ARE EMOTIONAL MEMORIES HARDER TO INTENTIONALLY FORGET? A META-ANALYSIS

by © Kelsi Hall

A Thesis submitted to the School of Graduate Studies

in partial fulfillment of the

Requirements for the degree of Master of Science in Experimental Psychology

Department of Psychology

Memorial University of Newfoundland

2020

St. John's Newfoundland and Labrador

Abstract

The current meta-analysis explored whether emotional memories are less susceptible to item-method directed forgetting than neutral memories. The final sample used consisted of 31 studies with 36 independent samples. Basic analyses revealed superior memory for remember (R) than forget (F) items, with directed forgetting effects observed for both neutral, $M = 19.6\%$, $CI_{95\%}$ [16.1%, 23.1%], and emotional, $M = 15.1\%$, $CI_{95\%}$ [12.4%, 17.7%], conditions. Directed forgetting in either valence condition was larger for (a) words than complex stimuli; (b) recall than recognition tests; (c) studies that used recall prior to recognition; and, (d) studies that included buffer items. Comparison within-studies revealed diminished directed forgetting of emotional items compared to neutral, with an average difference of 4.2%, $CI_{95\%}$ [2.0%, 6.4%]. However, this finding varied, meaning that whether – and to what degree – emotional memories are more resilient than neutral memories depends on the methodological features of the study in question. Larger differences were present in studies where emotional items were more arousing than neutral and when buffer items were included. These findings suggest that emotional memories are (often) more resilient to intentional forgetting than neutral memories.

General Summary

The aim of this thesis is to determine whether emotional memories are more difficult to forget than non-emotional memories. To this end, I reviewed all published research conducted applying the item-method directed forgetting task to emotional items. This task entails participants being presented with a series of items such as words, images, etc., each followed by an instruction to either remember (R) or forget (F) that item. After being tested for all items, participants typically remember more R items than F items; this is called the directed forgetting effect. I found less directed forgetting for emotional items on average in the studies compared to neutral items. The difference was influenced by study characteristics such that a larger difference between emotion conditions was found in studies that used more arousing emotional items than neutral items, and for studies that included buffer items at the beginning and end of their study phase. Also, overall for both item types, the effect of directed forgetting was larger in studies that used: (a) single word stimuli compared to other stimuli including images, faces and events; (b) recall tests instead of recognition; (c) when a recall test was used prior to recognition testing; and, (d) when buffer items were included at the beginning and end of the study phase. Overall, this thesis suggests that people have a harder time forgetting emotional memories than neutral memories.

Table of Contents

List of Tables

List of Figures

Are Emotional Memories Harder to Intentionally Forget? A Meta-Analysis

Memory is an important aspect of daily life, allowing us to recall important details such as birthdays, phone numbers, and even important life events. However, forgetting is also essential as it allows us to update our memory to function more efficiently (Anderson & Hanslmayr, 2014; Fawcett & Hulbert, 2020). For example, when you are trying to remember where you parked your car in a busy parking lot, it is easier if you have forgotten all the previous spots you have parked (Bjork, 1970). However, recent evidence suggests that some memories are less easily forgotten than others (e.g., Hauswald et al., 2010). Emotional experiences in particular appear to be more resilient in memory than neutral experiences (Bailey & Chapman, 2012; Hauswald et al., 2010; Nowicka et al., 2011; Yang et al., 2016) – and this resilience seems to apply even when the experience in question is unwanted (e.g., Butler & James, 2010). For example, reoccurring memories of traumatic events are a symptom of some psychiatric disorders such as posttraumatic stress disorder (PTSD; American Psychiatric Association, 2013). However, not all researchers agree that there is a difference in the degree to which emotional and neutral study items are affected by intentional forgetting (e.g., Gallant & Yang, 2014; Marchewka et al., 2016; Yang et al., 2012). To address this debate, the present thesis aims to further investigate the ability to control emotional memories using directed forgetting, as it will clarify the mixed pattern of results found previously and broaden our understanding of intentional forgetting.

Intentional Forgetting

The process of intentionally removing unwanted information from memory, known as intentional forgetting, has been studied in the laboratory using a variety of memory control paradigms such as item-method and list-method directed forgetting (Macleod, 1999), the think no-think paradigm (Anderson & Green, 2001), and retrieval-induced forgetting (Anderson et al., 1994). These paradigms highlight different mechanisms that can lead to the suppression of unwanted information, often resulting in its relative inaccessibility. Due to their relevance to the present thesis, I will begin with an overview of each, with an emphasis on item-method directed forgetting, which is the focus of the present meta-analysis.

Item-Method Directed Forgetting

In an item-method paradigm, participants are presented with a series of items, such as images or words (Quinlan et al., 2010), and each item is followed by an instruction to either remember (R) the item for a future test or to forget (F) the item as it will not be tested further. Following presentation of the study items, participants are tested on their memory for all items, including those they had been instructed to forget, using either recall (e.g. Lee & Hsu, 2013; Moulds & Bryant, 2008) or recognition (e.g. Quinlan & Taylor, 2014; Gallant, 2013). Participants typically remember more R-items than F-items; this is known as the directed forgetting effect (Johnson, 1994; MacLeod, 1998). With respect to the mechanisms involved, it has historically been theorized that participants initially engage in maintenance rehearsal to keep the item in working memory, awaiting the memory instruction (Macleod, 1975). If an R instruction is presented, participants work to encode the item further through elaborative rehearsal; if an F instruction is presented, they stop rehearsing the item (Anderson & Hansylmayr, 2014; Basden & Basden, 1998; Jing et al., 2019; Johnson, 1994).

There has been some debate as to the exact processes that lead to the removal of F-items from the rehearsal set, allowing the selective rehearsal of R-items. This has been supported in part by studies showing that participants take longer to respond to a visual probe following F-items than R-items, providing evidence of one or more active processes following the presentation of an F instruction that require cognitive effort, therefore slowing reaction to the probe (Fawcett &

Taylor, 2008; Fawcett et al., 2013b; but see, Lee 2018; Scholz & Dutke, 2019). This has also been replicated in studies investigating inhibition of return (IOR), where the study item serves as the cue for the task. In that paradigm, participants take longer to respond to targets presented at the same location as previous F-items, as compared to R-items (Thompson et al., 2014; Thompson & Taylor, 2015). Together, these findings suggest that active processes work to withdraw processing resources away from the item following a forget instruction.

Evidence from neuroimaging studies have supported similar conclusions, showing activation of frontal brain regions during enactment of an F instruction (Wylie et al., 2008; see Anderson & Hanslmayr, 2014, for a review). For example, Rizio and Dennis (2013) found activation of a right-lateralized frontoparietal control network to be association with successful, intentional forgetting, whereas incidental forgetting was instead associated with activity in left-lateralized brain regions thought to support encoding. This lends credibility to the argument that intentional and incidental forgetting rely on separate underlying processes (for a review, see Anderson & Hanslmayr, 2014; Rizio & Dennis, 2013; Wierzba et al., 2018; Wylie et al., 2008).

Notably, others disagree with such an active account of forgetting and instead have argued that intentional forgetting is a natural consequence of selectively rehearsing the R-items while passively excluding the F-items from consideration (Basden & Basden, 1996). In line with this theory, Sahakyan and Foster (2009) found when analyzing participants' explanations of the strategies they used during F trials that some participants indicated that they rehearsed previous R items to help in forgetting the F-item. It has been speculated that it is this rehearsal that might slow responses to secondary tasks following an F instruction. Notably, this theory was refuted by Fawcett et al. (2013b), who replicated the finding of slower probe responses following F than R instructions even on the first trial of the study phase, when no R-items would be available for rehearsal. In support of the selective rehearsal theory, Lee and Lee (2011) varied post-cue intervals in the directed forgetting task and found long intervals lead to greater memory for F-items. However, this was reduced using a secondary counting task. The authors suggest when the post-cue interval is long following the F instruction, participants have a harder time ignoring the item and begin to automatically process it, leading to higher memory for F-items compared to the short interval group.

This theory has since been supported with evidence from pupil data. Lee (2018) investigated the role of spatial attention in the item-method directed forgetting task using eye tracking to determine both the role of looking behaviour and cognitive effort as indicated by changes in pupil size. An auditory memory cue was delivered during visual presentation of the study word. Their analysis of the eye tracking data revealed that participants were more likely to look away from the words following a forget and ignore cue, indicating a shift of spatial attention. Pupil size was larger following remember cues than forget or baseline "ignore" cues (for which participants were instructed to ignore the words), indicating greater cognitive effort being used to remember items than to forget or ignore them (for a replication, see Scholz & Dutke, 2019). These findings conflict with the active account of forgetting but support the selective rehearsal account as they indicate no active processes are being used following the F instruction. This remains an area of active theoretical development.

List-Method Directed Forgetting

Whereas item-method directed forgetting permits exploration of how we disengage from information we have just processed, list-method directed forgetting involves pushing an entire set of information out of mind in favour of learning a new set. Although both list-method and item-method directed forgetting were once thought to share the same underlying mechanisms (Bjork, 1972), more recent theorists disagree (Basden et al, 1993). List-method directed forgetting involves presenting participants with an entire list of items, then instructing them to either remember or forget that list before presenting them with the next list, which they are instructed to remember. Typically, this task uses a between-subjects design where half of the participants are told to remember both lists. When tested using recall for all items, memory for remember lists is typically superior to memory for forget lists (Macleod, 1999; Pastötter et al., 2016). List-method directed forgetting is often described in terms of costs and benefits, with costs as the reduction in recall after participants are told to forget List 1, compared to the benefits of increased recall for List 2 which the participants are instructed to remember. This is in comparison to the remember all condition, where participants are instructed to remember both lists (Sahakyan & Kelley, 2002). However, the effect is typically not found, or is severely reduced when recognition testing is used (e.g. Pastötter et al., 2016; Macleod, 1999). Historically this phenomenon was theorized to be a result of retrieval inhibition, which was thought to prevent participants from recalling words from the F list. According to that theoretical perspective, learning List 2 required that List 1 be inhibited, decreasing the probability of recalling these items at test (Basden et al., 1993; Macleod, 1999). Recognition testing was thought to release List 1 from inhibition (Basden et al., 1993).

Later, research has suggested that context plays a role in this task (Sahakyan & Kelley, 2002). According to this idea, when participants are told to forget List 1, there is a shift in internal context prior to the presentation of List 2. This context shift makes it more difficult for participants to recall List 1 items at test, as their internal context while recalling the items is mismatched relative to their prior internal context while studying List 1, preventing the use of context cues to aid their memory (Hanczakowski et al., 2012; Sahakyan & Kelley, 2002). Sahakyan and Kelley (2002) emphasized the role of context in the paradigm by using a modification that enhanced the context shift following the initial list. This modification entailed a phase where participants envisioned what life would be like if they were invisible. The context shift modification led to worse recall for List 1 compared to the control group regardless of whether participants were instructed to remember or forget that list prior to studying List 2. In a follow-up experiment, they investigated the role of context reinstatement by instructing participants to reflect back to the context under which they studied List 1 prior to the test phase. Their results indicated that reinstating the previous context led to enhanced recall of List 1, apart from the control group which found no difference in recall due to this manipulation. These results support the importance of context in the task.

Think/No-Think

The next paradigm to be discussed is think/no-think (TNT), which was designed to be a lab analog of attempting to supress unwanted memories (such as the loss of a loved one) from coming to mind and interfering with everyday life. In this task, participants are presented with cue-target pairs (e.g., Roach-Ordeal, Grass-Tent, Cold-Lamp) which they are instructed to learn for a later phase. These pairs are studied until participants can recall the target item when given the cue according to a prespecified performance criterion (e.g., 75% correct). Next, during the think/no-think phase, participants are presented with a series of cues, some of which they are instructed to bring the associated target to mind (think trials; e.g., Roach), and others they are instructed to keep the associated target from coming to mind (no-think trials; e.g., Grass). Participants are subsequently tested for memory of all items. A subset of items from the learning phase that were not included in the think/no-think phase serve as a baseline for memory performance (e.g. Cold-Lamp; Anderson & Green, 2001). Typically, memory for no-think targets (e.g. Tent) is worse compared to baseline targets (e.g. Lamp; Depue et al., 2006; Anderson & Green, 2001). This finding is referred to as suppression-induced forgetting (SIF).

The dominant theoretical framework used to explain SIF in this paradigm points to the engagement of one or more inhibitory processes. This account suggests that a no-think cue enacts one or more inhibitory control processes used to block the target item from coming to mind and push it from mind should it emerge (Anderson & Green, 2001; Anderson & Levy, 2009; Anderson & Hanslmayr, 2014). These processes make no-think targets (e.g. Grass) less likely to be recalled at test. In a typical task, during the test phase, participants are tested on their memory for the targets using the same cues as the preceding phases, which is referred to as same probe testing (e.g. Anderson & Green, 2001). However, perhaps the strongest evidence supporting this perspective derives from work using independent probes at test (e.g. testing for Lamp using Dark). For example, Anderson and Green (2001) found that memory for no-think targets was impaired during the testing phase, even when the probes used were independent of those used in the previous phases. From this, those researchers inferred that memory for the target itself is suppressed, not just the association between the target and the probe.

