10.1523/JNEUROSCI.2640-12.2012

Levy, B. J., & Wagner, A. D. (2011). Cognitive control and right ventrolateral prefrontal cortex: reflexive reorienting, motor inhibition, and action updating. *Annals of the New York Academy of Sciences*, *1224*(1), 40-62. doi:10.1111/j.1749-6632.2011.05958.x

Lewis, D. J. (1979). Psychobiology of active and inactive memory. *Psychological Bulletin*, *86*(5), 1054. doi:10.1037/0033-2909.86.5.1054

Lewis, D. J., Bregman, N. J., & Mahan, J. J. (1972). Cue-dependent amnesia in rats. *Journal of comparative and physiological psychology*, *81*(2), 243.

Littel, M., Remijn, M., Tinga, A. M., Engelhard, I. M., & van den Hout, M. A. (2017). Stress enhances the memory-degrading effects of eye movements on emotionally neutral memories. Clinical Psychological Science, 5(2), 316-324.

Loerinc, A. G., Meuret, A. E., Twohig, M. P., Rosenfield, D., Bluett, E. J., & Craske, M. G. (2015). Response rates for CBT for anxiety disorders: need for standardized criteria. *Clinical psychology review, 42*, 72-82.

Loftus, E. F., & Palmer, J. C. (1974). Reconstruction of automobile destruction: An example of the interaction between language and memory. *Journal of Verbal Learning and Verbal Behavior, 13*(5), 585-589. doi:10.1016/S0022-5371(74)80011-3

Logan, G.D. (1994). On the ability to inhibit thought and action: A users' guide to the stop signal paradigm. In: Dagenbach Dand Carr TH (eds.) Inhibitory Processes in Attention, Memory and Language, pp. 189–239. San Diego, CA: Academic Press.

Lohr, J. M., Tolin, D. F., & Lilienfeld, S. O. (1999). Efficacy of eye movement desensitization and reprocessing: Implications for behavior therapy. *Behavior Therapy*, *29*(1), 123-156.

Lonergan, M. H., Olivera-Figueroa, L. A., Pitman, R. K., & Brunet, A. (2013). Propranolol's effects on the consolidation and reconsolidation of long-term emotional memory in healthy participants: a meta-analysis. *Journal of Psychiatry & Neuroscience, 38*(4), 222-31. doi: 10.1503/jpn.120111

Mace, J. H. (2004). Involuntary autobiographical memories are highly dependent on abstract cuing: The Proustian view is incorrect. *Applied Cognitive Psychology, 18*(7), 893-899. doi: 10.1002/acp.1020

Maes, E., Boddez, Y., Alfei Palloni, J. M., Krypotos, A. M., D'Hooge, R., De Houwer, J., & Beckers, T. (2016). The elusive nature of the blocking effect: 15 failures to replicate. *Journal of Experimental Psychology: General*. *145*(9), e49. doi:10.1037/xge0000200

Makel, M. C., Plucker, J. A., & Hegarty, B. (2012). Replications in psychology research how often do they really occur?. Perspectives on Psychological Science, 7(6), 537-542. DOI: 10.1177/1745691612460688

Marks, D. F. (1995). New directions for mental imagery research. *Journal of Mental Imagery, 19,* 153-167.

Marx, B. P., Marshall, P. J., & Castro, F. (2008). The moderating effects of stimulus valence and arousal on memory suppression. *Emotion, 8(2),* 199-207. doi:10.1037/1528-3542.8.2.199.

Marzi, T., Regina, A., & Righi, S. (2014). Emotions shape memory suppression in trait anxiety. *Frontiers in Psychology*, advance online publication. doi:10.3389/fpsyg.2013.01001

Maslany, A. J., & Campbell, J. I. (2013). Failures to replicate hyper-retrieval-induced forgetting in arithmetic memory. Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale, 67(1), 72. doi: 10.1037/a0031138

Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods*, *44*(2), 314-324. doi: 10.3758/s13428-011-0168-7

Matthijssen, S. J. M. A., van Schie, K., & van den Hout, M. A., (2017). The Effect of Modality Specific Taxing during Memory Recall: Results from a Random Interval Repetition Task.

Matzke, D., Nieuwenhuis, S., van Rijn, H., Slagter, H. A., van der Molen, M. W., & Wagenmakers, E. J. (2015). The effect of horizontal eye movements on free recall: A preregistered adversarial collaboration. *Journal of Experimental Psychology: General*, *144*(1), e1. doi:10.1037/xge0000038

Maxfield, L., Melnyk, W.T., & Hayman, C. A. G. (2008). A working memory explanation for the effects of eye movements in EMDR. *Journal of EMDR Practice and Research, 2*, 247-261. doi:10.1891/1933-3196.2.4.247

May, J., Andrade, J., Panabokke, N., & Kavanagh, D. (2010). Visuospatial tasks suppress craving for cigarettes. *Behaviour Research and Therapy, 48*(6), 476-485. doi:10.1016/j.brat.2010.02.001

McGaugh, J. L. (1966). Time-dependent processes in memory storage. *Science*, *153*(3742), 1351-1358.

McGaugh, J. L. (2000). Memory – a century of consolidation. *Science, 287*(5451), 248-251. doi: 10.1126/science.287.5451.248

McGeoch, J. A. (1942). The psychology of human learning: An introduction.

McNally, R. J. (1999). EMDR and mesmerism: A comparative historical analysis. *Journal of Anxiety Disorders, 13*(1), 225-236.

McNally, R. J. (2007). Mechanisms of exposure therapy: How neuroscience can improve psychological treatments for anxiety disorders. *Clinical Psychology Review, 27*(6), 750–759.

Meier, B., König, A., Parak, S., & Henke, K. (2011). Suppressed, but not forgotten. *Swiss Journal of Psychology*.

Misanin, J. R., Miller, R. R., & Lewis, D. J. (1968). Retrograde amnesia produced by electroconvulsive shock after reactivation of a consolidated memory trace. *Science*, *160*(3827), 554-555.

Monden, R., Vos, S., Morey, R., Wagenmakers, E. J., Jonge, P., & Roest, A. M. (2016). Toward evidence-based medical statistics: a Bayesian analysis of double-blind placebo-controlled antidepressant trials in the treatment of anxiety disorders. *International journal of methods in psychiatric research*, *25*(4), 299-308. doi: 10.1002/mpr.1507

Moors, A., De Houwer, J., Hermans, D., Wanmaker, S., van Schie, K., Van Harmelen, A. L., ... & Brysbaert, M. (2013). Norms of valence, arousal, dominance, and age of acquisition for 4,300 Dutch words. *Behavior research methods*, *45*(1), 169-177. doi 10.3758/s13428-012-0243-8

Morey, R. D., & Rouder, J. N. (2015). Bayesfactor: Computation of bayes factors for common designs [Computer software manual]. Retrieved from http://CRAN.R-project.org/package=BayesFactor (R package version 0.9.2+)

Mulder, J., Hoijtink, H., & de Leeuw, C. (2012). Biems: A Fortran90 program for calculating Bayes factors for inequality and equality constrained models. *Journal of Statistical Software, 46*(2), 1-39.

Mulder, J., Hoijtink, H., & Klugkist, I. (2010). Equality and inequality constrained multivariate linear models: Objective model selection using constrained posterior priors. *Journal of Statistical Planning and Inference, 140*(4), 887-906. doi:10.1016/j.jspi.2009.09.022

Mulder, J., Klugkist, I., van de Schoot, R., Meeus, W. H., Selfhout, M., & Hoijtink, H. (2009). Bayesian model selection of informative hypotheses for repeated measurements. *Journal of Mathematical Psychology, 53*(6), 530-546.

Müller, G. E., & Pilzecker, A. (1900). *Experimentelle beiträge zur lehre vom gedächtniss* (Vol. 1). JA Barth.

Muris, P., & Merckelbach, H. (1999). Traumatic memories, eye movements, phobia, and panic: A critical note on the proliferation of EMDR. *Journal of Anxiety Disorders, 13*, 209–223.

Murray, B. D., Anderson, M. C., & Kensinger, E. A. (2015). Older Adults Can Suppress Unwanted Memories When Given an Appropriate Strategy. *Psychology and Aging*. Advance online publication. doi:10.1037/a0038611

Murray, B. D., Muscatell, K. A., and Kensinger, E. A. (2011). Effects of emotion and age on performance on a think/no-think memory task. *Psychology and Aging*, 26, 940-955.

Nader, K., & Einarsson, E. Ö. (2010). Memory reconsolidation: an update. *Annals of the New York Academy of Sciences*, *1191*(1), 27-41.

Nader, K. (2015). Reconsolidation and the Dynamic Nature of Memory. *Cold Spring Harbor perspectives in biology*, *7*(10), a021782. doi: 10.1101/cshperspect.a021782

Nader, K., Schafe, G. E., & LeDoux, J. E. (2000). Fear memories require protein synthesis in the amygdala for reconsolidation after retrieval. *Nature, 406*, 722–726. doi:10.1038/35021052

National Institute for Clinical Excellence. (2011). Common mental health disorders: Identification and pathways to care. In Clinical guideline 123: . London, UK: National Institute for Health and Clinical Excellence.

Neria, Y., Nandi, A., & Galea, S. (2008). Post-traumatic stress disorder following disasters: a systematic review. *Psychological medicine, 38*(04), 467-480.

Nixon, R. D., Cain, N., Nehmy, T., & Seymour, M. (2009). The influence of thought suppression and cognitive load on intrusions and memory processes following an analogue stressor. *Behavior Therapy, 40*(4), 368-379.doi: 10.1016/j.beth.2008.10.004

Nørby, S. (2017). Forgetting and emotion regulation in mental health, anxiety and depression. *Memory*, 1-22.

Nørby, S., Lange, M., & Larsen, A. (2010). Forgetting to forget: On the duration of voluntary suppression of neutral and emotional memories. *Acta Psychologica, 133(1),* 73-80. doi:10.1016/j.actpsy.2009.10.002

Noreen, S, & MacLeod, M. D. (2013). It's All in the Detail: Intentional Forgetting of Autobiographical Memories Using the Autobiographical Think/No-Think Task. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39*(2), 375-393. doi:10.1037/a0028888

Noreen, S., & MacLeod, M. D. (2014). To think or not to think, that is the question: Individual differences in suppression and rebound effects in autobiographical memory. *Acta psychologica, 145*, 84-97. doi:10.1016/j.actpsy.2013.10.011

Noreen, S., & MacLeod, M. D. (2015). What do we really know about cognitive inhibition? Task demands and inhibitory effects across a range of memory and behavioural tasks. *PloS one*, *10*(8), e0134951.