Further evidence in support of the inhibitory account comes from neuroimaging studies, which have again implicated a frontoparietal control network (similar to that described when discussing item-method directed forgetting) engaged during no-think trials. This has predominantly involved increased activity in the dorso-lateral prefrontal cortex (DLPFC), a region also associated with stopping motor responses (Anderson & Hanslmayr 2014). In the context of the think/no-think paradigm, the DLPFC is thought to prevent targets from coming to mind during no-think trials while also downregulating hippocampal activity, leading to worse memory for those items at test (Anderson et al., 2004). Although research in this area is still developing, one explanation as to how memories become suppressed in this manner points toward a mechanism akin to the nonmonotonic plasticity hypothesis proposed by Newman and Norman (2010). According to this hypothesis, memories compete for the focus of attention, becoming activated to varying degrees depending on available cues. However, it is important that the mind remains orderly, meaning that our brains should favour a system biased toward whatever memory is "winning" that competition, rather than offering them equal fielding. As a result, retrieval is thought to follow a process of biased competition such that competitors become relatively destabilized insofar as they interfere with retrieval. Specifically, this theory posits that when a competing memory is strongly or weakly activated at retrieval, their representations are unaffected – because they either fail to compete or are too strongly activated to be controlled effectively. However, when a memory is moderately activated, meaning that it is competing for – but failing to achieve – the focus of attention, it becomes downregulated (for an application of this theoretical framework to the think/no-think paradigm see, Detre et al., 2013).

However, not all evidence points toward the role of inhibitory processes, and an alternate account is that SIF is caused instead by interference. This is the idea that on no-think trials participants bring something else to mind in order to avoid thinking of the unwanted target and therefore forming a new association with the cue. When prompted with the cue at test, the new item interferes with recall of the old target (e.g. Racsmány et al., 2011). Tomlinson et al. (2009) used a computational model of interference and argued that cue-independent forgetting in the think/no-think paradigm – previously attributed to inhibition – could be explained by interference. These authors posit that cue-independent forgetting in that paradigm is a result of participants learning to associate the cue with something else on no-think trials to avoid thinking of the related target, therefore supressing memory for the original target. However, the inhibition account is still the favored explanation of the SIF effect. For example, Wang et al., (2015) found that encouraging participants to use an interference-based strategy (e.g., distracting themselves with an alternate word) during no-think trials reduced recall for only same-probes whereas a suppression-based strategy instead reduced recall for both same-probes and independent probes. This demonstrates that while both accounts can be used to demonstrate forgetting during the task, only the inhibition account can support forgetting for the items while using independent probes.

Retrieval-Induced Forgetting

The final intentional forgetting paradigm to be discussed is retrieval-induced forgetting (RIF). In this task, categorized word lists are learned (e.g. fruit-orange, fruit-apple), where each category is associated with a number of exemplars. During a practice phase, participants are cued to recall some, but not all, exemplars associated with a certain category (e.g. fruit-or). To successfully recall the correct exemplar (e.g. orange), participants must suppress the competing exemplars (e.g. apple). These are typically less likely to be remembered when tested for all items compared to items from baseline categories which are not included in the practise phase (Anderson et al., 1994).

Aligned with the think/no-think paradigm, the mechanisms within this paradigm are often thought to rely heavily on inhibitory processes (Anderson et al., 1994; Weller et al, 2013). This is supported by the fact that words suppressed during the task are still unlikely to be remembered when using independent cues during test (e.g. Weller et al., 2013). As with the think/no-think task, this paradigm can also be explained by the nonmonotonic plasticity hypothesis, as it introduces competition between targets associated with the same cue. For example, Kereztes and Racsmany (2013) found that items that were recalled with moderate reaction times during the retrieval practice phase, compared to those recalled very quickly or slowly, were more likely to be forgotten during the task. They posit that when items are recalled very quickly it is unlikely that the use of control processes will be necessary to stop competing targets from coming to mind, as there is a strong relationship in memory between the cue and target. When items are recalled very slowly, there is little effect of RIF. However, when items are recalled moderately quickly, it is more likely control processes will be elicited to reduce competition between the targets as the relationship between the cue and prompted target is weaker, leading to subsequent forgetting of competing targets.

Memory for Emotional Stimuli

Whereas the preceding sections discuss the cognitive mechanisms involved in controlling unwanted or outdated information, not all information is equal in this respect; some types of information tend to be more difficult to forget than others. More specifically, research has shown that emotional memories seem to be particularly difficult to intentionally forget (e.g., Hauswald et al., 2010). Expansive research has investigated whether directed forgetting specifically is affected by the emotional nature of the stimuli used, producing mixed results (Taylor et al., 2018; Yang et al., 2016); therefore, further investigation is needed to determine the true impact. Research on memory for emotional items suggests they have a priority in our cognitive systems (e.g. Kensinger & Corkin, 2003), and therefore indicating that they may be more difficult to intentionally forget. The following sections will give an overview of research on the effect of emotion on memory and our cognitive systems before getting into how emotion affects directed forgetting specifically.

However, prior to commencing my overview, it is worth pointing out that the term ‘emotion’ tends to be poorly operationalized in most studies. For example, many studies conflate arousal and valence and fail to consider the underlying the array of emotions that comprise the valence categories (positive and negative) and the differing effects the specific emotions may elicit (e.g., Kissler & Herbert, 2013; Minor & Herzmann, 2019; Richardson et al., 2004). Dominance is another dimension of emotion referring to whether participants feel in control or controlled by the stimuli, however, dominance is often highly correlated with valence (Scott et al., 2018, Warriner et al., 2013) and is therefore less researched. Ignoring the multifaceted nature of emotions can mislead researchers as different aspects of emotions can have differing effects on our cognitive processes. For example, within the 'negative' category emotions such as anger and disgust differ in the strength of response elicited by participants in both attention (Liu et al., 2015) and memory (Ferré et al., 2017). Also, anger tends to capture attention quicker than happiness (Feldmann-Wüstefeld et al., 2010). When these specific emotional components are not considered, key elements underlying any effects found are not isolated. This adds noise to the field of research as variation due to certain components may be unreported. As most literature in the area tends to not specify which categories of emotion are being manipulated, most of this review will refer to emotion more generally in its comparison to neutral stimuli.

An enhancement in memory for emotional material has been well established in the literature (see Crowley et al., 2019 for a review). Memory for both negative and positive words has been shown to be enhanced when compared to neutral words (Adelman & Estes, 2013); this enhancement extends to the context in which the word is presented (Kensinger & Corkin, 2003). Minor and Herzmann (2019) aimed to determine how emotion influences both item and source memory for images, but also determine the neural correlates involved by monitoring event-related potentials (ERP). They presented participants with two blocks of intermixed neutral (e.g., a farm animal) and negative (e.g., an injured arm) images and after 24 hours tested participants on both item memory for the images as well as source memory for the block in which they were presented. In terms of behavioural results, they found better item memory for negative images, as negative images were recognized more accurately than neutral images. However, in terms of source memory, they found no difference due to valence. Their ERP results showed that for images recognized with the correct source, stronger old/new effects over the left hemisphere were found for negative images, indicating an enhanced representation in memory. Old/new effects are an implicit indication of whether a participant recognizes an item as previously studied (Wilding, 2000). This indicates that emotional items are not only typically better remembered, but additional details also tend to be better represented.

From a neuroscience perspective, the amygdala is thought to play a key role in the enhancement of emotional items over neutral items in memory (see Hamann, 2001 for a review). Strange et al. (2004) isolated the unique roles played by the amygdala and hippocampus in emotional memory by inhibiting the connection between the two regions using a drug while also monitoring using functional magnetic resonance imaging (fMRI) as participants learned neutral and emotional words for a later test. Compared to the drug group, which was administered a drug to block the beta-adrenergic system, memory of the emotional words in the placebo group was associated with greater activity in the amygdala during encoding and subsequent activation of the hippocampus during recall. This highlights the role of the amygdala in remembering emotional stimuli, specifically in relation to noradrenaline and its influence on the hippocampus. Richardson et al. (2004) further investigated the role of the amygdala and hippocampus in emotional memory by presenting participants who have hippocampal damage, with or without damage to the amygdala, with neutral and emotional items followed by a recognition test. Hippocampal damage was associated with reduced memory for all items, regardless of emotional condition. Damage to the amygdala was associated with decreased memory for only emotional items. Similarly, Segal et al. (2012) found increased levels of noradrenaline after participants viewed negative images, which was later associated with enhanced ability for participants to recognize previously seen neutral images compared to similar neutral images. This line of research highlights the importance of the amygdala and its connection to the hippocampus in emotional memory.

Research on emotion also suggests that emotional information is attended to and processed differently than neutral information (Bradley et al., 2003; Cuthbert et al., 2000; Schupp et al., 2003). Lexical decisions are made more rapidly for emotional words, indicating speeded processing when compared to neutral words (Kissler & Herbert, 2013). Similarly, memory for words presented rapidly on screen followed by a mask have been shown to be greater when the words are emotional (Zeelenberg et al., 2006). Faster detection of emotional material is theorized to occur due to its evolutionary value, as quickly processing threats is exceptionally important to survival (Ohman, et al., 2001; LoBue & Matthews, 2013). Similar results have occurred for positive stimuli as well. For example, Brosch et al. (2008) found newborn faces led to faster detection of a subsequent probe presented in the same location than did neutral adult faces. Also, attention tends to be drawn to emotional stimuli, which has been shown to negatively impact performance when it is irrelevant to the task at hand (Hindi Attar & Müller, 2012). This enhanced processing also extends to top-down processes such as the participants' perception of the items. Lee et al. (2010) used fearful and neutral faces with varying levels of noise over the images, along with full-noise images to investigate changes in ERP activity while viewing each image. They showed participants the faces one at a time and had participants indicate whether the faces were fearful or neutral. They found that when participants indicated they viewed a fearful face compared to a neutral one, there was an enhanced response in regions associated with motivational significance. This further demonstrates how the human cognitive system prioritizes emotion.

The experience of emotion is comprised of two components, valence and arousal. The former refers to the affective state elicited by the stimuli (e.g. pleasant or unpleasant) and the latter refers to the level of intensity or excitement elicited by the stimuli (Lang et al., 1993). These components are thought to have different influences on cognitive function. Kensinger and Schacter (2006) investigated the differing effects of arousal and valence on neural activity by presenting participants with a series of negative, positive, and neutral images and words while using fMRI. Arousal was associated with activity in the dorsal medial PFC, amygdala, and the ventromedial PFC, which have all been shown to be involved in processing emotion-evoking stimuli (e.g., Dolcos et al., 2004; Strange et al., 2004). Valence had differing effects for each emotion category; negative stimuli were associated with activity in the lateral PFC, whereas positive stimuli were associated with activity in the medial PFC which has been shown to be related to reward (Kim et al., 2010). Dolcos et al. (2004) found similar results with differences in arousal being associated with more activation in the dorsomedial PFC while differences in valence were associated with increased activity in the ventromedial PFC. This indicates that arousal and valence are processed distinctly; the former is associated with regions that process emotional stimuli, and the latter is associated with regions related to reinforcement.

Some research has suggested that arousal has a greater influence on memory than valence (e.g. Kensinger & Corkin, 2004; Dolcos et al., 2004). Kensinger and Corkin (2004) found that memory for non-arousing negative words (e.g., sorrow) was associated with activity in the left PFC suggesting the use of rehearsal strategies, whereas memory for arousing negative words (e.g., slaughter) was associated with activity in both the amygdala and hippocampus, suggesting a more automatic process. In a follow-up experiment, they used a divided attention manipulation to further disentangle the effect of arousal and valence on memory. The memory enhancement for the negative non-arousing words compared to neutral words diminished when participants' attention was divided during word presentation, confirming the reliance on rehearsal strategies requiring attention. The arousing negative condition still led to enhanced memory when attention was divided, supporting the automaticity of memory for arousing stimuli. Szőllősi and Racsmány (2020) demonstrated the differential impact of arousal and valence in a discrimination task where participants were presented with old, new, or similar negative, positive, and neutrally valenced images, with high and low arousal conditions for the emotional categories. Discrimination for both emotional categories was greater compared to neutral, with greater discrimination for negative items compared to positive. When comparing the roles of valence and arousal, high arousing emotional images in both conditions led to better discrimination compared to their low arousing counterparts. They suggest the enhancement of memory representations due to arousal allowed for the reduction in interference between old and similar images needed to complete the task successfully. Contrary to this claim, Adelman and Estes (2013) found that when arousal was controlled for, valence independently influenced memory performance, with better memory for words at both extremes of valence. However, they found no effect of arousal on memory in their study, but this finding is only reflective of word stimuli; arousal and valence have been shown to have differing effects depending on the type of stimuli used (Sutton & Lutz, 2019).

Overall, when compared to neutral stimuli, emotional stimuli tend to be processed faster (e.g. Kissler & Herbert, 2013), capture attention more easily (e.g. Hindi Attar & Müller, 2012), and lead to enhanced memory (e.g. Adelman & Estes, 2013). However, these processes are affected by the level of valence and arousal of the stimuli used and some research suggests arousal plays a greater role in this enhancement (e.g. Szőllősi & Rascmány, 2020). The combination of both the cognitive and neurophysiological responses to emotional stimuli as described above closely relate to the proposed focal enhanced theory of emotional memory by Kensinger (2009). Whereby emotionally arousing stimuli leads to the use of emotion specific processes and therefore the emotional aspects of these stimuli are preserved in more detail, although they suggest negative valence may lead to increased enhancement in this process. This line of research suggests that emotion is prioritized by our cognitive systems and therefore may be more difficult to forget intentionally.

Intentional Forgetting of Emotional Stimuli

Although past studies have consistently demonstrated a directed forgetting using neutral common nouns (e.g., Bjork, 1970; MacLeod, 1975; MacLeod, 1989; Woodward & Bjork, 1971, Lee, 2018), pictures of neutral objects (e.g., Quinlan et al., 2010) or visual scenes (e.g., Hauswald & Kissler, 2008), and abstract symbols (Hourihan et al., 2009), there are also a growing number of studies using emotionally valenced materials (e.g., Bailey & Chapman, 2012; Hauswald et al., 2010; Yang et al., 2012). Most of these studies have been predicated on the idea that emotional material should be harder to forget intentionally – supported by the literature summarized in the preceding section. However, past experiments exploring this question in the context of item-method directed forgetting have proven inconsistent. These findings are summarized below.

Emotionally valenced words have led to a variety of outcomes as some studies have found no difference in the magnitude of the directed forgetting effect for neutral words in relation to negatively or positively valenced words (e.g., Berger, et al., 2018; Gallant, 2013), whereas other studies have found a smaller or non-significant directed forgetting effect for negatively or positively valenced words compared to neutral words (e.g., Bailey & Chapman, 2012; Yang et al., 2016). To date, only one study has used emotionally valenced sentences and has found a smaller directed forgetting effect for emotional sentences compared to neutral sentences (Lee & Hsu, 2013).

Studies using emotionally valenced images have produced similarly mixed results. For instance, Nowicka and colleagues (2011) and Zwissler and colleagues (2011), found smaller directed forgetting effects for emotional than neutral images, whereas Hauswald et al. (2010) failed to observe a significant directed forgetting effect for emotional images. Studies in children (aged 8 to 12) likewise produced no directed forgetting effect for negative images in recognition memory (although performance did not differ significantly between neutral and emotional conditions when measured using sensitivity; Augusti & Melinder, 2012). Otani and colleagues (2012) failed to observe a significant directed forgetting effect for negative images but observed a significant directed forgetting effect of similar magnitude for both neutral and positive images. Other studies have found no effect of valence on the magnitude of directed forgetting for images (Yang et al. 2012; Taylor et al., 2018) or faces (Quinlan & Taylor, 2014). In short, conclusions have varied broadly from one study to the next as to whether emotional stimuli of any kind are harder to forget in an item-method task.

Unconstrained variation in arousal has been discussed by several of the aforementioned studies as one possible reason for the reported discrepancies across this literature (e.g., Hauswald et al., 2010). Some studies have found that matching the valence conditions for arousal has eliminated the effect (e.g. Gallant et al., 2018; Yang et al., 2012). Allowing the level of arousal to vary instead between valence conditions has been shown to produce differences in the magnitude of directed forgetting, with relatively less constrained arousal favouring less directed forgetting for the arousing stimuli (e.g., Hauswald et al., 2010). Notably, others have found no effect of emotion when the emotional stimuli used were more arousing than the neutral stimuli (e.g., Taylor et al., 2018). Similarly, in a study comparing neutral, positive and negative valences at low and high arousal, the effect of arousal was inconsistent (Bailey & Chapman, 2012). Together, these studies are suggestive that arousal – and the failure to adequately operationalize emotion in general – might contribute to some of the between-study variability observed in this literature.

Neuroimaging has been utilized to isolate the mechanisms which may be responsible for the discrepancy in the magnitude of directed forgetting between emotional and neutral stimuli. Some studies using neuroimaging have revealed that intentionally forgetting emotional stimuli may be more difficult than neutral stimuli, even when the behavioural outcome indicates no effect of emotion (e.g. Yang et al., 2012). Yang et al. (2012) used negative and neutral images in their task while also monitoring neural patterns using ERP. An effect of valence on neural patterns was evident in a component known as the N2, thought to reflect the enactment of cognitive control processes, as enhanced activity was present while intentionally forgetting emotional items compared to neutral items (Yang et al., 2012; Nowicka et al., 2011). However, contrary to this finding Yang et al. (2016) used fMRI to compare the mechanisms involved in directed forgetting of negative and neutral words and found that following the forget instruction for neutral words there was enhanced activity in the right prefrontal cortex, but for negative words this was not the case. The authors suggest that emotion attenuates the effectiveness of control processes typically used during directed forgetting, leading to less forgetting of emotional items. These studies suggest that there are inconsistencies in the mechanisms thought to contribute to differences in intentional forgetting for emotional items compared to neutral items, but overall, they suggest that emotional items are more difficult to forget intentionally.

Difficulty intentionally forgetting emotional stimuli has been demonstrated in other paradigms as well. Xia and Evans (2020) used a modified version of the think/no-think task using neutral and negative images and intermixed trials of both an episodic task that required retrieval of the location the image was previously presented, and a perceptual task where participants indicated where the image was when queried. When the image was negative for the perceptual task, ERP analysis indicated that participants retrieved context information from the study phase even though it was unnecessary, indicating they were unable to inhibit the details from coming to mind. Another think/no-think task found that highly arousing negative items were more resistant to suppression during the task than neutral items (Marx et al., 2008). Payne and Corrigan (2007) used a list-method directed forgetting task where the lists were comprised of either neutral or emotional (both positive and negative) items. They found no directed forgetting for the emotional lists. Finally, in a retrieval-induced forgetting task, Dehli and Brennen (2009) found that both negative and positive words were resistant to the RIF effect compared to neutrally valenced words. These findings, and the previously mentioned research on the memory enhancement for emotional items provide converging evidence that emotional items are more difficult to intentionally forget.

The Current Study

As summarized in the preceding sections, emotional memories pose a unique problem when unwanted, because modern theoretical perspectives and empirical evidence might suggest they are difficult to expunge. For that reason, it is particularly important to explore the intentional forgetting of emotional items using paradigms such as the item-method directed forgetting paradigm, because doing so might help us understand traumatic memories and their role within clinical populations, such as those with depression (e.g., Wingenfeld et al., 2013; Yang et al., 2016; Kuehl et al., 2017), post-traumatic stress disorder (e.g., Geraerts & McNally, 2008; Zwissler et al., 2012; Baumann et al., 2013) and anxiety (Noel et al., 2011). However, it is also important in non-clinical populations because many of the experiences that we wish to intentionally forget in everyday life are also emotional in nature, such as a hurtful comment or painful medical procedure. Thus, in a normal non-clinical population, can the same process that helps us intentionally forget neutral experiences, such as an outdated phone number, also help us forget these emotionally meaningful experiences, such as the scene of a car accident?

To address this question, the present thesis reports a meta-analysis of item-method directed forgetting studies using emotional stimuli in non-clinical populations. As summarized above, the findings of these studies have been mixed; while some researchers report finding a smaller or non-significant directed forgetting effect for emotional items compared to neutral items (e.g., Bailey & Chapman, 2012; Hauswald et al., 2010; Lee & Hsu, 2013; Nowicka et al., 2011; Otani et al., 2012), other researchers report finding no difference in the directed forgetting effect for emotional and neutral items (e.g., Gallant & Yang, 2014; Liang et al., 2011; Marchewka et al., 2016; Quinlan & Taylor, 2014; Taylor et al., 2018; Yang et al., 2012; Zwissler et al., 2011). The aim of my thesis is to meta-analyze the research exploring the item-method directed forgetting of emotional information to determine whether emotional material is truly more resistant to intentional forgetting than neutral material. To the extent that there is a difference in the directed forgetting effect for emotional and neutral items, I also believe that it is crucial to investigate whether the magnitude of any such difference is moderated by specific study characteristics. The analyses were completed in two phases, first exploring the magnitude of the directed forgetting within each valence condition, and then comparing differences amongst the valence conditions. For the initial analyses, we predicted greater directed forgetting for words than images (e.g., Quinlan et al., 2010) and for recall than recognition (e.g., Titz & Verhaeghen, 2010); for comparisons between valence conditions, we expected greater differences for studies with relatively more arousing emotional than neutral stimuli (e.g. Bailey & Chapman, 2012). Other moderators are also considered on an exploratory basis (e.g., whether buffer items were presented during the study phase).

Method

Literature Search

We conducted a search of the online resources Google Scholar, PsycINFO, PsycARTICLES, and PubMED using the following Boolean search phrase: ("item-method" OR "item method") AND ("directed forgetting" OR "intentional forgetting") AND (“emotion” OR "emotional" OR "valence" OR "negative" OR "positive"). The search was conducted until November 2018 and restricted to English-language articles. The search was supplemented by reference review, review articles, and consultation with experts in the field.

Study Inclusion Criteria

Articles reporting at least one estimate of item-method directed forgetting as measured by recall or recognition, within a non-clinical population, using emotional images, words, faces, or events were considered for inclusion. Articles were excluded if they 1) used a clinical sample; 2) reported only a review or meta-analysis; 3) did not use an item-method directed forgetting task (e.g., think/no-think, list-method directed forgetting, retrieval-induced forgetting); 4) reported a sample with a mean age below 17 or greater than 40; 5) did not include an emotional condition; 6) were reported in a language other than English; 7) were inaccessible and the corresponding author failed to provide a copy; 8) reported an animal model of directed forgetting; 9) reported data from another, already included source; and/or, 10) provided insufficient information to calculate needed effect sizes.

Data Extraction

Each article was coded in consultation with my supervisor; all coding decisions were documented and discussed until a consensus emerged. In this manner I also coded methodological features for use as moderators, including the stimulus type (word or complex, memory task (recall or recognition), average valence and arousal of the stimuli set, whether participants engaged in a recall task prior to the recognition task (if included), list size (defined as the total number of items included in the study phase), whether buffer items were included preceding and/or following the study items (buffers, no buffers), and whether or not a neurophysiological measure (e.g., EEG) was gathered concurrent to the study portion of the task (this final variable was included to evaluate folk wisdom sometimes discussed at conferences about how the use of secondary measures such as EEG or fMRI encourage more attentive performance). Because the scale used for valence and arousal varied across studies, these values were standardized by dividing each by its maximum possible value for that study, producing a value ranging from 0 to 1; valence ratings for which the scale ranged from positive to negative (rather than the more common negative to positive) were inverted to align with the remaining data. Moderator analyses applied to differences in the magnitude between neutral and emotional conditions used instead the difference in the arousal and valence ratings for those conditions (reflecting the degree to which emotional items were more arousing or valenced than the neutral items).

Effect Size Calculation

Effect sizes were calculated as raw mean differences using the equations appropriate for within-subject designs provided by Borenstein et al. (2010, Chapter 4). I first calculated effect sizes estimating the magnitude of the directed forgetting effect (R – F) separately for the neutral, negative, and positive conditions within each experiment. An effect size was also calculated for the combination of the negative and positive conditions, which I refer to as the emotional condition. Next I calculated effect sizes estimating differences in the magnitude of directed forgetting for each of these conditions. I did this by subtracting performance for the R- and F-items for each (accounting for correlations amongst these scores) and then calculating an effect size comparing the resulting values. For the effect sizes comparing performance in the R and F conditions of each valence condition, positive values indicate greater performance for R-items compared to F-items; for effect sizes comparing the magnitude of directed forgetting across valence conditions (e.g., Neutral – Emotional), positive values indicate a larger directed forgetting effect in the neutral condition except for the comparison between the positive and negative conditions (i.e., Positive – Negative) wherein positive values indicate a larger directed forgetting effect in the positive condition. My preference was to focus on the combined (positive + negative) emotional condition, as it made best use of the available data, and further because each of the comparisons between the neutral and emotional conditions produced similar results.

For recall, the mean percentage of items recalled in each condition was used for the above calculations; for recognition, I instead used the percentage of "hits" for old items, although analogous calculations were also undertaken based on a measure of sensitivity (*d'*). My preference was to analyze recall accuracy and "hits" because they share a common scale and could therefore be aggregated using raw difference scores. Nonetheless, given that "hits" risk conflating sensitivity and response bias, supplementary analyses were undertaken to verify whether findings observed in the main analyses were also present after contributions due to response bias had been removed.[1] I used values of *d'* reported in-text or derived from raw data where possible, with the remaining values calculated using aggregate hits and false alarms.

---

[1] I was not specifically concerned about the possibility of response bias contaminating the present findings owing to the fact that almost all included studies used a single false alarm rate for each valence condition. Because the directed forgetting effect was calculated as the difference between the R and F items within a given valence condition (which shared a common false alarm rate), I expected that differences between these conditions were already driven in large part by differences in sensitivity.

Throughout the calculations, standard deviations were required for each condition to estimate the standard error of the appropriate difference score. In cases where the standard deviations were unavailable from either the text or study's authors, I imputed the relevant value as the average of the standard deviations available for that dependent measure. Only a small number of standard deviations were imputed in this manner for my analysis of recall accuracy and recognition "hits"; however, few studies reported *d'* directly, requiring the imputation of most of the standard deviations used in the calculation of those models. For that reason, additional sensitivity analyses were undertaken of the *d'* models weighting studies by their sample sizes, rather than incorporating their standard errors into the model fit. Because the sample-size weighted models produced results qualitatively similar to – albeit more liberal than – the models using the standard errors, I have chosen to report only the latter in-text.

Following calculation of the effect sizes, separate Bayesian random- and mixed-effects models were generated using the *brms 2.9.0* (Bürkner, 2017, 2018) package in *R 3.6.1* (R Core Team, 2018). Each model reflected a separate multi-level regression where the overall, aggregate effect was estimated based on the weighted combination of the included studies. The weight given to a study was determined by the sampling error for that effect, which was incorporated into the model as a form of measurement error, effectively weighting studies based on the precision of their estimate (i.e., as reflected by their associated standard error and driven by factors such as sample size). These models were considered multi-level because they each included a random intercept reflecting variation across studies. In short, this approach assumes that studies differ in the "true" effect being estimated in the sampled population; differences could be due to any manner in which those studies differ from one another. Following from this approach, the goal of modern meta-analysis is most often not to estimate a single aggregate effect, but rather to characterize the distribution of effects represented in the literature and explore sources of between-study heterogeneity (reflecting variation in these "true" effects). The estimate corresponding to each study was assumed to be sampled from a normal distribution centred on the aggregate effect with a standard deviation which was itself estimated within the model; the standard deviation in this case would be referred to as τ in a more traditional meta-analysis, with a large standard deviation reflecting relatively high between-study heterogeneity and small standard deviation reflecting relatively low between-study heterogeneity. In summary, this modelling approach estimates both an aggregate effect (reflecting the expected effect in a "typical" study) as well as the underlying range of expected "true" effects after incorporating model assumptions and reasonable prior knowledge (discussed below) concerning the possible range of values.