Nosek, B. A., Spies, J. R., & Motyl, M. (2012). Scientific utopia II. Restructuring incentives and practices to promote truth over publishability. *Perspectives on Psychological Science, 7*(6), 615-631. doi:10.1177/1745691612459058

Olatunji, B. O., Cisler, J. M., & Deacon, B. J. (2010). Efficacy of cognitive behavioral therapy for anxiety disorders: a review of meta-analytic findings. *Psychiatric Clinics of North America, 33*(3), 557-577.

Olatunji, B. O., Leen-Feldner, E. W., Feldner, M. T., & Forsyth, J. P. (2007). Experimental psychopathology. Handbook of research methods in abnormal and clinical psychology, 47-60.

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science, 349*(6251), aac4716.

Parsons, R. G., & Ressler, K. J. (2013). Implications of memory modulation for post-traumatic stress and fear disorders. *Nature neuroscience, 16*(2), 146-153. doi:10.1038/nn.3296

Pashler, H., & Wagenmakers, E. J. (2012). Editors' Introduction to the Special Section on Replicability in Psychological Science A Crisis of Confidence?. *Perspectives on Psychological Science, 7*(6), 528-530. doi:10.1177/1745691612465253

Paz-Alonso, P. M., Ghetti, S., Matlen, B. J., Anderson, M. C., & Bunge, S. A. (2009). Memory suppression is an active process that improves over childhood. *Frontiers in Human Neuroscience, 3,* 1-6. doi:10.3389/neuro.09.024.2009

Pearson, D. G., & Sawyer, T. (2011). Effects of dual task interference on memory intrusions for affective images. *International Journal of Cognitive Therapy, 4*(2), 122-133. doi: 10.1521/ijct.2011.4.2.122

Philippot, P., & Hermans, D. (2006). Experimental psychopathology: From laboratory studies to clinical practice. *Psychologica belgica, 46*(1-2).

Pitman, R. K. (1989). Post-traumatic stress disorder, hormones, and memory. *Biological Psychiatry, 26*(3), 221-223.

Pitman, R. K. (2011). Will reconsolidation blockade offer a novel treatment for posttraumatic stress disorder?. *Frontiers in behavioral neuroscience, 5*, 11. doi:10.3389/fnbeh.2011.00011. eCollection 2011.

Potts, R., & Shanks, D. R. (2012). Can testing immunize memories against interference?. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*(6), 1780.

Psychology Software Tools, Inc. [E-Prime 2.0]. (2012). Retrieved from http://www.pstnet.com.

Raudenbush, S.W., Bryk, A.S, & Congdon, R. (2004). HLM 6 for Windows [Computer software]. Skokie, IL: Scientific Software International, Inc.

R Core Team (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Redick, T. S., Broadway, J. M., Meier, M. E., Kuriakose, P. S., Unsworth, N., Kane, M. J., & Engle, R. W. (2012). Measuring working memory capacity with automated complex span tasks. *European Journal of Psychological Assessment, 28*(3), 164–171. doi:10.1027/1015-5759/a000123

Redick, T. S., Calvo, A., Gay, C. E., & Engle, R. W. (2011). Working memory capacity and go/no-go task performance: selective effects of updating, maintenance, and inhibition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(2),

Redick, T. S., Heitz, R. P., & Engle, R. W. (2007). Working memory capacity and inhibition: Cognitive and social consequences. *Inhibition in cognition*, 125-142.

Restle, F. (1970). Speed of adding and comparing numbers. *Journal of Experimental Psychology*, *83*(2 Pt 1), 274-278.

Reisberg, D., Pearson, D. G., & Kosslyn, S. M. (2003). Intuitions and introspections about imagery: The role of imagery experience in shaping an investigator's theoretical views. *Applied Cognitive Psychology*, *17*(2), 147-160. doi: 10.1002/acp.858

Riccio, D. C., Millin, P. M., & Bogart, A. R. (2006). Reconsolidation: A brief history, a retrieval view, and some recent issues. *Learning & Memory*, *13*(5), 536-544.

Robitzsch, A., Grund, S., Henke, T., & Robitzsch, M. A. (2016). Package 'miceadds'.

Roediger, H. L., & Karpicke, J. D. (2006). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science, 17*(3), 249-255. doi:10.1111/j.1467-9280.2006.01693.x

Rogers, S., & Silver, S. M. (2002). Is EMDR an exposure therapy? A review of trauma protocols. *Journal of Clinical Psychology*, *58*(1), 43-59.

Rosen, G. M. (1995). On the origin of eye movement desensitization.*Journal of Behavior Therapy and Experimental Psychiatry, 26*, 121-122

Rosen, G. M., Lohr, J. M., McNally, R. J., & Herbert, J. D. (1998). Power therapies, miraculous claims, and the cures that fail. *Behavioural and Cognitive Psychotherapy, 26*(2), 99-101.

Rothbaum, B. O, Meadows, E. A., Resick, P., & Foy, D. W. (2000). In E. B. Foa, T. M. Keane, & M. J. Friedman (Eds.), *Effective treatments for PTSD: Practice guidelines from the International Society for Traumatic Stress Studies*. New York, NY, US: Guilford Press 320–325.

Samuelson, K. W., Neylan, T. C., Metzler, T. J., Lenoci, M., Rothlind, J., Henn-Haase, C., ... & Marmar, C. R. (2006). Neuropsychological functioning in posttraumatic stress disorder and alcohol abuse. *Neuropsychology, 20*(6), 716. doi: 10.1037/0894-4105.20.6.716

Santiago, P. N., Ursano, R. J., Gray, C. L., Pynoos, R. S., Spiegel, D., Lewis-Fernandez, R., ... & Fullerton, C. S. (2013). A systematic review of PTSD prevalence and trajectories in DSM-5 defined trauma exposed populations: intentional and non-intentional traumatic events. *PloS one, 8*(4), e59236. doi:10.1371/journal.pone.0059236

Sara, S. J. (2000). Retrieval and reconsolidation: toward a neurobiology of remembering. *Learning & Memory*, *7*(2), 73-84.

Schie, K. van, Geraerts, E., & Anderson, M. C. (2013). Emotional and non-emotional memories are suppressible under direct suppression instructions. *Cognition and Emotion, 27*, 1122-1131. doi:10.1080/02699931.2013.765387

Schiller, D., Monfils, M., Raio, C. M., Johnson, D. C., Le Doux, J. E., & Phelps, E. A. (2010). Preventing the return of fear in humans using reconsolidation update mechanisms. *Nature, 463*, 49-52. doi:10.1038/nature08637

Schiller, D., and Phelps, E. A. (2011). Does reconsolidation occur in humans? *Front. Behav. Neurosci*. 5:24. doi:10.3389/fnbeh.2011.00024

Schlagman, S., & Kvavilashvili, L. (2008). Involuntary autobiographical memories in and outside the laboratory: How different are they from voluntary autobiographical memories?. *Memory & Cognition, 36*(5), 920-932. doi:10.3758/MC.36.5.920

Schneider, A. M., & Sherman, W. (1968). Amnesia: a function of the temporal relation of footshock to electroconvulsive shock. *Science*, *159*(3811), 219-221.

Schroyens, N., Beckers, T., & Kindt, M. (2017). In search for boundary conditions of reconsolidation: A failure of fear memory interference. *Frontiers in behavioral neuroscience*, *11*.

Schubert, S. J., Lee, C. W., & Drummond, P. D. (2011). The efficacy and psychophysiological correlates of dual-attention tasks in eye movement desensitization and reprocessing (EMDR). *Journal of anxiety disorders*, *25*(1), 1-11. doi: 10.1016/j.janxdis.2010.06.024

Schwabe, L., Nader, K., & Pruessner, J. C. (2013). β-Adrenergic blockade during reactivation reduces the subjective feeling of remembering associated with emotional episodic memories. *Biological Psychology, 92*(2), 227-232. doi:10.1016/j.biopsycho.2012.10.003

Schwabe, L., Nader, K., & Pruessner, J. C. (2014). Reconsolidation of human memory: brain mechanisms and clinical relevance. *Biological Psychiatry, 76*(4), 274-280. doi:10.1016/j.biopsych.2014.03.008

Schwabe, L., & Wolf, O. T. (2009). New episodic learning interferes with the reconsolidation of autobiographical memories. *PLoS One, 4*(10), e7519. doi: 10.1371/journal.pone.0007519

Schweizer, S., Samimi, Z., Hasani, J., Moradi, A., Mirdoraghi, F., & Khaleghi, M. (2017). Improving cognitive control in adolescents with post-traumatic stress disorder (PTSD). *Behaviour Research and Therapy*, *93*, 88-94. doi: 10.1016/j.brat.2017.03.017

Scully, I. D., Napper, L. E., & Hupbach, A. (2017). Does reactivation trigger episodic memory change? A meta-analysis. *Neurobiology of learning and memory*, *142*, 99-107.

Seidler, G. H., & Wagner, F. E. (2006). Comparing the efficacy of EMDR and trauma-focused Cognitive-Behavioral Therapy in the Treatment of PTSD: a meta–analytic study. *Psychological Medicine, 36*, 1515-1522.

Sergent, C., Baillet, S., & Dehaene, S. (2005). Timing of the brain events underlying access to consciousness during the attentional blink. *Nature neuroscience, 8*(10), 1391-1400. doi:10.1038/nn1549

Sergent, C., & Dehaene, S. (2004). Is consciousness a gradual phenomenon? Evidence for an all-or-none bifurcation during the attentional blink. *Psychological science, 15*(11), 720-728. doi: 10.1111/j.0956-7976.2004.00748.x

Sevenster, D., Beckers, T., & Kindt, M. (2012). Retrieval per se is not sufficient to trigger reconsolidation of human fear memory. *Neurobiology of learning and memory, 97*(3), 338-345. doi:10.1016/j.nlm.2012.01.009

Sevenster, D., Beckers, T., & Kindt, M. (2013). Prediction error governs pharmacologically induced amnesia for learned fear. *Science, 339*(6121), 830-833. doi:10.1126/science.1231357

Sevenster, D., Beckers, T., & Kindt, M. (2014). Prediction error demarcates the transition from retrieval, to reconsolidation, to new learning. *Learning & Memory*, *21*(11), 580-584.

Shapiro, F. (1989). Eye movement desensitization: A new treatment for post-traumatic stress disorder. *Journal of behavior therapy and experimental psychiatry, 20*(3), 211-217.

Shapiro, F. (1995). Eye movement desensitization and reprocessing: Basic principles, protocols and procedures. New York: The Guilford Press.

Shipherd, J. C., & Beck, J. G. (2005). The role of thought suppression in posttraumatic stress disorder. *Behavior Therapy. 36*, 277–87.

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological science*, 0956797611417632. doi:10.1177/0956797611417632

Simons, D. J., & Chabris, C. F. (2011). What people believe about how memory works: A representative survey of the US population. *PloS one*, *6*(8), e22757.

Simonsohn, U. (2015). Small telescopes detectability and the evaluation of replication results. *Psychological Science*, *26*(5), 559-569. doi:10.1177/0956797614567341

Smeets, M. A. M, Dijs, W. M., Pervan, I., Engelhard, I. M., & Van den Hout, M. A. (2012): Time-course of eye movement-related decrease in vividness and emotionality of unpleasant autobiographical memories, *Memory, 20*(4,) 346-357. doi:10.1080/09658211.2012.665462

Soeter, M., & Kindt, M. (2011). Disrupting reconsolidation: pharmacological and behavioral manipulations. *Learning & Memory, 18*(6), 357-366. doi:10.1101/lm.2148511

Soeter, M., & Kindt, M. (2012). Erasing fear for an imagined threat event. *Psychoneuroendocrinology, 37*(11), 1769-1779. doi:10.1016/j.psyneuen.2012.03.011

Soeter, M., & Kindt, M. (2013). High trait anxiety: A challenge for disrupting fear memory reconsolidation. *PloS one*, *8*(11), e75239.

Soeter, M., & Kindt, M. (2015). An abrupt transformation of phobic behavior after a post-retrieval amnesic agent. *Biological psychiatry*, *78*(12), 880-886.

Spielberger, C. D., Gorsuch, R. L., Lushene, R., Vagg, P. R., & Jacobs, G. A. (1983). State-trait anxiety inventory (form Y). Palo Alto. *CA: Mind Garden*.

Srinagesh, K. (2006). *The principles of experimental research*. Oxford: Butterworth-Heinemann.

Stephens, E., Braid, A., & Hertel, P. T. (2013). Suppression-induced reduction in the specificity of autobiographical memories. *Clinical Psychological Science, 1*(2), 16. doi:10.1177/2167702612467773

Stuart, A. D., Holmes, E. A., & Brewin, C. R. (2006). The influence of a visuospatial grounding task on intrusive images of a traumatic film. *Behaviour Research and Therapy*, *44*(4), 611-619. doi:10.1016/j.brat.2005.04.004

Suzuki, A., Josselyn, S. A., Frankland, P. W., Masushige, S., Silva, A. J., & Kida, S. (2004). Memory reconsolidation and extinction have distinct temporal and biochemical signatures. *The Journal of Neuroscience, 24*(20), 4787-4795.

Tabachnick, B.G., Fidell, L.S., 2013. Using, 6th ed. Pearson Education, Boston.

Tadmor, A., McNally, R. J., & Engelhard, I. M. (2016). Reducing the negative valence of stressful memories through emotionally valenced, modality-specific tasks. *Journal of behavior therapy and experimental psychiatry, 53*, 92-98. doi:10.1016/j.jbtep.2014.11.001

12

Thomaes, K., Engelhard, I. M., Sijbrandij, M., Cath, D. C., & Van den Heuvel, O. A. (2016). Degrading traumatic memories with eye movements: a pilot functional MRI study in PTSD. *European journal of psychotraumatology, 7*(1), 31371.

Thome, J., Koppe, G., Hauschild, S., Liebke, L., Schmahl, C., Lis, S., & Bohus, M. (2016). Modification of fear memory by pharmacological and behavioural interventions during reconsolidation. *PloS one, 11*(8), e0161044.

Treanor, M., Brown, L. A., Rissman, J., & Craske, M. G. (2017). Can Memories of Traumatic Experiences or Addiction Be Erased or Modified? A Critical Review of Research on the Disruption of Memory Reconsolidation and Its Applications. *Perspectives on Psychological Science*, *12*(2), 290-305.

Tulving, E., & Pearlstone, Z. (1966). Availability versus accessibility of information in memory for words. *Journal of Verbal Learning and Verbal Behavior, 5*, 381–391. doi: 10.1016/S0022-5371(66)80048-8

van de Schoot, R., Broere, J. J., Perryck, K. H., Zondervan-Zwijnenburg, M., & Van Loey, N. E. (2015). Analyzing small data sets using Bayesian estimation: the case of posttraumatic stress symptoms following mechanical ventilation in burn survivors. *European journal of Psychotraumatology, 6*. doi:10.3402/ejpt.v6.25216

van Etten, M. L., & Taylor, S. (1998). Comparative efficacy of treatments for post-traumatic stress disorder: a meta-analysis. *Clinical Psychology & Psychotherapy, 5*, 126–144

van den Hout, M. A., Bartelski, N., & Engelhard, I. M. (2013*a*). On EMDR: Eye movements during retrieval reduce subjective vividness and objective memory accessibility during future recall. *Cognition & Emotion, 27*(1), 177-183. doi 10.1080/02699931.2012.691087

van den Hout, M. A., Eidhof, M. B., Verboom, J., Littel, M, & Engelhard, I. M. (2014). Blurring of emotional and non-emotional memories by taxing working memory during recall. *Cognition & Emotion, 28*(4), 717-727. doi: 10.1080/02699931.2013.848785

van den Hout, M. A., & Engelhard, I. M. (2012). How does EMDR work?. *Journal of Experimental Psychopathology, 3*(5), 724-738. doi:10.5127/jep.028212

van den Hout, M. A., Engelhard, I. M., Beetsma, D., Slofstra, C., Hornsveld, H., Houtveen, J., & Leer, A. (2011*a*). EMDR and mindfulness. Eye movements and attentional breathing tax working memory and reduce vividness and emotionality of aversive ideation. *Journal of Behavior Therapy and Experimental Psychiatry, 42*(4), 423-431.

van den Hout, M. A., Engelhard, I. M., & McNally, R. J. (2016). Thoughts on experimental psychopathology. *Psychopathology Review, 4*(2), 141-154. doi:10.5127/pr.045115 van den Hout, M. A., Engelhard, I. M., Rijkeboer, M. M., Koekebakker, J., Hornsveld, H., Leer, A., Toffolo, M. B. J., & Akse, N. (2011*b*). EMDR: Eye movements superior to beeps in taxing working memory and reducing vividness of recollections. *Behaviour Research and Therapy, 49*(2), 92-98. doi:10.1016/j.brat.2010.11.003

van den Hout, M. A., Engelhard, I. M., Smeets, M. A. M., Horsnveld, H., Hoogeveen, E., de Heer, E., Toffolo, M. B. J., & Rijkeboer, M. M. (2010). Counting during recall: Taxing of working memory and reduced vividness and emotionality of negative memories. *Applied Cognitive Psychology, 24*, 303-311. doi:10.1002/acp.1677

van den Hout, M., Muris, P., Salemink, E., & Kindt, M. (2001). Autobiographical memories become less vivid and emotional after eye movements. *British Journal of Clinical Psychology, 40*, 121-130.

van den Hout, M. A., Rijkeboer, M. M., Engelhard, I. M., Klugkist, I., Hornsveld, H., Toffolo, M. J. B., & Cath, D. C. (2012). Tones inferior to eye movements in the EMDR treatment of PTSD. *Behaviour Research and Therapy, 50*, 275–279. doi:10.1016/j.brat.2012.02.001

van der Does (2002). BDI-II-NL. Handleiding. De Nederlandse versie van de Beck Depression Inventory-2nd edition. Lisse: Harcourt Test Publishers

van der Ploeg, H. M., Defares, P. B., & Spielberger, C. D. (2000). Handleiding bij de zelfbeoordelingsvragenlijst. *Een Nederlandstalige bewerking van de Spielberger State-Trait Anxiety Inventory [Manual for the State-Trait Anxiety Inventory. A Dutch translation]. Lisse, The Netherlands.*

van der Ploeg, E., Mooren, T., Kleber, R. J., van der Velden, P. G., & Brom, D. (2004). Construct validation of the Dutch version of the impact of event scale. *Psychological assessment*, *16*(1), 16.

Vandierendonck, G., De Vooght, K., & Van der Goten, A. (1998). Interfering with the central executive by means of a random interval repetition task. *The Quarterly Journal of Experimental Psychology: Section A, 51*(1), 197-218 .doi:10.1080/713755748

van Schie, K., van Veen, S. C., Engelhard, I. M., Klugkist, I., & van den Hout, M. A. (2016). Blurring emotional memories using eye movements: Individual differences and speed of eye movements. *European Journal of Psychotraumatology, 7*. doi: 10.3402/ejpt.v7.29476.

van Schie, K., van Veen, S. C., Hagenaars, M. A., & Engelhard, I. M. (2017). Exploring effects of dual-tasks on reduction of intrusive memories following analogue trauma

van Veen, S. C., Engelhard, I. M., & van den Hout, M. A. (2016). The effects of eye movements on emotional memories: Using an objective measure of cognitive load. *European Journal of Psychotraumatology, 7*. doi:10.3402/ejpt.v7.30122 van Veen, S. C., van Schie, K., Wijngaards-de Meij, L. D., Littel, M., Engelhard, I. M., & van den Hout, M. A. (2015). Speed matters: relationship between speed of eye movements and modification of aversive autobiographical memories. *Frontiers in Psychiatry, 6*. doi:10.3389/fpsyt.2015.00045

Vervliet, B., & Raes, F. (2013). Criteria of validity in experimental psychopathology: application to models of anxiety and depression. *Psychological medicine, 43*(11), 2241-2244.