Because several studies reported both recall and recognition performance, I was also able to incorporate a random slope into the analysis of task effects. Random slopes in this case would reflect between-study variability in the difference between recall and recognition, allowing for the possibility that some studies show larger differences than others. The assumptions underlying random slopes are functionally the same as those for random intercepts, assuming that the difference between recall and recognition for a given study is sampled from a normal distribution centred on the aggregate difference between these conditions, with a standard deviation which was itself estimated within the model.

Each model was run using four chains with 10,000 iterations per chain. Each chain represented an independent version of the model, implemented with random starting values for each parameter to ensure that the parameter estimates produced by the model were not influenced by the starting values or random noise during the fitting procedure. As a result, 40,000 iterations were run in total, although the first 5,000 iterations were removed from each chain for the warmup period (allowing time for the model to calibrate itself), resulting in 20,000 total iterations that served as the posterior distribution for each model. I fit a separate model for each moderator analysis; this was necessary due to incomplete or inconsistent reporting across studies. The modelling approach used is similar to previous work from the NeuroFog laboratory (e.g., Fawcett et al., 2016; Fawcett & Ozubko, 2016; see also, Fawcett et al., in press; Fawcett et al., 2019). Models were fit and evaluated for convergence using standard practices (R-hat < 1.01; Gelman & Hill, 2007).

My priors for each model were mildly informed by expert knowledge derived from consultation with my supervisor and consideration of the broader literature on human memory; these priors are summarized below. When comparing the magnitude of directed forgetting within a given valence, my prior expectations relating to the intercept of each model assumed that the average effect in a typical sample should range somewhere between -30% and 30%. I further assumed that the standard deviation pertaining to random effects would most likely range between 0 and 30%; this broadly permits the "true" effect within any given sample to vary anywhere from -90% to 90%. The prior for slopes within the moderator models were represented by a normal distribution centred at 0 with a standard deviation of 30. The corresponding prior expectations for the analogous analyses using *d'* were calibrated such that the average effect in a typical sample should range somewhere between -1.00 and 1.00, with the standard deviation pertaining to random effects ranging between 0 and 1.00. This broadly permits the "true" effect within any given sample to vary anywhere from -3.00 to 3.00. The prior for slopes were represented by a normal distribution centred at 0 with a standard deviation of 0.50.

For the comparison of the magnitude of directed forgetting across valences, my prior expectations relating to the intercept of each model assumed that the average effect in a typical sample should range somewhere between -20% and 20%. I further assumed that the standard deviations pertaining to random effects would most likely range between 0% and 20%; this broadly permits the "true" effect within any given sample to vary anywhere from -60% to 60%. The prior for slopes within the moderator models were reflected by a normal distribution centred at 0 with a standard deviation of 20. The corresponding prior expectations for the analogous analyses using *d'* were calibrated such that the average effect in a typical sample should range somewhere between -0.50 and 0.50, with the standard deviation pertaining to random effects ranging between 0 and 0.50. This broadly permits the "true" effect within any given sample to vary anywhere from -1.50 to 1.50, with values outside this range possible albeit improbable. The prior for slopes were reflected by a normal distribution centred at 0 with a standard deviation of .25.

Finally, although my focus will be on evaluating the magnitude of directed forgetting within and across valence conditions, it is also important to consider variation in the magnitude of the underlying effects between studies. This heterogeneity will be quantified using prediction intervals (IntHoult et al., 2016) calculated within each of the primary (i.e., non-moderator) models. Prediction intervals reflect the range of probable "true" effects that would be expected should a new study be conducted similar to those included in the analysis; as such, these intervals address the question as to what degree methodological differences (e.g. differences in demographics or methods across studies) within this literature are liable to contribute to variability observed across the included estimates, separate from sources such as sampling or measurement error.

Results

Description of Studies

Of the 607 studies initially identified by our search, 31 were ultimately included in the final sample, providing 36 independent effect sizes (see Figure 1). Articles contributing effect sizes to one or more of our analyses are indicated in our reference section by an asterisk (*). Study characteristics are summarized in Table 1.

Figure 1.

Meta-analysis inclusion flowchart

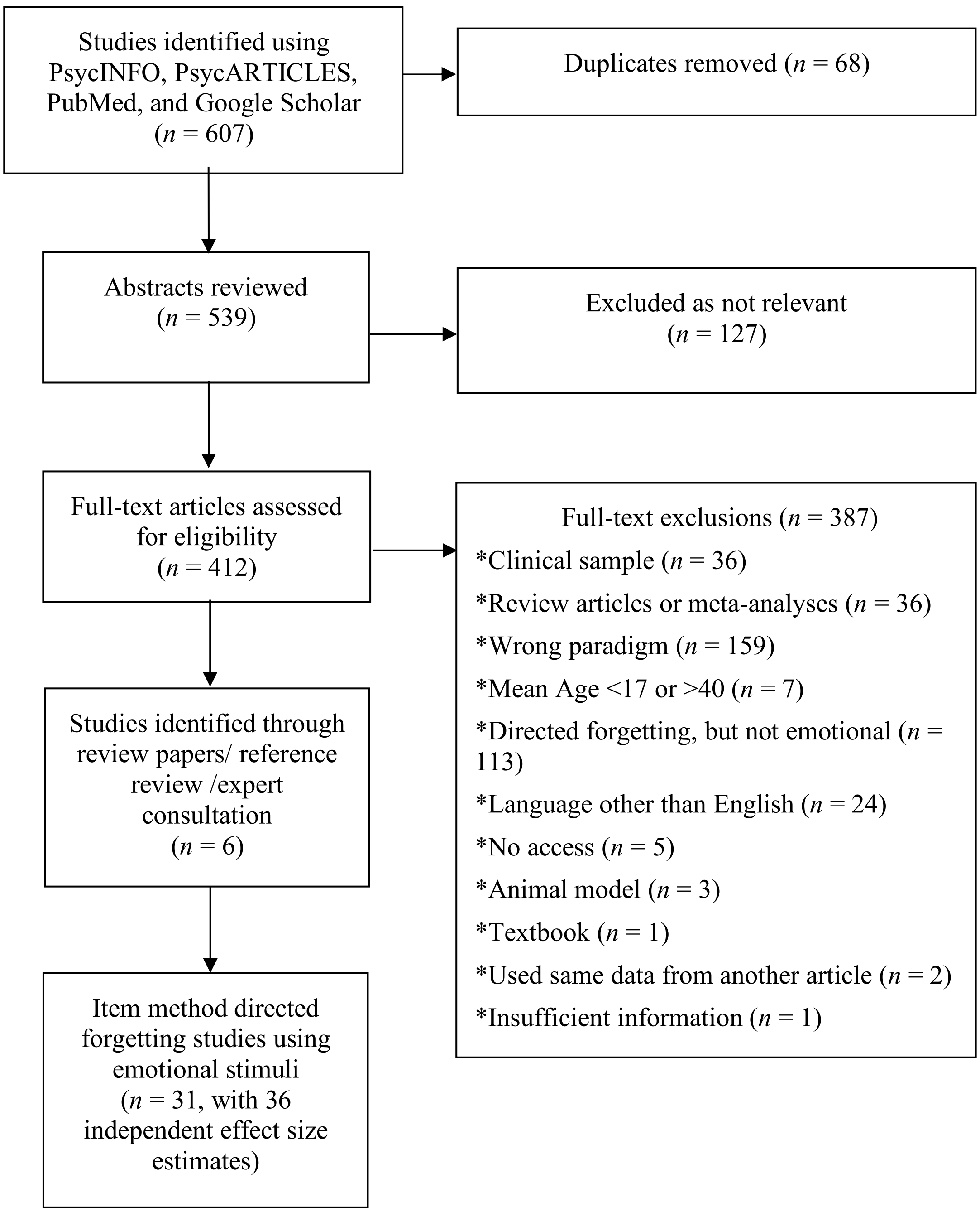


Table 1.

Study Characteristics

						Arousal			Valence				
Paper	**Stimuli**	**Task**	**Recall First**	**Buffers**	**Valences Included**	**Neu.**	**Emo.**	**Diff.**	**Neu.**	**Emo**	**Diff.**	**List Size**	**Secondary Measure**
Bailey and Chapman (2012, Experiment 1)	Words	Both	Yes	No	Neutral, Positive, Negative	–	–	–	–	–	–	240	No
Bailey and Chapman (2012, Experiment 2)	Words	Both	Yes	No	Neutral, Positive, Negative	–	–	–	–	–	–	240	Yes
Berger et al. (2018)	Words	Recognition	No	No	Neutral, Positive, Negative	0.45	0.64	0.19	0.59	0.21	0.38	90	No
Brandt et al. (2013)	Words	Recognition	No	No	Neutral, Negative	–	–	–	–	–	–	80	Yes
Gallant (2013)	Words	Recognition	No	Yes	Neutral, Positive, Negative	0.47	0.5	0.03	0.56	0.24	0.32	60	No
Gallant and Dyson (2016)	Words	Recognition	No	No	Neutral, Positive, Negative	0.44	0.68	0.24	0.57	0.22	0.35	240	Yes
Gallant et al. (2018)	Words	Recognition	No	Yes	Neutral, Positive, Negative	0.47	0.5	0.03	0.56	0.24	0.32	60	Yes
Gamboa et al. (2017)	Words	Both	Yes	Yes	Neutral, Negative	–	–	–	–	–	–	80	No
Hauswald et al. (2010)	Pictures	Recognition	No	No	Neutral, Negative	0.31	0.67	0.35	0.58	0.29	0.29	120	Yes
Korfine and Hooley (2000)	Words	Both	Yes	No	Neutral, Positive, Negative	0.46	0.60	0.15	0.65	0.24	0.41	42	No
Lee and Hsu (2013, Experiment 1)	Events	Recall	No	No	Neutral, Negative	0.48	0.81	.33	0.66	0.31	0.36	60	No
Lee and Hsu (2013, Experiment 2)	Events	Recall	No	No	Neutral, Negative	0.48	0.81	.33	0.66	0.31	0.36	60	No
Lee and Hsu (2013, Experiment 3)	Events	Recall	No	No	Neutral, Negative	0.48	0.81	.33	0.66	0.31	0.36	60	No
Li et al. (2017)	Events	Recall	No	No	Neutral, Positive, Negative	0.43	0.57	0.14	0.57	0.27	0.30	42	No
Liu, Chen and Cheng (2017)	Words	Recognition	No	Yes	Neutral, Negative	0.35	0.67	0.32	0.56	0.26	0.30	240	Yes
Marchewka et al. (2016)	Pictures	Recognition	No	No	Neutral, Negative	0.17	0.36	0.19	0.62	0.33	0.29	240	Yes
McNally et al. (1998)	Words	Recall	No	Yes	Neutral, Positive, Negative	0.35	0.58	0.23	0.61	0.24	0.37	30	No
McNally et al. (1999)	Words	Recall	Yes	Yes	Neutral, Positive, Negative	–	–	–	–	–	–	48	No
Moulds and Bryant (2002)	Words	Recall	No	Yes	Neutral, Positive, Negative	–	–	–	0.69	0.16	0.54	30	No
Moulds and Bryant (2008)	Words	Recall	No	Yes	Neutral, Positive, Negative	–	–	–	0.69	0.16	0.54	30	No

Nowicka et al. (2011)	Pictures	Recognition	No	No	Neutral, Negative	0.42	0.64	0.21	0.67	0.31	0.36	120	Yes
Otani et al. (2012)	Pictures	Recall	No	No	Neutral, Positive, Negative	0.29	0.58	0.29	0.52	0.23	0.29	30	No
Patrick et al. (2015)	Words	Recognition	No	No	Neutral, Negative	0.46	0.65	0.19	0.50	0.21	0.29	140	Yes
Pierguidi et al. (2016)	Faces	Recognition	No	No	Neutral, Positive, Negative	0.36	0.59	0.23	0.56	0.29	0.27	216	Yes
Quinlan and Taylor (2014, Experiment 1)	Faces	Recognition	No	No	Neutral, Positive, Negative	–	–	–	–	–	–	60	No
Quinlan and Taylor (2014, Experiment 2)	Faces	Recognition	No	No	Neutral, Positive, Negative	–	–	–	–	–	–	60	No
Tay and Yang (2017)	Faces	Recognition	No	No	Positive, Negative	–	0.52	0.02	–	–	–	48	No
Taylor et al. (2018, Experiment 1)	Pictures	Recognition	No	No	Neutral, Negative, Positive	0.36	0.55	0.19	0.55	0.27	0.29	96	No
Taylor et al. (2018, Experiment 2)	Pictures	Recognition	No	No	Neutral, Negative, Positive	0.36	0.55	0.19	0.55	0.27	0.29	144	No
Teckan et al. (2008)	Words	Both	Yes	Yes	Neutral, Positive, Negative	–	–	–	–	–	–	54	No
Wierzba et al. (2018)	Words	Recognition	No	No	Neutral, Negative	0.35	0.62	0.27	0.59	0.41	0.18	120	Yes
Wilhelm et al. (1996)	Words	Both	Yes	Yes	Neutral, Positive, Negative	0.37	0.54	0.18	0.62	0.26	0.36	48	No
Yang et al. (2012)	Pictures	Recognition	No	No	Neutral, Negative	–	–	–	–	–	–	280	Yes
Yang et al. (2016, Experiment 1)	Words	Recognition	No	Yes	Neutral, Negative	–	–	–	–	–	–	64	Yes
Yang et al. (Experiment 2)	Words	Recognition	No	Yes	Neutral, Negative	–	–	–	–	–	–	160	Yes
Zwissler et al. (2011)	Pictures	Recognition	No	No	Neutral, Positive	0.37	0.53	0.16	0.50	0.20	0.30	36	No

Note. Stimuli = stimuli type; Task = memory task used; Recall First = whether the recognition test was preceded by a recall test; Buffers = whether buffer items were included; Valences included = the valence conditions used; Arousal = the arousal rating provided for the stimuli used for the neutral and emotional items and the difference in rating between valences; Valence = valence rating provided for the stimuli used for the neutral and emotional items and the difference between the valence rating between the neutral and emotional items; Study List = the total number of study items presented in the learning phase; Secondary Measure = whether a neural or psychophysiological measure was used during the task.