Verwoerd, J., de Jong, P. J., & Wessel, I. (2008). Low attentional control and the development of intrusive memories following a laboratory stressor. *Journal of Psychopathology and Behavioral Assessment, 30*(4), 291-297. doi:10.1007/s10862-008-9080-6

Verwoerd, J., Wessel, I., de Jong, P. J., & Nieuwenhuis, M. M. (2009). Preferential processing of visual trauma-film reminders predicts subsequent intrusive memories. Cognition and Emotion, 23(8), 1537-1551. doi:10.1080/02699930802457952

Verwoerd, J., Wessel, I., de Jong, P. J., Nieuwenhuis, M. M., & Huntjens, R. J. (2011). Pre-stressor interference control and intrusive memories. Cognitive therapy and research, 35(2), 161-170. doi:10.1007/s10608-010-9335-x

Vincken, M. J., Meesters, C., Engelhard, I. M., & Schouten, E. (2012). Psychometric qualities of the White Bear Suppression Inventory in a Dutch sample of children and adolescents. *Personality and Individual Differences*, *52*(3), 301-305. doi:10.1016/j.paid.2011.10.023

Walker, M. P., Brakefield, T., Hobson, J. A., & Stickgold, R. (2003). Dissociable stages of human memory consolidation and reconsolidation. *Nature, 425*(6958), 616-620. doi:10.1038/nature01930

Wang, Y., Cao, Z., Zhu, Z., Cai, H., & Wu, Y. (2015). Cue-independent forgetting by intentional suppression–Evidence for inhibition as the mechanism of intentional forgetting. *Cognition*, *143*, 31-35. doi:10.1016/j.cognition.2015.05.025

Wang, S. H., de Oliveira Alvares, L., & Nader, K. (2009). Cellular and systems mechanisms of memory strength as a constraint on auditory fear reconsolidation. *Nature Neuroscience, 12*(7), 905-912. doi:10.1038/nn.2350

Weathers, F. W., Litz, B. T., Keane, T. M., Palmieri, P. A., Marx, B. P., & Schnurr, P. P. (2013). The ptsd checklist for dsm-5 (pcl-5). Scale available from the National Center for PTSD at www. ptsd. va. gov.

Wegner, D. M., Schneider, D. J., Carter, S. R., & White, T. L. (1987). Paradoxical effects of thought suppression. *Journal of Personality and Social Psychology, 53*, 5–13. doi:10.1037/0022-3514.53.1.5

Weidmann, A., Conradi, A., Gröger, K., Fehm, L., & Fydrich, T. (2009). Using stressful films to analyze risk factors for PTSD in analogue experimental studies–which film works best?. *Anxiety, Stress, & Coping, 22*(5), 549-569. doi: 10.1080/10615800802541986

Wenzlaff, R. M., & Wegner, D. M. (2000). Thought suppression. *Annual review of psychology, 51*(1), 59-91. doi:10.1146/annurev.psych.51.1.59.

Wetzels, R., & Wagenmakers, E. J. (2012). A default Bayesian hypothesis test for correlations and partial correlations. *Psychonomic Bulletin & Review, 19*(6), 1057-1064.doi: 10.3758/s13423-012-0295-x

Weidmann, A., Conradi, A., Gröger, K., Fehm, L., & Fydrich, T. (2009). Using stressful films to analyze risk factors for PTSD in analogue experimental studies–which film works best?. *Anxiety, Stress, & Coping, 22*(5), 549-569. doi: 10.1080/10615800802541986

Wichert, S., Wolf, O. T., & Schwabe, L. (2011). Reactivation, interference, and reconsolidation: Are recent and remote memories likewise susceptible?. *Behavioral Neuroscience, 125*(5), 699. doi: 10.1037/a0025235

Wichert, S., Wolf, O. T., & Schwabe, L. (2013a). Changing memories after reactivation: A one-time opportunity?. *Neurobiology of Learning and Memory, 99*, 38-49. doi:10.1016/j.nlm.2012.11.001

Wichert, S., Wolf, O. T., & Schwabe, L. (2013b). Updating of episodic memories depends on the strength of new learning after memory reactivation. *Behavioral Neuroscience, 127*(3), 331. doi:10.1037/a0032028

Wilson, S. A., Becker, L. A., & Tinker, R. H. (1995). Eye movement desensitization and reprocessing (EMDR) treatment for psychologically traumatized individuals. *Journal of Consulting and Clinical Psychology, 63*, 928-937.

Winters, B. D., Tucci, M. C., & DaCosta-Furtado, M. (2009). Older and stronger object memories are selectively destabilized by reactivation in the presence of new information. *Learning & Memory, 16*(9), 545-553. doi:10.1101/lm.1509909

Wood, N. E., Rosasco, M. L., Suris, A. M., Spring, J. D., Marin, M. F., Lasko, N. B., ... & Pitman, R. K. (2015). Pharmacological blockade of memory reconsolidation in posttraumatic stress disorder: Three negative psychophysiological studies. *Psychiatry research, 225*(1), 31-39. doi:10.1016/j.psychres.2014.09.005

Wulff, K., Gatti, S., Wettstein, J. G., & Foster, R. G. (2010). Sleep and circadian rhythm disruption in psychiatric and neurodegenerative disease. *Nature Reviews Neuroscience, 11*(8), 589-599. doi:10.1038/nrn2868

Yalch, M. M. (2016). Applying Bayesian statistics to the study of psychological trauma: A suggestion for future research. *Psychological Trauma: Theory, Research, Practice, and Policy, 8*(2), 249-257. doi:10.1037/tra0000096

Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of memory and language, 46*(3), 441-517. doi:10.1006/jmla.2002.2864

Zvolensky, M. J., Lejuez, C. W., Stuart, G. L., & Curtin, J. J. (2001). Experimental psychopathology in psychological science. *Review of General Psychology, 5*(4), 371.

Zwaan, R. A., & Pecher, D. (2012). Revisiting mental simulation in language comprehension: Six replication attempts. *PLoS ONE, 7*(12), e51382. doi:10.1371/journal.pone.0051382

DUTCH SUMMARY

NEDERLANDSE SAMENVATTING

Verandering van Ongewenste Herinneringen
Een Beschouwing van Drie Methoden

Vrijwel iedereen maakt op een gegeven moment in zijn of haar leven een traumatische gebeurtenis mee. Hoewel een meerderheid hier goed mee om kan gaan, ontwikkelt een substantiële minderheid (tussen 5.6% en 22%) een posttraumatische stressstoornis (PTSS) in de maanden na deze gebeurtenis. PTSS wordt onder andere gekenmerkt door herbeleving van het trauma (bijv. in de vorm van intrusies of nachtmerries), vermijding van zaken die doen denken aan de gebeurtenis, negatieve verandering in stemming en verandering in arousal en reactiviteit. Er zijn voor PTSS psychologische behandelingen beschikbaar, maar die zijn helaas niet voor iedereen even effectief. Een deel van de patiënten verbetert niet als gevolg van de behandeling (en is behandelresistent) of valt terug na een initieel succesvol lijkende behandeling. Er is dus ruimte voor verbetering in de behandeling van PTSS.

Een aantal wetenschappers stelt dat verbetering in psychologische behandeling bereikt kan worden door onderzoek te richten op de vraag *waarom* iets werkt, en niet uitsluitend op *of* iets werkt. Zodra we weten welk mechanisme onderliggend is aan een effectieve interventie kunnen we proberen om deze interventie naar onze hand te zetten en aan te passen aan individuele verschillen tussen mensen; een stap in de richting van gepersonaliseerde psychotherapie. Een manier om te onderzoeken waarom bestaande en nieuwe interventies werken is door een experimentele benadering toe te passen op psychopathologie. Het doel van experimentele psychopathologie is om middels gecontroleerde (lab)experimenten oorzaak-gevolgrelaties bloot te leggen om op deze manier het ontstaan en de behandeling van psychopathologie beter te begrijpen. In dit proefschrift wordt deze experimentele psychopathologie benadering toegepast op drie verschillende interventies die tot doel hebben om ongewenste herinneringen te veranderen, namelijk: (1) het belasten van het werkgeheugen door duale-taken, (2) het verstoren en updaten van het geheugen tijdens reconsolidatie, en (3) het onderdrukken van ongewenste intrusieve herinneringen.

Sectie 1: het belasten van het werkgeheugen door duale-taken

Een eerste interventie om ongewilde, aversieve herinneringen te veranderen is door doelbewust het naarste beeld van de herinnering (de hotspot) op te halen en tegelijkertijd een duale-taak uit te voeren die het werkgeheugen belast, zoals het maken van horizontale oogbewegingen (bijv. door de vinger van de therapeut te volgen of – in het lab – een bewegende stip op een beeldscherm). Deze duale-taak interventie wordt in de klinische praktijk gebruikt in de behandeling Eye Movement Desensitization and Reprocessing

(EMDR). Als gevolg van de duale-taak interventie neemt de hotspot af in hoe naar en helder deze ervaren wordt. De werkgeheugentheorie biedt een verklaring voor dit effect; het tegelijkertijd maken van oogbewegingen en het ophalen van de hotspot zorgt voor competitie om de beperkte bronnen van het werkgeheugen. Het gevolg is dat het ophalen van de hotspot wordt belemmerd en deze daardoor daalt in helderheid en onaangenaamheid. Daarna wordt deze nieuwe, minder heldere herinnering opgeslagen in het langetermijngeheugen in de plaats van de oude, veel heldere herinnering. Uit deze werkgeheugentheorie vloeien een aantal veronderstellingen voort die in de hoofdstukken in deze sectie zijn getoetst.

Een veelvoud aan experimenten laat inmiddels zien dat een duale taak, zoals het maken van oogbewegingen (of een andere taak die het werkgeheugen voldoende belast), in combinatie met het ophalen van een herinnering zorgt voor dalingen in helderheid en/of onaangenaamheid ten opzichte van het alleen ophalen van een herinnering. De experimenten in **hoofdstukken 3, 4** en **6** laten deze dalingen ook zien en dit sterkt het idee dat competitie in het werkgeheugen tot dit effect leidt. De afwezigheid van deze effecten in **hoofdstuk 2** en **5** is mogelijk het gevolg van de opzet van de experimenten: de manier waarop meerdere herinneringen opgehaald moesten worden (**hoofdstuk 2**) of de afwezigheid van een standaard controleconditie (**hoofdstuk 5**).

Voorts bieden **hoofdstukken 3** en **4** aanvullend bewijs voor een werkgeheugenverklaring; experimenten in beide hoofstukken laten onafhankelijk van elkaar zien dat een hogere snelheid van oogbewegingen (hogere mate van belasting van het werkgeheugen door de duale-taak) samengaat met grotere dalingen in helderheid en/of onaangenaamheid. Er is echter ook afwezigheid van bewijs voor bepaalde hypotheses in deze hoofdstukken en dat oogt in tegenspraak met de werkgeheugenverklaring. Zo onderzochten we hier of de effectiviteit van een duale-taak vergroot kan worden door deze aan te passen aan individuele verschillen. De werkgeheugentheorie voorspelt dat dit mogelijk is omdat de mate van competitie in het werkgeheugen afhankelijk is van drie elementen, namelijk (1) de belasting van de herinnering, (2) de belasting van de duale-taak en (3) de werkgeheugencapaciteit van een individu. De werkgeheugentheorie voorspelt dat personen met een kleine werkgeheugencapaciteit meer profiteren van langzame oogbewegingen, terwijl personen met grote werkgeheugencapaciteit meer profiteren van snelle oogbewegingen. Eenzelfde voorspelling gaat op voor laaglevendige herinneringen (die minder belastend zijn op het werkgeheugen) en hooglevendige herinneringen (die meer belastend zijn op het werkgeheugen). Echter, in geen van beide experimenten vinden wij bewijs dat aanpassingen aan deze karakteristieken de interventie effectiever maakten.

In **hoofdstukken 5** en **6** onderzochten we of duale-taak interventies ook een effect hebben op het (onvrijwillig) herbeleven van een traumatische ervaring. Herbeleving is één van de kernsymptomen van PTSS en kan tot uiting komen in herhalende, zich opdringende herinneringen die intrusies worden genoemd. Uit onderzoek bij PTSS patiënten is inmiddels bekend dat een meerderheid van de intrusies gerelateerd is aan de hotspot. Aangezien duale-taken de helderheid en/of onaangenaamheid van de hotspot verminderen, veranderen ze daarmee mogelijk ook de prikkelbaarheid van de hotspot en het aantal intrusies dat daaruit voortvloeit. Een belangrijke vraag hierbij is hoe veranderingen in helderheid en onaangenaamheid uiteindelijk leiden tot afnames in PTSS symptomen. Voorlopige resultaten van twee experimenten in **hoofdstuk 5** laten zien dat duale-taak interventies van korte duur geen effect hebben op het aantal intrusies. Echter, duale-taak interventies van lange duur, ongeacht de modaliteit van de duale-taak (visuospatieel vs. verbaal), verminderen het aantal intrusies wel. Een nog uit te voeren derde experiment moet uitsluitsel geven of het effect specifiek komt door het ophalen van de herinnering gelijktijdig met het uitvoeren van de duale-taak of dat het uitsluitend komt door het ophalen van de herinnering. Verder doen de experimenten in dit hoofdstuk vermoeden dat algemene belasting en modaliteitspecifieke belasting van het werkgeheugen beide een rol spelen in het verminderen van intrusies, daar de verbale taak ietwat meer belastend was voor het werkgeheugen dan de visuospatiële taak maar zij hetzelfde effect hadden op het aantal intrusies.

In **hoofdstuk 6** is getoetst of het ophalen van de herinnering een noodzakelijk onderdeel is om intrusies te verminderen. Eerdere onderzoeken laten namelijk zien dat enkel het uitvoeren van een visuospatiële taak (bijv. het spelen van Tetris) zonder het ophalen van de herinnering ook minder intrusies tot gevolg heeft. Tegen verwachtingen in vinden wij in ons experiment dat het spelen van Tetris met óf zonder het ophalen van een herinnering beide niet tot minder intrusies leidt. Bovendien zijn dalingen in helderheid en/of onaangenaamheid niet voorspellend voor de hoeveelheid intrusies die deelnemers ervaren. Dit roept vragen op over de repliceerbaarheid van het 'Tetris-effect', maar ook over de relatie tussen veranderingen in hoe de hotspot ervaren wordt en de hoeveelheid intrusies die een persoon ervaart. Het is mogelijk dat deze veranderingen in helderheid en aangenaamheid van de hotspot niet noodzakelijkerwijs een passende reflectie zijn van therapeutische verandering zoals bijvoorbeeld in de hoeveelheid intrusies die een persoon ervaart.

Sectie 2: het verstoren en updaten van het geheugen tijdens reconsolidatie

Wanneer iemand een traumatische gebeurtenis meemaakt, zorgen consolidatieprocessen ervoor dat deze gebeurtenis wordt opgeslagen in ons brein. Tot voor kort werd aangenomen dat deze opslag permanent was en de herinnering bijgevolg onveranderbaar. Recentelijk is bekend geworden dat het reactiveren, ofwel het ophalen, van een geconsolideerde herinnering ervoor kan zorgen dat deze herinnering terugkeert naar een labiele staat voordat de herinnering weer wordt heropgeslagen in het langetermijngeheugen (reconsolidatie). Tijdens dit reconsolidatieproces kunnen gereactiveerde herinneringen verstoord of aangepast worden door gedragsinterventies toe te passen. In dit proefschrift zijn dit gedragsinterventies waarin nieuwe informatie geleerd wordt tijdens het reconsolidatieproces.

In **hoofdstuk 7** voerden we een replicatie uit van een eerdere studie die bewijs toonde voor geheugenverandering tijdens het reconsolidatieproces. In **hoofdstuk 8** hebben we onderzocht of het veranderen van een sterk geconsolideerde herinnering afhankelijk is van de sterkte van de interventie tijdens reconsolidatie. In de experimenten in beide hoofdstukken vinden we echter geen bewijs dat herinneringen uitsluitend veranderd kunnen worden tijdens het reconsolidatieproces. Opmerkelijk was dat we vonden dat geconsolideerde herinneringen veranderden als het gevolg van het leren van nieuw informatie *ongeacht* of dit vooraf ging door reactivatie. Dit duidt niet op veranderingen in een herinnering als gevolg van processen die spelen tijdens reconsolidatie; dit duidt op retroactieve interferentie. Het leren van iets nieuws zorgt simpelweg voor verandering in iets ouds. Deze bevindingen staan in contrast met een groot aantal onderzoeken dat bewijs laten zien voor verandering van herinneringen tijdens reconsolidatie. De twee experimenten in dit proefschrift staan niet op zichzelf; er zijn inmiddels steeds meer onderzoeken die geen bewijs kunnen vinden voor het reconsolidatieproces. Dit duidt erop dat (geheugenverandering tijdens) reconsolidatie nauw luistert en mogelijk eenvoudig beïnvloedt wordt door andere variabelen die het optreden van het effect kunnen beletten, zoals de leeftijd van de herinnering, de mate van angst die de persoon ervaart of de aanwezigheid van nieuwe informatie die de informatie in de oude herinnering update.

De afwezigheid van bewijs voor reconsolidatie doet twijfelen aan de robuustheid van dit proces, maar biedt ook kansen om systematisch te onderzoeken wanneer het proces wel en niet optreedt. Hier moet ook rekening gehouden worden met de klinische toepasbaarheid. Tot op heden wordt namelijk in veel onderzoek naar reconsolidatie gebruik gemaakt van farmacologische interventies, terwijl in psychotherapie voor PTSS voornamelijk gebruik wordt gemaakt van gedragsinterventies (bijv. exposure, imagery rescripting). Toekomstig onderzoek dient zich derhalve te richten op gedragsinterventies die herinneringen kunnen veranderen tijdens reconsolidatie waarbij gekeken dient te worden waarom verandering tijdens reconsolidatie niet optreedt en hoe dit mogelijk te overkomen is.

Sectie 3: het onderdrukken van ongewenste intrusieve herinneringen

De derde interventie die we onderzocht hebben is het stoppen of onderdrukken van het ophalen van een herinnering (ook wel *retrieval stopping* of *retrieval suppression* genoemd). Dit stelt individuen in staat om het ophalen van een ongewenste herinnering te voorkomen of het ophalen te onderbreken. Het vermogen dit te doen is vaak bestudeerd in contexten waarin personen worden blootgesteld aan een stimulus die hen onmiddellijk herinnert aan deze ongewenste herinneringen. Dit is vooral relevant in het licht van PTSS. Cognitieve theorieën over PTSS stellen namelijk dat stimuli (bijv. objecten, personen, etc.) geassocieerd worden met de traumatische gebeurtenis en dat wanneer deze stimuli later tegengekomen worden, ze regelmatig zorgen voor het – bijna reflexief – ophalen van het trauma. In **hoofdstuk 9** en **10** is bekeken of dit ophaalproces gestopt kon worden en wat de effecten zijn op het onthouden van de herinnering en de intrusiviteit van deze herinnering.

In **hoofdstuk 9** is onderzocht of de mogelijkheid om ongewilde herinneringen te onderdrukken afhankelijk is van de valentie (neutraal vs. negatief) van wat men probeert te onderdrukken en de strategie die hiervoor gebruikt wordt. In dit hoofdstuk laten we zien dat zowel neutrale als negatieve herinneringen onderdrukt kunnen worden wanneer gebruik wordt gemaakt van een 'directe suppressie' strategie. Wanneer deze strategie gebruikt wordt probeert men het bewustzijn leeg te houden van de ongewilde herinnering. Tevens wordt de ongewilde herinnering niet vervangen voor iets anders. In tegenstelling tot een andere veelgebruikte strategie 'gedachtenvervanging' (waar er wel aan iets anders gedacht wordt), doet directe suppressie een beroep op het kernmechanisme van inhibitie. Inhibitie in deze context verwijst naar het stoppen of onderbreken van het ophalen van een specifieke herinnering. Gedachtenvervanging doet een beroep op interferentie en zorgt er daardoor voor dat de *relatie* tussen stimulus en herinnering wordt verstoord. Dat betekent dus dat een andere stimulus nog steeds de traumatische herinnering kan oproepen, aangezien die relatie nog steeds onaangetast is. Inhibitie daarentegen zorgt ervoor dat de ongewilde *herinnering* in het geheugen ontoegankelijk wordt gemaakt en vergeten wordt. Het gevolg is dus dat ook andere stimuli de traumatische herinnering slecht kunnen oproepen, omdat de herinnering zelf geïnhibeerd is en niet slechts één specifieke relatie.