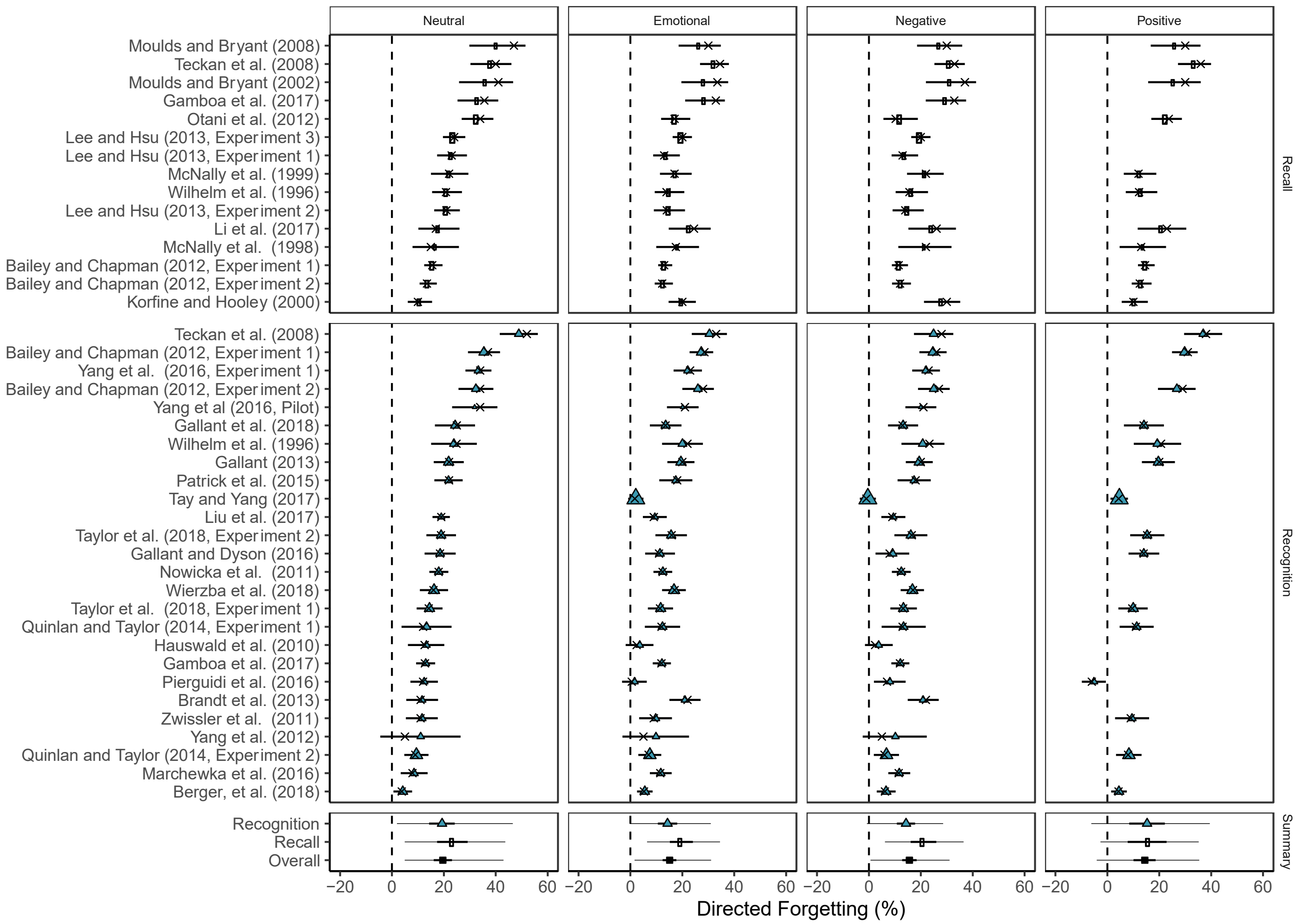
Directed Forgetting for Neutral and Emotional Conditions

The initial analyses focused on estimating the magnitude of directed forgetting within each of the valence conditions. For the sake of exposition – and to maximize statistical power – my focus is on the neutral and emotional conditions, owing to the fact that the negative and positive conditions were similar across all models and did not differ when compared directly (see the following section). Although I report the aggregate effects for the negative and positive estimates in-text, all moderator models pertaining to those conditions are summarized instead in Appendix A.

As depicted in Figure 2, a directed forgetting effect was observed in each of the valence conditions, ranging in magnitude from 14.4% (positive condition) to 19.6% (neutral condition). Notably, the magnitude of the directed forgetting effect was numerically larger in the neutral condition than each of the valence conditions, whereas positive and negative effects were of numerically similar magnitude. These comparisons are explored further in the following section.

***Figure 2*.**

Magnitude of Directed Forgetting (%) for Each Valence Condition: Neutral, Emotional, Negative and/or Positive items as a Function of Task Type



Note. (Yellow Circles: Recall, Blue Triangles: Recognition). Symbols and error bars represent posterior estimates and their corresponding 95% confidence intervals. X's represent the empirical values reported in the relevant article. Symbol size is scaled to reflect relative sample size. Estimates provided in the bottom panel represent aggregate effects; in this panel, thick lines reflect 95% confidence intervals and thin lines reflect 95% prediction intervals.

Despite clear evidence of an effect in a "typical" study, prediction intervals also revealed substantial heterogeneity across studies. Dispersion of the "true" effects was numerically broader in the positive, $PI_{95\%}$ [-4.3%, 35.2%], and neutral conditions, $PI_{95\%}$ [4.7%, 43.5%], than in the negative, $PI_{95\%}$ [0.5%, 31.1%], and emotional, $PI_{95\%}$ [1.3%, 31.3%], conditions. Of these prediction intervals, all but the positive valence condition excluded negative values, indicating that the "true" effect for most studies with methods similar to those in the present analysis should indicate some degree of directed forgetting – at least for the neutral, negative or emotional conditions – although some of those effects are likely close to 0%.

As summarized in Table 2 the effect of each moderator was consistent across the neutral and emotional conditions. In particular, the magnitude of directed forgetting was greatest for comparisons (a) using single words than complex stimuli (pictures, faces and narrative events); (b) using recall than recognition; (c) using recognition for which the recognition test was preceded by a recall test; and, (d) for which the study phase included buffer items to minimize recency and primacy effects. The effect of list size also demonstrated a trend toward favouring a larger magnitude of directed forgetting for smaller lists. There was minimal evidence to support an effect of valence, arousal, or the inclusion of a secondary measure (e.g., electroencephalography, functional magnetic resonance imaging).

Table 2.

Effect of Moderators on the Magnitude of Directed forgetting for Neutral and Emotional Items

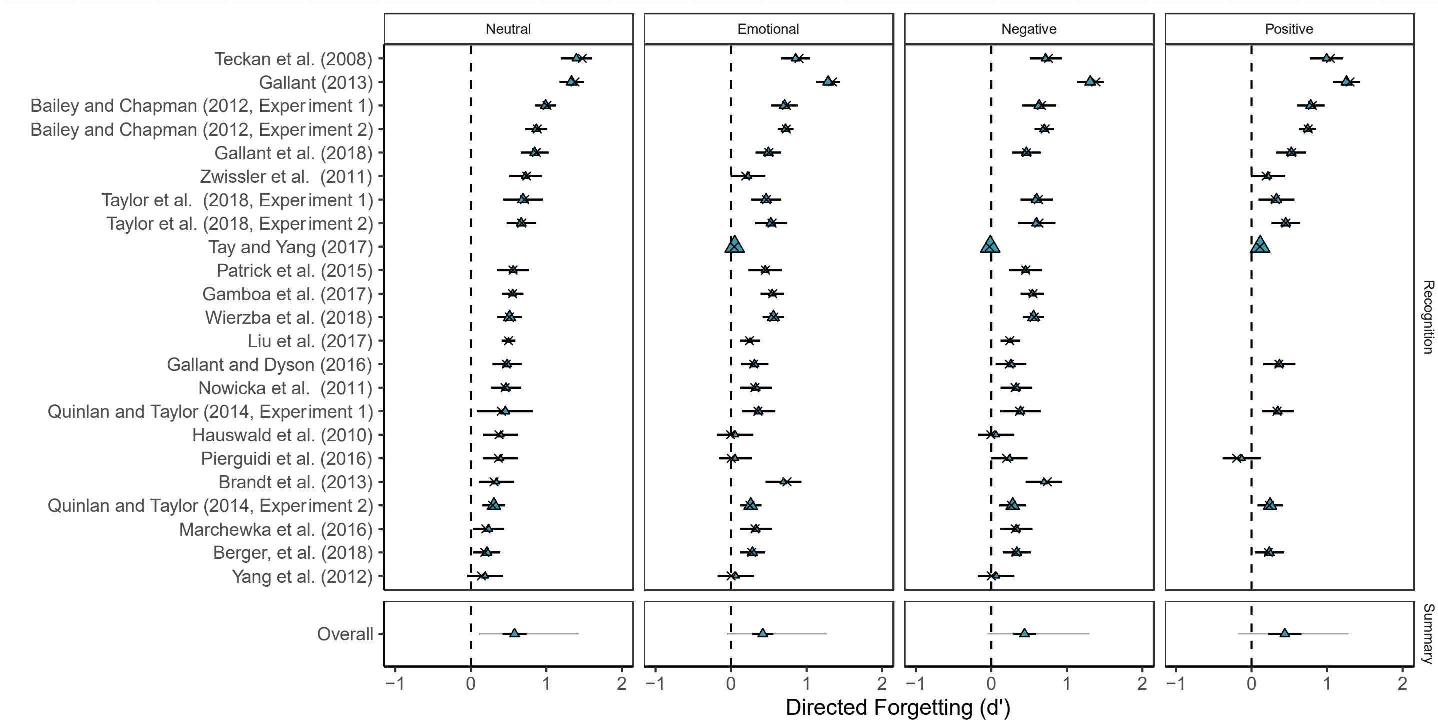
Moderator	*k*	*M*	Difference (%)	*p*
Neutral Items				
Stimulus				
Complex	14	16.2 [10.9, 21.4]		
Words	26	22 [17.8, 26.4]	5.8 [-0.9, 12.7]	0.96
Task				
Recall	15	23.5 [17.4, 29.2]		
Recognition	25	19.3 [14.3, 24.2]	-4.2 [-11.7, 3.4]	0.87
Arousal	25	2.93 [-19.4, 35.7]		0.56
Valence	29	19.5 [-23.1, 61.0]		0.82
Recall First				
No recall	20	16.2 [11.6, 20.6]	15.0 [5.1, 24.9]	1
Recall	5	31.3 [21.9, 40.1]		
List Size	40	-3.1 [-6.8, 0.47]		0.95
Buffer				
No buffers	24	15.7 [12, 19.2]	12.0 [5.8, 18.3]	1
Buffers	15	27.6 [22.6, 32.7]		
Secondary measure				
No measure	25	20.7 [16.2, 25.1]	-2.6 [-9.7, 4.4]	0.77
Measure	15	18.1 [12.4, 23.6]		
Emotional Items				
Stimulus				
Complex	15	10.3 [6.7, 13.8]		
Words	26	18.4 [15.4, 21.4]	8.1 [3.6, 12.8]	1
Task				
Recall	15	19.6 [15.2, 24.1]		
Recognition	26	14.3 [10.6, 18.0]	-5.4 [-11.0, 0.2]	0.97
Arousal	26	0.3 [-27.0, 28.0]		0.51
Valence	29	-22.3 [-59.9, 15.7]		0.87
Recall First	26	12.3 [5.0, 20.1]		1
No Recall	21	11.7 [8.3, 15.0]	12.3 [5.0, 20.0]	1.00
Recall	5	24.0 [17.3, 30.8]		
Buffer				
No buffers	25	12.2 [9.4, 15.0]	8.1 [2.6, 13.1]	1
Buffers	15	20.3 [16.1, 24.4]		
Secondary measure		-		
No Measure	26	16.2 [12.9, 19.6]	-3.1 [-8.5, 2.3]	0.87
Measure	15	13.2 [8.9, 17.5]		

Note. k = the number of observations for each moderator or each level of the moderator; *M* = the mean estimate of directed forgetting for each level of the moderator, in the case of continuous moderators it indicates the slope, 95% confidence interval presented in brackets; Difference (%) = difference in magnitude of directed forgetting between levels of the moderator, 95% confidence interval presented in brackets; *p* = Bayesian p-value reflecting confidence in the direction of the effect (e.g., *p* = .95 for a positive effect means 95% confidence that the effect is positive).

These effects persisted when measured using *d'* rather than hits, as summarized in Figure 3 and Table 3, although the effect of list size no longer excluded 0 (no difference) as a credible value for either the neutral or emotional models.

Figure 3.

Magnitude of Directed Forgetting (d') for Each Valence Condition: Neutral, Emotional, Negative and/or Positive items



Note. Triangles and error bars represent posterior estimates and their corresponding 95% confidence intervals. X's represent the empirical values reported in the relevant article. Symbol size is scaled to reflect relative sample size. Estimates provided in the bottom panel represent aggregate effects; in this panel, thick lines reflect 95% confidence intervals and thin lines reflect 95% prediction intervals.

Table 3.