In **hoofdstuk 10** onderzochten we of directe suppressie niet enkel leidt tot het vergeten van de ongewenste herinnering, maar of het er ook voor zorgt dat er een afname is in intrusies, het ongewild denken aan de herinnering. We onderzochten of dit afhankelijk is van hoe lang een ongewenste herinnering onderdrukt moest worden. We vonden dat het aantal intrusies afnam met het aantal keren dat de ongewenste herinnering onderdrukt werd. Kortom, met oefening wordt men beter in het onderdrukken. Daarnaast vonden we dat personen meer intrusies ervoeren wanneer ze lange tijd achter elkaar een ongewenste herinnering moesten onderdrukken (vergeleken met een korte tijd).

Hoewel het vergeten van herinneringen en het verminderen van intrusies aan deze herinnering door gedachtenonderdrukking mogelijk is, zijn er nog een aantal belangrijke onbeantwoorde vragen in het licht van klinische toepasbaarheid. Is gedachtenonderdrukking te trainen, en zo ja, is het een effect dat behouden blijft op de lange termijn? Er is bewijs dat het onderdrukken op kortetermijn te trainen is, maar er is er nog onduidelijkheid over de langetermijneffecten. Daarnaast is er de vraag of het door iedereen op ieder moment aangewend kan worden om ergens niet aan te denken. Er is immers grote variabiliteit in de mate waarin personen beschikken over inhibitievermogen. Hoogangstige personen en PTSS patiënten lijken over een kleiner inhibitievermogen te bezitten. Of dit een risicofactor is voor het ontwikkelen van angststoornissen of trauma- en stress gerelateerde stoornissen of dat het eerder een gevolg is, behoeft nog een antwoord. Tot dat daar meer duidelijkheid over is kan gedachtenonderdrukking het best ingezet worden door diegenen zonder tekorten in hun inhibitievermogen.

Een link tussen methoden

Hoewel er in dit proefschrift ogenschijnlijk afzonderlijke methoden zijn gebruikt, zijn ze allen sterk gerelateerd aan een aantal fenomenen en concepten, zoals geheugenverandering, emotieregulatie en vergeten. Tot op zekere hoogte overlappen deze methoden en waar overlap is kan vruchtbare grond zijn voor nieuw onderzoek. Het belasten van het werkgeheugen door duale-taken veronderstelt bijvoorbeeld dat bij het bewerkstelligen van langetermijneffecten reconsolidatieprocessen betrokken moeten zijn. Reconsolidatie is mogelijk ook betrokken bij gedachtenonderdrukking (of vertoont op z'n minst parellellen met dit onderzoek). Het blijkt namelijk zo te zijn dat het vergeten van ongewilde herinneringen het meest effectief gedaan kan worden als de ongewilde herinnering kortstondig in het bewustzijn aanwezig is. Dit veronderstelt dat, net als bij reconsolidatie, (re)activatie van een herinnering deze meer vatbaar maakt voor verandering. Inhibitieprocessen die betrokken zijn bij gedachtenonderdrukking zijn mogelijk ook betrokken bij de belasting van het werkgeheugen door duale taken. Het is namelijk denkbaar dat het belasten van het werkgeheugen beter verklaard kan worden als het continue en snel afwisselen van twee verschillende taken (het ophalen van de herinnering en de afleidende taak). Om succesvol op de ene taak te focussen dient de ander (tijdelijk) geïnhibeerd te worden.

Naast overlap in de mechanismen zijn er ook andere raakvlakken. Vrijwel alle interventies hebben bijvoorbeeld een effect op de *toegankelijkheid* van de herinnering en niet per se op de *beschikbaarheid*. Toegankelijkheid reflecteert met welk gemak informatie uit het geheugen gehaald kan worden. Beschikbaarheid heeft betrekking op of de informatie überhaupt aanwezig is in het geheugen. Tevens is er convergentie in hoe de verschillende methoden bepalen of er geheugenverandering heeft plaatsgevonden (bijv. het aantal onthouden of vergeten items; een schatting van hoe onaangenaam een herinnering is; het aantal intrusies) en de mogelijk betrokken hersengebieden (de prefrontale cortex en de amygdala).

Mogelijke klinische implicaties

De experimenten in dit proefschrift beschrijven relatief fundamenteel werk aangaande het veranderen van ongewenste herinneringen. Desalniettemin is het onderwerp klinisch zeer relevant en zijn er een aantal klinische implicaties te benoemen; sommige daarvan zijn direct toepasbaar in het werkveld, andere zijn suggesties voor verder onderzoek. Uit **hoofdstuk 3** en **4** volgt duidelijk dat het verhogen van de snelheid van oogbewegingen de effectiviteit van de interventie vergroot. Het voorlopige bewijs uit **hoofdstuk 5** wekt de suggestie dat duale-taak interventies ook in staat zijn om (intrusieve) herbeleving van een ongewenste herinnering te verminderen.

De resultaten uit **hoofdstuk 7** en **8** naar het verstoren en updaten van herinneringen tijdens geheugenreconsoldiatie suggereren dat het gepast is een pas op de plaats te maken voordat interventies gebaseerd op het reconsolidatieproces worden ingezet in de klinische praktijk. Een optie is om in een gecontroleerde labomgeving te onderzoeken wanneer geheugenverandering tijdens reconsolidatie optreedt (eventueel in een klinische populatie). Er moet hierbij rekening gehouden worden met kenmerken die inherent zijn aan de klinische populatie(s) die we hopen te behandelen, maar die verandering in de weg kunnen staan, zoals sterkte van een herinnering, individuele angstgevoelens en stress.

De experimenten uit **hoofdstuk 9** en **10** laten zien dat ongewenste herinneringen vergeten kunnen worden en dat intrusies verminderd kunnen worden door gedachtenonderdrukking. Dit lijkt klinisch toepasbaar, maar is niet aan te raden aangezien PTSS patiënten en hoogangstige individuen inhibitietekorten vertonen waardoor gedachtenonderdrukking kan falen.

Conclusie

Dit proefschrift laat zien dat ongewenste herinneringen veranderd kunnen worden en dat verschillende interventies ingezet kunnen worden om dit doel te bereiken. Het sterkste bewijs voor geheugenverandering komt uit de onderzoeken waarin duale-taak interventies zijn ingezet of waarin er gebruik wordt gemaakt van gedachtenonderdrukking. De resultaten van duale-taak interventies zijn veelbelovend en leiden tot heldere klinische implicaties (bijv. hogere snelheid van oogbewegingen tijdens EMDR). Wat betreft gedragsinterventies (het leren van nieuwe informatie) tijdens reconsolidatie is ons onderzoek strijdig met de grote hoeveelheid bewijs die geheugenverandering laat zien tijdens reconsolidatie. Hoewel gedragsinterventies tijdens reconsolidatie implicaties kunnen hebben voor de klinische praktijk, is het noodzakelijk om eerst beter te begrijpen wanneer (geheugenverandering tijdens) reconsolidatie optreedt, wanneer dat niet het geval is en hoe niet-veranderde herinneringen mogelijk toch veranderd kunnen worden. De klinische toepasbaarheid van gedachtenonderdrukking is vooralsnog beperkt gezien de onduidelijkheid over de trainbaarheid, langetermijneffecten en algemene toepasbaarheid. Desalniettemin hebben deze interventies duidelijke implicaties voor toekomstig onderzoek, wat op termijn mogelijk kan leiden tot verbetering van de effectiviteit van behandeling.

ACKNOWLEDGMENTS

DANKWOORD

Dankwoord

Het starten, intensief werken en het (uiteindelijk) voltooien van een PhD is niet alleen het gevolg van eigen doorzettingsvermogen, maar ook van een dosis geluk en van de mensen die je in je leven leert kennen. Mensen die jou beter maken, die jou iets gunnen en die deuren voor je openen of nieuwe perspectieven bieden op de wereld die je dacht te kennen. Die mensen wil ik hier bedanken, want zonder hen was alles anders.

Marcel, wat was het een eer en een voorrecht om met jou samen te werken. Ik ben je dankbaar voor veel dingen; vooral voor je aanstekelijke enthousiasme over de wetenschap, de vrijheid die me gegeven hebt (ook om soms helemaal vast te lopen) en de helpende hand die je me – wanneer nodig – bood (ons meester-gezel principe). Ik heb ook ontzettend genoten van al onze wandelingen in de velden waar we mooi konden brainstormen over nieuw onderzoek of over verklaringen voor huidig onderzoek. Naast je ijzersterke wetenschappelijke kant ben ik ook blij je als mens te hebben leren kennen. Nu, terugkijkend op de afgelopen vier jaar, kan ik me niet voorstellen hoe mijn promotie zou zijn geweest zonder jouw sterke verhalen, onze voortdurende uitwisseling van taalweetjes en jouw oprechtheid en humor. Je hebt op mij grote indruk gemaakt en ik hoop ik ook een beetje op jou. Hopelijk blijft je altijd bij waar de woorden fauteuil, silhouet en framboos vandaan komen, en dat je 'bij dezen' altijd met een –n zult schrijven.

Iris, wat heb ik veel van jou geleerd, in het bijzonder wetenschappelijk schrijven. Onder jouw begeleiding leerde ik dat wetenschappelijk schrijven een kunstvorm is die tot op het woord uitgeoefend kan worden. Jouw oog voor detail en scherpe feedback hebben mij veel geleerd en daardoor zijn een aantal van mijn artikelen kleine (ambachtelijke) meesterwerkjes geworden. Daarnaast ben ik je ook heel dankbaar voor je persoonlijke benadering en je interesse voor de persoon achter de promovendus. Ik kon en kan altijd bij je terecht voor (persoonlijk) advies. We hebben het ook nooit uitsluitend over de wetenschap, maar ook over 'de wereld en alles wat daarin gebeurde' of over mijn leven buiten de wetenschap. Vrijwel iedere afspraak begon dan ook met 'Ha Kevin, hoe gaat het thuis?'. Overigens zal ik ook nooit vergeten dat jij de eerste persoon bent die mij (altijd) met 'Ha' aansprak waardoor ik dat heb overgenomen en nu dagelijks gebruik.

Suzanne, aanstichter van de wolkenspotterij, paranimf en mijn wetenschappelijke wederhelft, wat zou mijn PhD tijd geweest zijn zonder jou? In het begin gingen we wat stroef, maar al snel veranderden wij in een goed geoliede machine die tot op vandaag hard draait (net als een ander befaamd wetenschappelijk duo). Er zijn maar weinig dingen die we *niet* samen hebben gedaan. We hebben talloze symposia bezocht (met New York en Boston als hoogtepunten), vele cursussen gevolgd en vele – maar dan ook vele – (onderzoeks)projecten opgezet en daarin ons 'good cop-bad cop' rollenspel geperfectioneerd (wanneer het aan kwam op studentbegeleiding). Van onze projecten waren er een aantal een groot succes (project schaap, het brainstormweekend) en andere aanzienlijk minder (over DJ Tony hebben we het nooit meer). De wetenschap was voor ons een drug die zorgde voor een high waardoor we op momenten niet meer wisten of we nou Suzanne of Kevin waren, of beide? Ik stel voor dat we die drug nog vaak gebruiken, want in tegenstelling tot andere drugs heeft deze alleen maar positieve effecten.

Sophie, medewolkenspotter, geharde Pokémonjager en paranimf. In jou vond ik in tijden een medestander om niet-werkgerelateerd gedrag te vertonen: we kletsten wanneer we hadden moeten werken, soms klaagden we een beetje en we speelden spelletjes. Dat alles ging altijd des te beter met een stukje donkere chocolade met sinaasappel. Dat we op ongeveer hetzelfde moment promoveren is een cadeautje. Zonder de pomodoros in het laatste jaar was alles veel langzamer gegaan (voor ons beide, vermoed ik) en zonder jou was ik waarschijnlijk de helft van de zaken van een promotie (en vele andere dingen in de afgelopen jaren) volledig vergeten te regelen. Wat ben ik blij dat jij er op de dag ook bij bent en naast me staat; het Utrechtse basketbalteam is zonder jou niet compleet.

Lieve, beste, toffe en fijne collega's, jullie hebben werken op de Universiteit Utrecht en specifiek op onze afdeling vele malen leuker gemaakt. Ik ben blij dat ieder van jullie voor langere of kortere tijd aanwezig is geweest tijdens mijn promotietraject. In het bijzonder wil ik mijn EPP familie bedanken: Angelos, Annemarie, Arne, Daniëlle, Dieuwke, Eliane, Elze, Evi-Anne, Gaëtan, Ilse, Karin, Katharina, Marianne, Marieke T, Marthe, Mitzy, Muriel, Nicole, Puck en Suzy, evenals studentassistenten Angelique, Kiki, Laurijn en Marloes. Bedankt voor de dagelijkse lunch, voor alle kopjes "koffie", alle gesprekken, borrels, uitjes, (inter)nationale congressen, promotiefeesten, summer schools, symposia in Heeze, en alle keren dat ik jullie kamer binnen mocht lopen voor mijn vraag, klaag of – okay, ik geef toe – imitatie. Ik ga jullie missen! Natuurlijk wil ik ook mijn andere (oud)-collega's van de K(G)P afdeling bedanken voor jullie betrokkenheid op welke manier dan ook: Alishia, Amal, Astrid, Aukje, Cate, Henk, Joris, Lot, Maarten, Manik, Marieke M, Marieke S, Marlies, Ninke, Paul, Ria, Sibe, Sosja, Tamara, Tracy en Zhongfang. Allemaal, ontzettend bedankt.

Een aantal collega's (naast promotoren en paranimfen) wil ik hier nog specifiek benoemen en bedanken. Mijn (oud)kamergenoten Marieke T, Arne, Marianne en Evi-Anne. Marieke T, voor je immer opgewekte karakter, je adviezen en je 'alles komt altijd goed' mentaliteit. Een mooie tegenhanger voor mijn 'komt het nog wel goed met deze studenten?'. De dubbele duimen omhoog ga ik nooit meer vergeten.

Arne, senioraio, en dubbelganger (alle andere aanduidingen laat ik omwille van de fatsoenlijkheid even buitenbeschouwing) bedankt voor je altijd scherpe brainstormsessies, maar ook voor alle keren dat we naar de sportschool gingen (en gaan), alle spellen die we hebben gespeeld, biertjes die we hebben gedronken en natuurlijk Leuven (x2). We hebben alles wel zo'n beetje gedaan wat gezond is om kamergenoten samen te laten doen, behalve misschien het schrijven van een artikel; begin jij alvast?

Marianne, in het begin was je mijn kamergenoot nog niet, maar was je de ervaren postdoc op het project. Mede door veel van jouw voorwerk (o.a. de 'helse' taken in Open Sesame) kon ik snel aan de slag met het echte werk. Niet alleen heb ik dankbaar voortgebouwd op jouw experimenten, maar heb ik ook de 'Marianne-techniek' van je overgenomen (in principe 'good cop-bad cop', maar dan uitgevoerd door één persoon). Tot slot, ik heb sterk het vermoeden dat ik dankzij jouw profetie in je afscheidsbriefje weer terug ben op de Erasmus Universiteit Rotterdam.

Evi-Anne, met jou als kamergenoot had ik eindelijk iemand die ook 'kattenkwaad' wilde uithalen en dat maakte me zeer gelukkig. Het gevolg was dat Arne's hoofd voor ons een geliefd doelwit werd, zowel letterlijk (stressballen) als figuurlijk (labs die hem ineens ontnomen werden door jouw nep e-mails). Ook andere collega's waren hun leven niet meer zeker door jouw snode plannen; met als hoogtepunt mijn zeer vroege (doch geniale) beursaanvraag in ruil voor drie chocoladerepen. Natuurlijk hebben we niet alleen maar kattenkwaad uitgehaald tijdens kostbare werktijd, we hebben gelukkig ook belangrijke dingen gedaan zoals het maken van de GOLD challenge; werken kan tenslotte altijd nog.

Angelos, als ik jou in het begin van mijn aio-tijd niet begeesterd had zien presenteren over Bayesiaanse statistiek was ik hier waarschijnlijk nooit zelf mee aan de slag gegaan (*There is the wrong way and the Bayesian way.*) Vervolgens betekende dat wel dat ik (toen jij als postdoc in Utrecht kwam werken) soms met statistische vragen bij je aanklopten, die je vaak direct voor me kon oplossen. Jouw hulp kent werkelijk geen grenzen. You R amazing!

Muriel, omdat ik altijd bij je binnen mocht lopen met al mijn vragen en ingevingen over het traumafilmparadigma (ook als het eigenlijk niet uitkwam). Suzy, voor je eindeloze enthousiasme en natuurlijk voor de tijd in Metz.

Elke, het is alweer enige tijd geleden dat we elkaar gesproken en gezien hebben, maar dit dankwoord zou niet compleet zijn zonder jou te bedanken. Hoewel ik "slechts" jouw onderzoeksassistent was opende je deuren en gaf je mij grote vrijheid. Je bracht mij in contact met Michael en onder jouw coachende leiding mocht ik mijn eigen onderzoek uitvoeren. Dat deed ik en dat opende mijns inziens weer een nieuwe deur, namelijk het mogen doen van een PhD bij Marcel en Iris. Dank voor jouw oneindig vertrouwen in mij.

Michael, I would also like to thank you. You provided me with guidance and help in setting up experiments even when I had not graduated yet, which eventually resulted in a first authored pre-PhD paper. You helped me acquire a scientific way of thinking and you took my thinking of experimental design to a whole new level. I am grateful for all that I have learned from working together with you and for those people that I have met through you. I hope that we will continue working together for a long time.

Karima, jij in het bijzonder bedankt voor alle labruimtes die je (steeds weer) voor me reserveerde, terwijl ik natuurlijk al lang klaar had moeten zijn met mijn onderzoeken. Christiaan en Marcel, zonder jullie hulp had ik sommige onderzoeken nooit kunnen uitvoeren. Ik wilde vaak van alles in mijn experimenten programmeren wat me zelf veelal niet lukte. Gelukkig hebben jullie mij hier vaak uit de brand kunnen helpen. Dit proefschrift had er heel anders uitgezien zonder jullie hulp. Ontzettend bedankt!

Beste coauteurs zonder jullie fijne en vaak rappe samenwerking was het nooit gelukt om al mijn artikelen te schrijven. Uiteraard wil ik ook alle studenten bedanken die betrokken zijn geweest bij het verzamelen van de data voor de onderzoeken, maar natuurlijk ook de studenten, vrienden en familie die een deel van hun tijd geïnvesteerd hebben om proefpersoon te zijn voor één van de onderzoeken; door jullie hulp of deelname zijn er data en is er dus een proefschrift!

De onderzoeksschool EPP bedankt voor alle mooie symposia en de EPP opleidingscommissie die het met mij aandurfde om een deel van de symposia door mij te laten organiseren. Anna, James, Bart, en Bruno, ik ben blij dat ik jullie op deze manier heb leren kennen. Promovendi Netwerk Utrecht (Prout) bedankt voor alle gezelligheid en de activiteiten die we georganiseerd hebben. ZonMw en de Faculteit Sociale Wetenschappen voor de financiering van mijn aanstelling.

Lieve, beste en fantastische vrienden, hoewel ik een groot deel van mijn (werkende) leven onderzoekend op de universiteit heb doorgebracht was het goed om soms mijn aandacht even volledig weg te nemen van onderzoek en van het proefschrift. Jullie waren hierin de spil en ieder van jullie heeft op zijn of haar eigen unieke manier daaraan bijgedragen. Om te ontspannen, maar ook om nieuwe ideeën te vormen, moet je tenslotte af en toe je gedachten verzetten en mooie dingen doen die helemaal niets met de wetenschap te maken hebben. Laten we daar altijd mee doorgaan!