Effect of Moderators on the Magnitude of Directed forgetting for Neutral and Emotional Items in d’

Moderator	***k***	***M***	**Difference(%)**	***p***
Neutral Items				
Stimulus				
Non- words	10	0.4 [0.2, 0.7]	0.3 [-0.0, 0.6]	0.96
Words	12	0.7 [0.5, 0.9]		
Arousal	14	0.3 [-0.6, 1.2]		0.76
Valence	16	-0.2 [-1.1, 0.8]		0.63
Recall First	22			
No Recall	18	0.5 [0.4, 0.7]	0.4 [0.1, 0.8]	0.99
Recall	4	0.9 [0.6, 1.2]		
List size	22	-0.1 [0.1, -0.3]		0.88
Buffers				
No Buffers	17	0.4 [0.2, 0.7]	0.4 [0.1, 0.7]	0.99
Buffers	5	0.7 [0.5, 0.9]		
Activity Measure				
No measure	10	0.7 [0.5, 0.9]	-0.3 [-0.5, 0.0]	0.95
Measure	12	0.5 [0.3 0.7]		
Emotional Items				
Stimulus				
Complex	11	0.2 [0.1, 0.4]	0.4 [0.1, 0.6]	1
Words	12	0.6 [0.4, 0.8]		
Arousal	15	-0.2 [-1.1, 0.7]		0.68
Valence	16	-0.1 [-1.0, 0.9]		0.56
Recall First	23			
No recall	19	0.4 [0.2, 0.5]	0.3 [0, 0.6]	0.97
Recall	4	0.7 [0.4, 1.0]		
List Size	23	-0.1 [-0.2, 0.1]		0.83
Buffers				
No Buffers	18	0.2 [0.1, 0.4]	0.3 [0.0, 0.6]	0.98
Buffers	5	0.6 [0.4, 0.8]		
Activity Measure				
No measure	11	0.5 [0.3, 0.7]	-0.2 [-0.4, 0.1]	0.86
Measure	12	0.4 [0.2, 0.5]		

Note. *k* = the number of observations for each moderator or each level of the moderator; *M* = the mean estimate of directed forgetting for each level of the moderator, in the case of continuous moderators it indicates the slope, 95% confidence interval presented in brackets; Difference (%) = difference in magnitude of directed forgetting between levels of the moderator, 95% confidence interval presented in brackets; *p* = Bayesian p-value reflecting confidence in the direction of the effect (e.g., $p = .95$ for a positive effect means 95% confidence that the effect is positive).

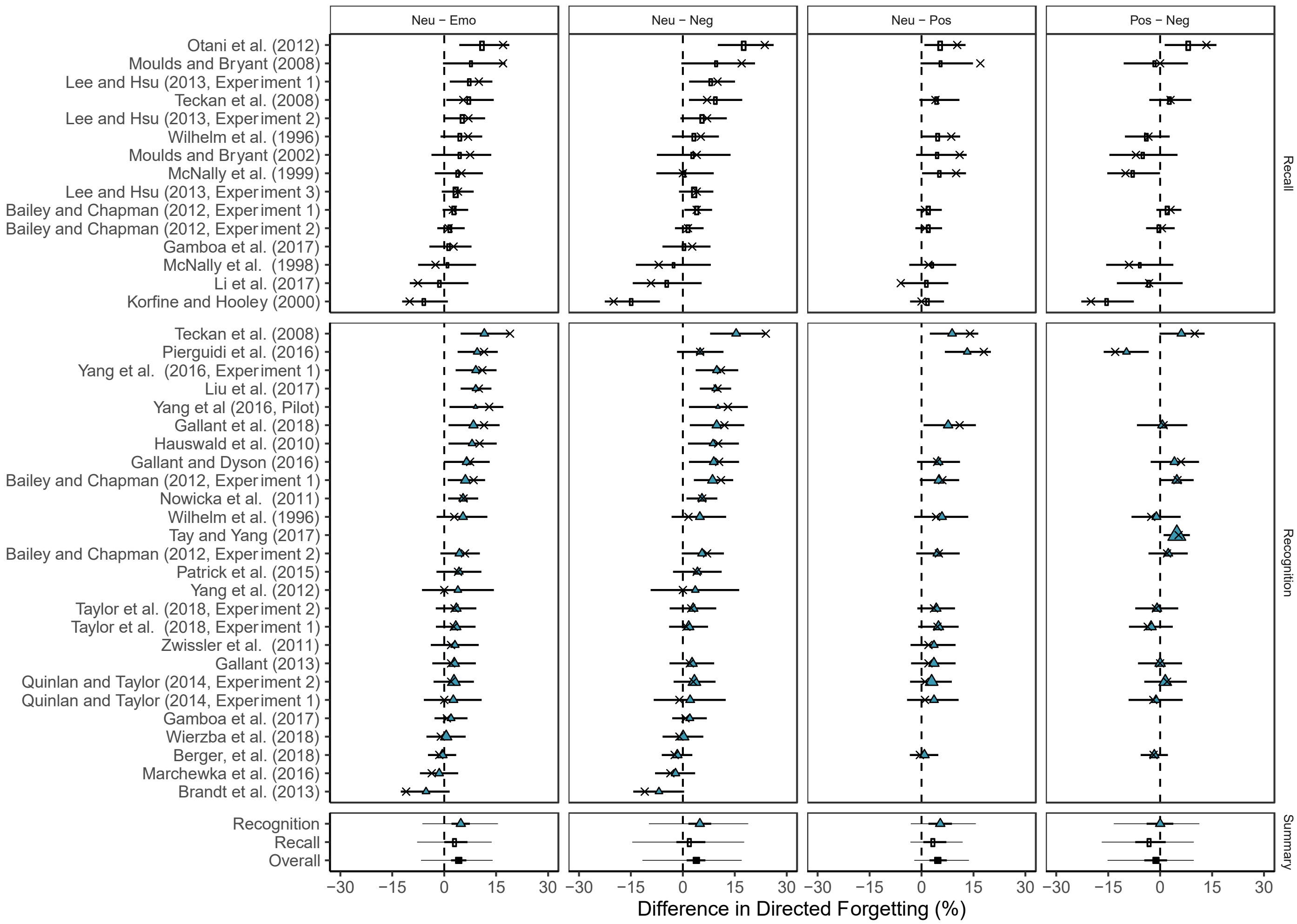
Comparing the Magnitude of Directed Forgetting Across Conditions

Having established the presence of a credible directed forgetting effect within a “typical” study using each valence, I next compared the magnitude of directed forgetting across valences. As before, my focus was on the neutral and emotional conditions, with all moderators pertaining to the negative and positive conditions summarized in Appendix A.

As depicted in Figure 4, on average directed forgetting tends to be smaller for emotional than neutral stimuli but does not differ between negative and positive stimuli. Even so, prediction intervals again revealed substantial heterogeneity between studies. Of particular interest, prediction intervals for the neutral-emotional comparison revealed that the probable “true” effects for studies with methods similar to those in the present sample would be expected to range from as low as -5.3% to as high as 13.7%. That is to say that whereas we would expect most (75%) studies to support the claim that emotional stimuli are less likely than neutral stimuli to be forgotten intentionally, this may not always be the case: For the remaining studies, similar – or even slightly superior – DF would be expected for emotional relative to neutral items.

Figure 4.

Mean difference in the magnitude of directed forgetting (%) between Neutral, Emotional, Negative and/or Positive items as a function of task type



Note. (Yellow Circles: Recall, Blue Triangles: Recognition). Symbols and error bars represent posterior estimates and their corresponding 95% confidence intervals. X's represent the empirical values reported in the relevant article. Symbol size is scaled to reflect relative sample size. Estimates provided in the bottom panel represent aggregate effects; in this panel, thick lines reflect 95% confidence intervals and thin lines reflect 95% prediction intervals.

The moderators summarized in Table 4 provide potential insight into when emotional material might – or might not – be less likely to be forgotten intentionally. In particular, differences in the magnitude of DF between emotional and neutral items were expected to be greatest for comparisons (a) for which the emotional items were more arousing than the neutral items; and, (b) for which the study phase included buffer items to minimize recency and primacy effects.

Table 4.

Effect of Moderators on the Difference in the Magnitude of Directed forgetting Between Emotion Conditions

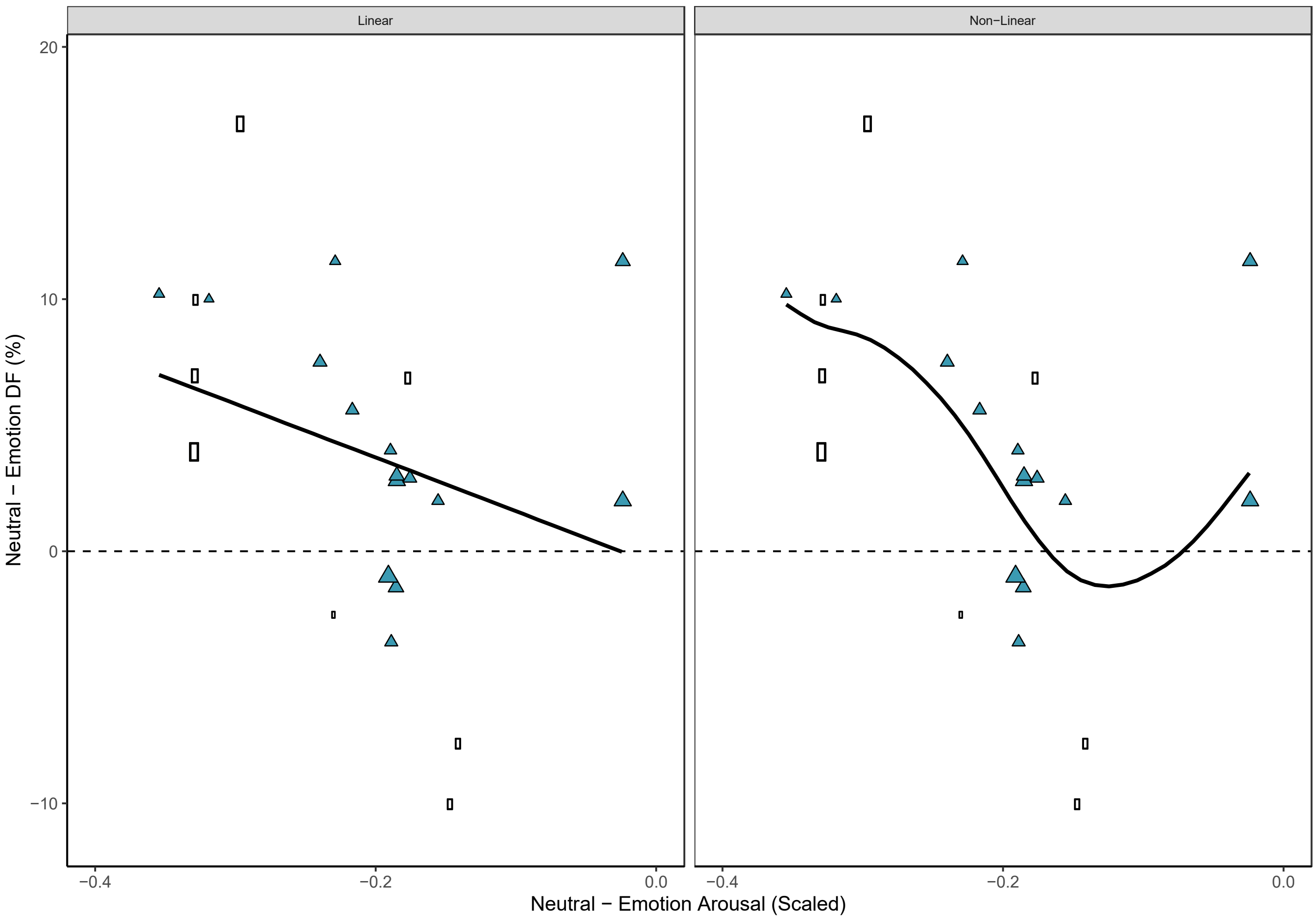
Moderator	***k***	***M***	**Difference (%)**	***p***
Neu-Emo				
Stimulus				
Complex	14	5.3 [1.8, 8.8]	-1.9 [-6.4, 2.6]	0.8
Words	26	3.5 [0.7, 6.3]		
Task				
Recall	15	3.4 [0.0, 6.7]	1.37 [-2.6, 5.2]	0.76
Recognition	25	4.8 [2.1, 7.4]		
Arousal	25	-21.2 [-44.6, 3.6]		0.96
Valence	29	2.3 [-25.4, 29.8]		0.56
Recall First				
No recall	28	3.9 [1.1, 6.8]	3.03 [-3.55, 9.7]	0.83
Recall	12	7.0 [1.1, 13.0]		
List size	40	0.2 [-2.2, 2.6]		0.58
Buffers				
No Buffers	24	3.0 [0.5, 5.5]	4.45 [-0.0, 8.9]	0.97
Buffers	15	7.5 [3.8, 11.1]		
Activity measure	40			
No measure	25	3.6 [0.8, 6.5]	1.4 [-3.2, 5.9]	0.73
Measure	15	5.0 [1.5, 8.5]		
Neg-Pos				
Stimulus				
Complex	7	-0.2 [-6.0, 5.7]	-2.2 [-9.5, 4.9]	0.74
Words	18	2.1 [-6.3, 2.1]		
Task				
Recall	11	-2.8 [-7.2, 1.6]	2.7 [-2.0, 7.5]	0.88
Recognition	14	0.0 [-3.9, 3.8]		
Arousal	16	-12.6 [-49.9, 25.5]		0.75
Valence	19	-10.52 [-43.4, 23.1]		0.73
Recall First				
No recall	15	-0.5 [-4.1, 3.0]	4.5 [-2.6, 11.5]	0.9
Recall	10	3.9 [-2.2, 10.1]		
List size	25	0.65 [-3.1, 4.3]		0.65
Buffers				
No Buffers	14	-0.7 [-5.1, 3.7]	-1.5 [-8.9, 5.8]	0.66
Buffers	10	-2.1 [-8.0, 3.7]		
Activity Measure				
No measure	20	-1.2 [-5.1, 2.5]	-0.1 [-8.1, 8.5]	0.52
Measure	5	-1.3 [-8.6, 6.3]		

Note. k = the number of observations for each moderator or each level of the moderator; M = the mean estimate of directed forgetting for each level of the moderator, in the case of continuous moderators it indicates the slope, 95% confidence interval presented in brackets; Difference (%) = difference in magnitude of directed forgetting between levels of the moderator, 95% confidence interval presented in brackets; *p* = Bayesian p-value reflecting confidence in the direction of the effect (e.g., *p* = .95 for a positive effect means 95% confidence that the effect is positive).

The effect of arousal was of particular interest, having been raised as a possible explanation for heterogeneity within this literature (e.g., Hauswald et al., 2010). As depicted in Figure 5, the difference in DF for neutral and emotional items was greatest when the emotional items were more arousing than the neutral items, with the effect dissipating as arousal reached parity; when arousal was matched, DF was predicted to be numerically equivalent for the neutral and emotional items, M = 0.0%, $CI_{95\%}$ [-5.5%, 5.6%].

Figure 5.

Effect of arousal on the difference in the magnitude of directed forgetting between neutral and emotional conditions as a function of task type

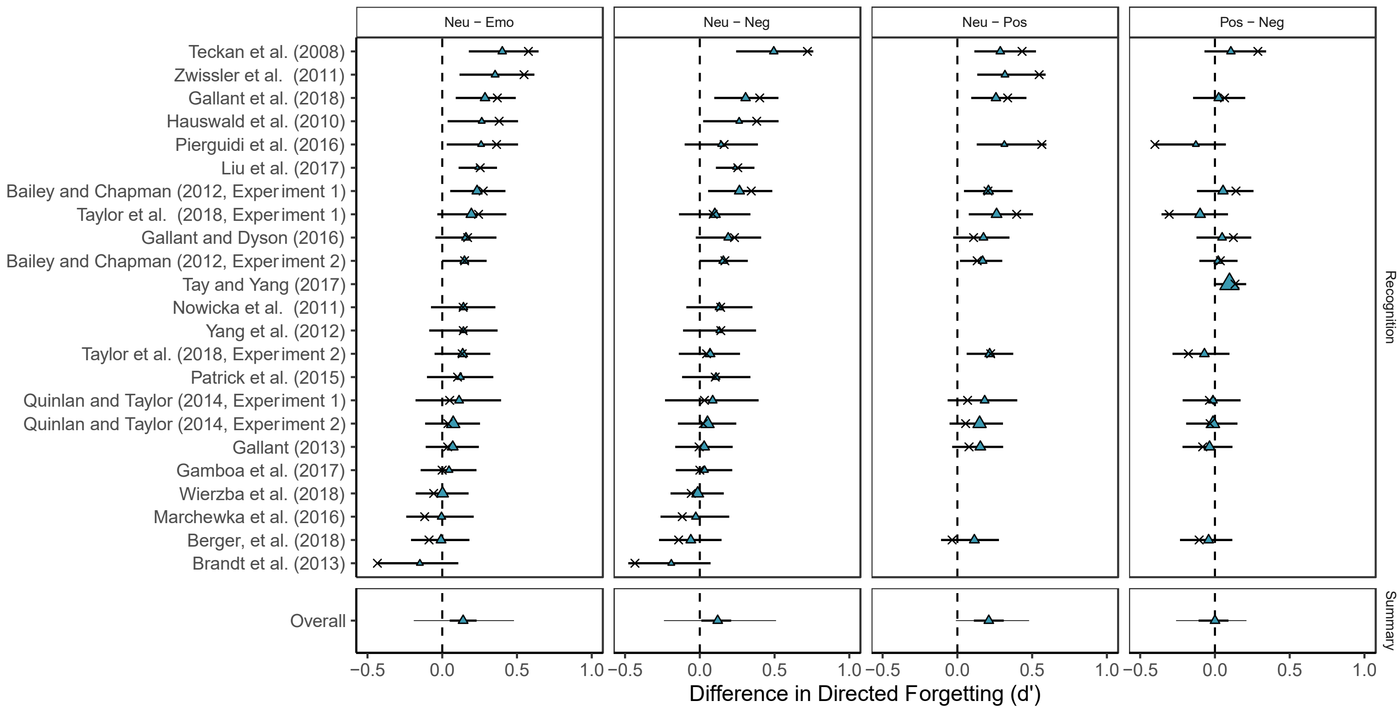


Note. (Yellow Circles: Recall, Blue Triangles: Recognition). X-axis indicates the mean difference in (scaled) arousal between conditions. The Y-axis indicates the difference in the magnitude of directed forgetting between conditions (%). The left panel represents the effect fitted as a linear function; the right panel represents the effect fitted as a non-linear function. Symbol size is scaled to reflect relative sample size.

However, unlike the preceding models, although the analyses of *d'* produced a similar pattern of differences in the overall magnitude of directed forgetting between the emotional and neutral conditions (see Figure 6), the same was not true for the moderator analyses. Despite the directionality of the effects matching those described in the preceding paragraph, none of the moderators reported in Table 5 were credible, possibly owing to the exclusion of half or more of the relevant effects, resulting in a loss of statistical power.

Figure 6.

Difference in the Magnitude of Directed Forgetting (d') for Each Valence Condition: Neutral, Emotional, Negative and/or Positive items.



Note. Triangles and error bars represent posterior estimates and their corresponding 95% confidence intervals. X's represent the empirical values reported in the relevant article. Symbol size is scaled to reflect relative sample size. Estimates provided in the bottom panel represent aggregate effects; in this panel, thick lines reflect 95% confidence intervals and thin lines reflect 95% prediction intervals.

Table 5.

Moderators Influencing the Difference in the Magnitude of Directed Forgetting Between Emotion Conditions in d'

Moderator	*k*	*M*	Difference (%)	*p*
Neu-Emo				
Stimulus				
Complex	10	0. 2 [0, 0.3]	-0.1 [-0.2, 0.1]	0.74
Words	12	0.1 [0, 0.2]		
Arousal	14	-0.05 [-0.5, 0.4]		0.59
Valence	16	0.03 [-0.44, 0.5]		0.54
Recall First	22			
No recall		0.1 [0, 0.2]	0.09 [-0.12, 0.3]	0.81
Recall		0.2 [0, 0.4]		
List size	22	0 [-0.09, 0.1]		0.52
Buffers				
No buffers	17	0.1 [0, 0.2]	0.1 [-0.1, 0.3]	0.86
Buffers	5	0.2 [0.1, 0.4]		
Activity Measure	22			
No measure	10	0.2 [0, 0.3]	-0.0 [-0.2, 0.1]	0.67
Measure	12	0.1 [0, 0.2]		
Neg-Pos				
Stimulus				
Complex	6	-0.1 [-0.2, 0.1]	0.12 [-0.1, 0.3]	0.89
Words	7	0.0 [-0.1, 0.2]		
Arousal	8	-0.0 [-0.5, 0.5]		0.55
Valence	9	-0.1 [-0.6, 0.4]		0.7
Recall First				
No recall	10	0.0 [-0.2, 0.1]	0.15 [-0.1, 0.4]	0.93
Recall	3	0.1 [-0.1, 0.3]		
List size	13	-0.0 [-0.1, 0.9]		0.64
Buffers				
No buffers	9	0.0 [-0.1, 0.1]	0.08 [-0.1, 0.3]	0.78
Buffers	4	0.1 [-0.1, 0.3]		
Activity Measure				
No measure	8	0 [-0.2, 0.1]	-0.01 [-0.2, 0.2]	0.55
Measure	5	0 [-0.2, 0.2]		

Note. k = the number of observations for each moderator or each level of the moderator; *M* = the mean estimate of directed forgetting for each level of the moderator, in the case of continuous moderators it indicates the slope, 95% confidence interval presented in brackets; Difference (%) = difference in magnitude of directed forgetting between levels of the moderator, 95% confidence interval presented in brackets *p* = Bayesian p-value reflecting confidence in the direction of the effect (e.g., *p* = .95 for a positive effect means 95% confidence that the effect is positive).

Discussion

The current study addressed whether emotional material is more resilient to item-method directed forgetting than neutral material. This question was motivated by several recent studies claiming either comparable (e.g., Quinlan & Taylor, 2014) or reduced/absent directed forgetting for emotional items (e.g., Hauswald, 2010). There had also been the question as to whether negative and positive items were themselves differently or equally affected by intentional forgetting (e.g., Otani et al., 2012). Initial models demonstrated a significant directed forgetting effect for each condition, although these effects were larger for (a) words than for complex stimuli; (b) studies that used buffer items compared to those that did not; (c) studies using recognition for which the recognition test was preceded by a recall test; and, (d) studies using fewer items at encoding. Models comparing the magnitude of directed forgetting between these conditions demonstrated a smaller directed forgetting effect on average for emotional (positive or negative) than neutral material, although minimal difference in the magnitude of directed forgetting between positive and negative items. Importantly, preliminary evidence supports the notion that reduced directed forgetting for emotional material may be driven in part by differences in arousal. In the following sections I will first address the findings in terms of the magnitude of directed forgetting within each emotional category followed by discussion of the comparison of directed forgetting across the emotional categories. Finally, I will discuss the implications of these findings in terms of their relevance to the emotion and directed forgetting literatures as well as some real-word applications.

Directed Forgetting for Neutral and Emotional Conditions

Each emotion category showed an effect of directed forgetting, with better memory for R-items compared to F-items. This is consistent with previous research showing an effect of directed forgetting for all emotion categories (Taylor et al., 2018; Quinlan & Taylor 2014; Gallant et al., 2018) and indicates that regardless of emotional valence, participants are able to effectively control the unwanted F-items in the task, leading to less memory for those items. However, this conflicts with previous findings such as those reported by Hauswald et al. (2010) and Otani et al. (2012) that have shown no effect of directed forgetting specifically for negative items. Therefore, our findings highlight the robustness of the effect. Moderators influencing the magnitude of directed forgetting within the emotional and neutral conditions will be explored in greater detail in the following sections.

Moderators Influencing Directed Forgetting Across Conditions

Although past work would have predicted a larger directed forgetting effect for recall compared to recognition (e.g. Titz & Verhaeghen., 2010), the current analysis demonstrated only weak support for this effect. Specifically, both the neutral and emotional models demonstrated a larger directed forgetting effect for recall than recognition, but the effect was weak and unconvincing in the neutral condition. I attribute this difference to the enhanced representation of emotional items in memory as they tend to be attended to and processed faster (e.g. Schupp et al., 2003), therefore while forget instructions may still impact memory for the item making it inaccessible for recall, participants may still be able to recognize the item. Even so, taken together, findings still tended to favour a directed forgetting effect for recall. This pattern may be expected because recall entails the participants having a strong representation of the item in memory, whereas participants may rely on familiarity for recognition (MacLeod & Kampe, 1996). In the context of directed forgetting it has been shown that forget instructions lead to a reduction in recollection and less detailed representations of the items in memory (e.g., Fawcett et al., 2013; Fawcett & al., 2016; Ahmad et al., 2019). This suggests that forget instructions may be more detrimental to recall than recognition, as recall relies on a more detailed memory of the item (Leshikar et al., 2017). However, the current results found inconsistencies for this pattern across conditions for this claim.

One moderator that consistently influenced the magnitude of directed forgetting across conditions was the type of stimuli used. The current analysis found a larger magnitude of directed forgetting for words compared to complex stimuli, which is not surprising given previous research in the area (e.g., Quinlan et al., 2010). Images typically have an advantage in memory compared to words, which is referred to as the picture superiority effect (Nelson et al., 1976), thought to be driven by the inherently distinctive nature of images compared to words (Ensor et al., 2019). Past research combining directed forgetting with a manipulation of distinctive encoding (i.e., using production) has shown that distinctiveness attenuates the directed forgetting effect (Hourihan & Macleod, 2008). Our results provide additional support for this claim. Notably, this finding was also observed in an earlier meta-analysis by Titz and Verhaeghen (2010), who found reduced directed forgetting for complex stimuli compared to words; however, their analysis included only a single study using images whereas ours included a mixture of 15 studies using images, faces or written narratives.

An unexpected finding is that studies that included buffers demonstrated greater directed forgetting than those that did not. As primacy and recency effects have been shown to influence memory for items presented serially (e.g. Wiswede et al., 2007), studies employ buffers to mitigate these effects. However, typical theories explaining primacy and recency effects would not account for this difference. Primacy effects are thought to be a result of enhanced rehearsal of items presented early in the list (Glenberg et al., 1980), but in the context of directed forgetting F-items should not receive enhanced rehearsal regardless of where they are presented in the study phase.

Recency effects are typically thought to occur due to the lack of subsequent items being presented, therefore the items are able to be maintained in working memory until testing (Craik, 1970), however in this task, following a forget instruction, participants work to push the item out of working memory (e.g., Fawcett & Taylor, 2008). Therefore, an effect of recency for F-items is unlikely and the inclusion of buffer items would have little effect. Another theoretical standpoint on primacy and recency effects may better explain the current finding. Some theorists suggest these effects occur due to the temporal distinctiveness of the items presented at the beginning and end of the learning phase (Bireta et al., 2018; Neath, 1993a; Neath, 1993b), and as previously mentioned, distinctiveness has been shown to lessen the effect of directed forgetting (Hourihan & Macleod, 2008). Therefore, one interpretation of the current results might be that items presented at the beginning and end of the study phase are temporally distinct, leading to less directed forgetting for those items. Using buffer items would work to eliminate this enhancement for F-items specifically presented at the beginning and end as they would otherwise not be enhanced in memory, and therefore may lead to a larger directed forgetting effect in the list overall.