Mijn (schoon)familie – in het bijzonder mijn ouders Peter & Marja – wil ik hier noemen en bedanken voor hun steun en continue interesse. Pap en mam, ook al wisten jullie soms niet wat me precies bezighield, dank voor jullie trots. De vrijheid die jullie mij gegeven hebben om mijn eigen weg te vinden en me te ontwikkelen tot wie ik ben is ongekend. Veel van jullie mooie eigenschappen zoals jullie oplossingsgerichte houding in combinatie met hard werken om je doelen te behalen, hebben mij in het leven zeker geholpen, maar in het bijzonder in de afgelopen vier jaar.

Lieve Bas, mijn grote liefde, er zijn woorden tekort om je te bedanken voor al je geduld, je begrip en de ruimte die je mij gegeven hebt om gepassioneerd met de wetenschap bezig te zijn. In de jaren dat we elkaar kennen, leverde je een onmisbare bijdrage aan mijn leven met je mooie gevoel voor humor, de hulp die me bood als ik vastliep of ergens over twijfelde, en met de wens om samen met mij de rest van de wereld (buiten de wetenschap) te ontdekken. Ik hoop dat je dit (en nog vele andere dingen) nog lange tijd voor me en met me wilt doen. Ik kijk heel erg uit naar de rest van ons leven samen.

ABOUT THE AUTHOR

Curriculum vitae

Kevin van Schie was born on 27 March 1987 in Heemstede, the Netherlands. In 2011 he obtained his master‘s degree in Clinical Psychology and Brain & Cognition (cum laude) at Erasmus University Rotterdam, while simultaneously completing the Advanced Research Program. After obtaining his master’s degree he worked as an academic teacher (tutor) and junior researcher. Under supervision of dr. Elke Geraerts and in close collaboration with prof. dr. Michael Anderson from the University of Cambridge, he began investigating suppression-induced forgetting and thought suppression. In 2013 he started his PhD project under supervision of prof. dr. Marcel van den Hout and prof. dr. Iris Engelhard at Utrecht University, where he investigated working mechanisms (i.e., taxing working memory & memory change during reconsolidation) underlying Eye Movement Desensitization and Reprocessing. During this time he was also an active member of the Experimental Psychopathology Graduate School’s Educational Committee and the PhD Network Utrecht. In August 2017 he returned to his alma mater where he now is assistant professor in clinical psychology.

International publications

van Schie, K., van Veen, S. C., Hendriks, Y. R., van den Hout, M. A., & Engelhard, I. M. (2017). Intervention strength does not differentially affect memory reconsolidation of strong memories, *Neurobiology of Learning and Memory, 144*, 174-185.

van Schie, K., van Veen, S. C., van den Hout, M. A., & Engelhard, I. M. (2017). Modification of episodic memories by novel learning: a failed replication study. *European Journal of Psychotraumatology, 8*(sup1), 1315291.

van Schie, K., & Anderson, M. C. (2017). Successfully controlling intrusive memories is harder when control must be sustained. *Memory 25*(9), 1201-1216.

Littel, M., **van Schie, K.**, & van den Hout, M. A. (2017). Exploring expectation effects in EMDR: does prior treatment knowledge affect the degrading effects of eye movements on memories? *European Journal of Psychotraumatology 8*(sup1), 1328954.

van Schie, K., van Veen, S. C., Engelhard, I. M., Klugkist, I., & van den Hout, M. A. (2016). Blurring emotional memories using eye movements: individual differences and speed of eye movements. *European Journal of Psychotraumatology, 7*(1), 29476.

van Schie, K., Wanmaker, S., Yocarini, I., & Bouwmeester, S. (2016). Psychometric qualities of the thought suppression inventory-revised in different age groups. *Personality And Individual Differences, 91*, 89-97.

van Schie, K., Engelhard, I. M., & van den Hout, M. A. (2015). Taxing working memory during retrieval of emotional memories does not reduce memory accessibility when cued with reminders. *Frontiers in Psychiatry, 6*, 16.

van Veen, S. C., **van Schie, K.**, Wijngaards-de Meij, L. D., Littel, M., Engelhard, I. M., & van den Hout, M. A. (2015). Speed matters: relationship between speed of eye movements and modification of aversive autobiographical memories. *Frontiers in Psychiatry, 6*, 45.

van Schie, K., Geraerts, E., & Anderson, M. C. (2013). Emotional and non-emotional memories are suppressible under direct suppression instructions. *Cognition & Emotion, 27*(6), 1122-1131.

Moors, A., De Houwer, J., Hermans, D., Wanmaker, S., **van Schie, K.**, Van Harmelen, A., De Schryver, M., De Winne, J., & Brysbaert, M. (2013). Norms of valence, arousal, dominance, and age of acquisition for 4300 Dutch words. *Behavior Research Methods, 45*(1), 169-177.

Publications in Dutch

van Schie, K., & van den Hout, M. A. (in press). Expres vergeten. *De Psycholoog*.

van Schie, K. & van Veen S. C. (2017). Snelheid doet er toe: snelheid van oogbewegingen en individuele verschillen bij het vervagen van emotionele herinneringen. *Tijdschrift voor Gedragstherapie, 1*, 21-36.

Submitted manuscripts

van Schie, K., Kessler, H., van den Hout, M. A., & Engelhard, I. M. (submitted). Reducing intrusive memories: the necessity of deliberate recall during taxation of working memory

Conference presentations & Invited presentations

van Schie, K., van Veen, S. C., Hagenaars, M. A., van den Hout M. A., & Engelhard, I. M. (2017). The impact of taxing working memory on the modulation of intrusive memories. *Oral presentation at Annual convention of Association for Psychological Science (APS), Boston, United States.*

van Schie, K., van Veen, S. C., Hagenaars, M. A., van den Hout M. A., & Engelhard, I. M. (2017). Kunnen oogbewegingen intrusieve herinneringen verminderen? *Oral presentation at the Dutch EMDR Conference, Utrecht, The Netherlands.*

van Schie, K. (2017). Changing memories we would rather not have. *Oral presentation at the EPP Research Day of the Dutch-Flemish postgraduate school 'Experimental Psychopathology' (EPP). Utrecht, The Netherlands.*

van Schie, K. (2017). How do eye movements in EMDR change memories? *Invited presentation for a Research Colloquium at Maastricht University, Maastricht, The Netherlands.*

van Schie, K., van Veen, S. C., Hendriks, Y. van den Hout, M. A., & Engelhard, I. M. (2016). How do different doses of behavioral interventions affect strong memories during reconsolidation? *Poster presentation at Conference of the European Association for Behavioural and Cognitive Therapies (EABCT), Stockholm, Sweden.*

van Schie, K., van Veen, S. C., van den Hout, M. A., & Engelhard, I. M. (2016). Modification of episodic memories by novel learning: a replication study. *Oral presentation at Conference of the European Association for Behavioural and Cognitive Therapies (EABCT), Stockholm, Sweden.*

van Schie, K., van Veen, S. C., van den Hout, M. A., & Engelhard, I. M. (2016). Modification of episodic memories during memory reconsolidation: a replication study. *Oral presentation at the 6th International Conference on Memory (ICOM), Budapest, Hungary.*

van Schie, K., van Veen, S. C., Klugkist, I., Engelhard, I. M., & van den Hout, M. A. (2016). Individual differences in working memory capacity do not affect the blurring of aversive memories using eye movements. *Invited oral presentation for Séminaire Eye Movement Doctoral Research, Metz, France.*

van Schie, K., van Veen, S. C., Klugkist, I., Engelhard, I. M., & van den Hout, M. A. (2016). Het vervagen van emotionele herinnering door oogbewegingen: individuele verschillen en de snelheid van oogbewegingen, *Oral presentation at the Dutch NtVP Conference. Lunteren, The Netherlands.*

van Schie, K. (2016). The effects of working memory capacity and speed of eye movements on blurring emotional memories using eye movements. *Oral presentation for Battle of the Brains competition at the EPP Research Day of the Dutch-Flemish postgraduate school 'Experimental Psychopathology' (EPP). Utrecht, The Netherlands.*

van Schie, K., Van Veen, S. C., Hendriks, Y., Engelhard, I. M., & Van den Hout, M. A. (2015). Changing reactivated strong memories with strong interference. *Poster presentation at the Summer school on Emotional Learning and Memory in Health and Psychopathology, Leuven, Belgium.*

van Schie, K., Van Veen, S. C. Klugkist, I., Engelhard, I. M., & Van den Hout, M. A. (2015). Snelheid van oogbewegingen aanpassen op de werkgeheugencapaciteit, *Oral presentation at the Dutch EMDR Conference. Nijmegen, The Netherlands.*

van Schie, K., van Veen, S. C., Engelhard, I. M., & van den Hout, M. A. (2015). Blurring aversive memories by dual tasking. *Poster session presented at Annual convention of Association for Psychological Science (APS), New York, United States.*

van Schie, .K (2015). Reconsolidation of human memory: fighting strong with strong? *Oral presentation for Battle of the Brains competition at the EPP Research Day of the Dutch-Flemish postgraduate school 'Experimental Psychopathology' (EPP). Utrecht, The Netherlands.*

van Schie, K., Engelhard, I. M., & van den Hout, M. A., (2015). Eye movements during retrieval do not reduce objective memory accessibility when cued with reminders, *Poster presentation at the International Convention of Psychological Science (ICPS). Amsterdam, The Netherlands.*

van Schie, K., Engelhard, I. M., & van den Hout (2014). Eye movements reduce accessibility of emotional memories. *Poster presentation at the 7th Helmholtz Retreat. Schoorl, The Netherlands.*

Awards & Prizes

Second place Article Prize. EPP research day 2016 of the Dutch-Flemish postgraduate school 'Experimental Psychopathology' (EPP). The Netherlands.

Second place Presentation in Battle of the Brains Competition. EPP research day 2016 of the Dutch-Flemish postgraduate school 'Experimental Psychopathology' (EPP). The Netherlands.

Travel Grant International Convention of Psychological Science. Amsterdam, the Netherlands