Among studies using recognition, those that also had a recall test preceding it demonstrated greater directed forgetting on average than those not preceded by a recall test. One explanation of this could be testing effects as R-items have been shown to be recalled in the task prior to F-items (e.g. Lee, 2013). This could lead to enhanced memory for those items and therefore a larger directed forgetting effect on the subsequent recognition test, since recall of items has been shown to enhance later recognition of those items (Roediger & Butler, 2011). This finding might also relate to the aforementioned nonmonotonic plasticity hypothesis, as items are competing in memory to be recalled (e.g. Detre et al, 2013; Kereztes & Racsmany, 2013). Recalling R-items at the beginning of the test (e.g., Lee, 2013), indicates that F-items may be less strongly represented, but still activated and interfering with recall of R-items. Therefore, these items may fall into the critical intermediate zone of interference and memory activation, becoming susceptible to downregulation, reducing the probability of recognition on the subsequent test. A similar argument might be made for a RIF-like process whereby retrieving R-items at recall elicits down-regulation of competing F-items via cognitive control, akin to the processes thought to arise during the retrieval practice phase of a typical RIF paradigm (Anderson et al., 1994).

Lastly, the present study found that using fewer items is associated with greater directed forgetting, which is consistent with previous research (Titz & Verhaeghen, 2010). This finding might be explained by participants performing worse on the task as the number of items increases. Raunch and Schmitt (2009) demonstrated that participants show fatigue and attenuated performance on a task that relies on cognitive control (i.e., a Stroop task) as the task progresses. In the context of directed forgetting this could mean that participants rehearse R-items less effectively as the task goes on while also becoming less efficient at pushing the F-items from mind. This could possibly lead to both relatively worse memory for R-items and somewhat greater memory for F-items, lessening the effect of directed forgetting. Further research should determine an appropriate list length to ensure differences at test are due to the effect of directed forgetting rather than other factors.

To summarize, the magnitude of directed forgetting in all conditions was moderated in part by the test task used; while previous findings suggest a greater magnitude of directed forgetting may be found in recall testing, this was inconsistent in the current results. The effect was also moderated by stimulus type, with less directed forgetting for complex stimuli than words, which is likely due to the distinctive nature of images. Furthermore, the inclusion of buffer items augmented the effect of directed forgetting, likely by means of reducing any primacy and recency effects for F-items presented early and late in the study phase. Recall testing prior to recognition testing also lead to increased directed forgetting, possibly due to testing effects. Lastly, there was less of an effect of directed forgetting as the number of items increased, suggesting participants may become fatigued and less efficient on the task as the number of items increases. Now that I have addressed the factors influencing the effect of directed forgetting within the emotion conditions, I will discuss both the difference in directed forgetting between conditions and the moderators influencing this difference.

Comparing the Magnitude of Directed Forgetting Between Neutral and Emotional Items

The present meta-analysis has also revealed the directed forgetting effect to be – on average – of smaller magnitude for emotional than neutral stimuli. This finding resolves the previous debate in the literature where numerous studies provided evidence for (Yang et al., 2016) or against (e.g., Taylor et al., 2018) an effect of emotion on directed forgetting. While this finding may be expected based on previous research (e.g., Kissler & Herbert, 2013; Minor & Hersmann, 2019) demonstrating that emotional stimuli receive preferential processing compared to neutral stimuli, others notably disagreed. In terms of the mechanisms possibly involved in this difference, emotion has been shown to lead to less efficient performance on other tasks requiring cognitive control as well, such as the Stroop task (see Song et al., 2017 for a meta-analysis). Also, research has shown that emotion interferes with inhibitory processes. For example, Rebetez et al. (2015) found that emotion attenuated performance on a stop-signal task and led to more proactive interference on a task where participants had to determine whether a word had appeared in the few trials previous. This demonstrates how emotion can impact cognitive control, which in terms of directed forgetting may impact the active control processes thought to be at play during forget trials (e.g., Fawcett & Taylor, 2008; Fawcett et al., 2013b). As noted in the introduction, similar results have been found using other intentional forgetting paradigms such as list-method directed forgetting (Payne & Corrigan, 2007), retrieval-induced forgetting (Dehli & Brennon, 2009), and for highly arousing negative words in the think/no-think task (Marx et al., 2008). The current findings suggest that not only are emotional memories typically remembered better, they are harder to intentionally forget.

However, it is also possible that unconstrained differences between the emotional and neutral stimulus sets other than valence or arousal may have contributed to the observed differences between these conditions. For example, memory for words can be impacted by characteristics such as word frequency, familiarity and length (Grühn, 2016), as well as concreteness (Fliessbach et al., 2006), orthographical (Glanc & Greene, 2007) and phonological neighbourhood size (Cortese et al., 2004), contextual diversity (Hicks et al., 2005) and/or semantic diversity (Hoffman et al., 2013). Differences in these characteristics – or other characteristics, such as age of acquisition (Marful et al., 2016) – could likewise influence directed forgetting. As summarized in Appendix B, the majority of the studies included in the current analysis failed to provide ratings for these characteristics or provide their wordlists to allow ratings to be calculated. Therefore, differences in these dimensions may have driven the apparent differences between the neutral and emotional conditions. However, the influence of these characteristics is both beyond the purview of the current thesis and would be difficult to analyze as too few studies provided ratings for these variables to reliably meta-analyze their influence. Therefore, firm conclusions as to whether emotion - both valence and arousal - is the main contributing factor to this difference cannot be drawn, as differences in these properties between the emotional and neutral stimuli used may have had an impact. Moderators of this effect that were analysed will be discussed in later sections. Further studies should focus on reporting more detailed descriptions of the stimuli they use to allow the influence of all possible confounds to be analyzed.

When comparing the magnitude of directed forgetting between positive and negative items, no difference was found. This finding is somewhat surprising as some past research using directed forgetting has suggested that negative items in particular are harder to intentionally forget than positive items (Gallant & Dyson 2016; Otani et al., 2012). Other research has demonstrated the differing effects of positively and negatively valenced items on memory more generally and found negative items lead to a stronger enhancement in memory (e.g. Inaba et al., 2005; Szőllősi & Racsmány, 2020). This also fits within the framework predicted by the NEVER model proposed by Bowen et al. (2018). This model predicts that enhanced sensory processing occurs for negative items at each step of memory formation leading to a more detailed memory representation for these items compared to other valances independent of arousal. Within the context of directed forgetting, this would lead to less directed forgetting for negative items compared to both positive and neutral items. Similarly, Migita et al. (2011) found enhancement for negative items in memory occurs mostly due to pre-attentive processes as the item is presented. In the context of the current paradigm, this means the negative items are enhanced in memory prior to the presentation of a memory instruction, further explaining the difficulty intentionally forgetting negative items. However, Migita et al. (2011) found no enhancement for positive items compared to neutral items. The current analysis conflicts with this finding, indicating that whether or not negative items lead to more enhanced memory, these items are impacted to a similar extent as positive items using directed forgetting. This adds to the understanding of emotional memory because it highlights the strength of positive items over our intentional control of memory, which has seen mixed results in the past. Although the effect of emotion on the magnitude of directed forgetting varies greatly from one study to the next, a moderator suspected to be responsible for this heterogeneity will be discussed next.

Moderators Influencing the Magnitude of Directed Forgetting Between Emotional Conditions

While attempting to explain the heterogeneity found between studies, one of the most promising moderators influencing the difference was arousal. In particular, the present findings offer preliminary support for similar claims made in the primary literature (e.g., Hauswald et al. 2010; Taylor et al., 2018). For example, Bailey and Chapman (2012) compared neutral, negative and positive items at high and low arousal levels and in their first study found a significant interaction between arousal, memory instruction and emotion, with less directed forgetting for emotional items at higher arousal levels. In their second experiment the interaction was not statistically significant. Similarly, Gallant and Dyson (2016) compared directed forgetting for negative, positive and neutral items at high and low levels of arousal, with emotional categories matched for arousal, with the neutral items rated lower for arousal in both conditions. They found when arousal was high, both neutral and negative items led to lower magnitudes of directed forgetting compared to positive items, but only negative items led to less directed forgetting when arousal was low. For positive items they saw enhanced memory for R-items when arousal was high, but no change in memory for F-items. Gallant et al. (2018) used items where arousal had been matched between valences and found no effect of valence on the magnitude of directed forgetting. These results and the finding of the current meta-analysis suggest that the decrease in the magnitude of directing for emotional items compared to neutral items is at least partly due to unconstrained differences in arousal rather than differences in valence, but this relationship still needs further investigation.

The effect of arousal was not present for my analysis of *d'*. As indicated above, this could be caused by a lack of statistical power, as the analysis of *d'* was based on far fewer studies than the overall models. Further, inspection of Figure 5 suggests the presence of two influential effects

(Gallant 2018; Zwissler et al., 2011) separated from the other data in terms of their closely matched arousal ratings, despite moderate to large neutral-emotional difference. In short, they could reflect outliers, the influence of which would be magnified in the *d'* models due to the small number of studies. Another possibility is that the effect of arousal is non-linear; this would be expected because once arousal has been matched closely between the emotional and neutral conditions, the nature of the effect should asymptote at 0, or even invert. To evaluate this possibility, a non-linear model was fit to the overall model (see Figure 5). Supporting this assertion, the model appears to decline rapidly, leveling off close to 0 (no difference). However, little is known about the nature of the curve beyond a standardized difference in arousal of -0.25, as demonstrated by the large error bars in that region of the plot. The same non-linear model could not be fit to the *d'* data, due to both the small sample of effects, and range restriction in the arousal ratings for those effects. A third possibility would be that the observed relation with arousal was driven in some way by differences in response bias, which then disappeared due to the use of *d'*. I view this possibility as relatively unlikely. As previously mentioned, since directed forgetting is measured by the difference in proportion of hits for both the remember and forget conditions in each emotion, with a shared false alarm rate, there is little opportunity for response bias to contribute. Nonetheless, this should be explored further by future research.

These findings highlight the importance of better operationalizing what is meant by 'emotion' in the context of memory research, as the current findings reflect different outcomes due to varying levels of arousal but little effect due to valence. Also, this further highlights the idea that arousal may have a bigger influence on our cognitive systems than valence. As previously stated, more arousing stimuli leads to more enhanced memory (Szőllősi & Racsmány 2020), but now the current results have revealed it also makes those memories more difficult to forget intentionally, compared to less arousing or neutral stimuli. This effect may arise as previous research has stated that arousing items are enhanced in memory without the need for elaborate rehearsal strategies, evidenced by enhanced recall while attention is elsewhere at the time of item presentation (Kensinger & Corkin, 2004); as directed forgetting relies on the enhanced rehearsal of some items over others, it is not surprising that there is less of an effect for arousing items.

The only other moderator to demonstrate a credible impact on the difference in directed forgetting between the emotional and neutral condition was the inclusion of buffers. As previously discussed, including buffer items may help to reduce primacy and recency effects and therefore lead to greater directed forgetting. In the context of this comparison, these effects might lead to enhanced memory for both neutral and emotional items presented at the beginning or the end of the study phase, whereas eliminating these effects would lead to a larger difference between conditions as the emotional items are already preferentially processed compared to the neutral items. However, this finding is largely speculative and further research would be required, for example studying serial position effects between neutral and emotional items within an item-method directed forgetting paradigm.

Implications for the Field and Applied Settings

Since our analysis has supported that emotional memories are harder to intentionally forget than neutral ones – at least insofar as those memories are also arousing – this sets the stage for future studies to further investigate the underlying processes that lead to this disparity. Using studies within healthy populations has given us insight into the difficulty of controlling emotional information at a baseline level. Also, studies that have included neuroimaging have indicated the difficulty in controlling these memories arise from either the enhanced inhibitory control needed to push these items out of memory, or the failure of these processes to be employed due to the emotional nature of the items (Yang et al., 2012; Yang et al., 2016). Having established the impact emotional items have on directed forgetting, interventions that aim to improve mechanisms such as cognitive control processes may have implications for controlling these items more efficiently. For example, Ducrocq et al. (2016) demonstrated how an internet-based program aimed at improving cognitive control processes in a visual search task could enhance these processes as well as enhance performance on a task that required an applied use of these processes, in this case tennis. They also found that these improvements persisted as more pressure was added to the participants. Although, many are critical of this type of training and believe improvements are not able to be applied beyond the task used during training (e.g., Sala & Gobet, 2019). Perhaps a similar program aimed at controlling emotional memories in particular could have implications for both the general population to better control emotional memories as well as people suffering with disorders that find the inability to control emotional memories especially detrimental to their lives (e.g. PTSD).

The current findings have implications for other areas outside memory paradigms. In a courtroom setting sometimes jurors are presented with evidence which they are subsequently told to disregard. There has been extensive research showing that jurors have difficulty putting aside and forgetting about this evidence when making their decision (Steblay et al., 2006). Dietvorst and Simonsohn (2019) have shown that making participants aware of the issue of using inadmissible evidence while making decisions and also providing a strong reason not to do so lowered the incidence of using it. The current findings suggest that if this evidence is especially emotional or arousing, jurors may have an even harder time forgetting it while considering a verdict, so being aware of this factor can ensure extra precautions are in place to lessen the impact.

Limitations

One limitation of our analysis is the rarity of studies including positive stimuli, especially those contributing to the *d'* models, which could suggest these analyses were underpowered and therefore it is difficult to draw any strong conclusions from the results. Further research should aim to include positive items in their studies when assessing the influence emotion has on directed forgetting. A second limitation is that not all articles provided arousal and valence ratings for their stimulus set. Having all ratings would have given a clearer picture of the influence these two factors have on directed forgetting. Also, the relatively inconsistent nature of the stimuli more generally restricted our ability to draw firm conclusions as to the origin of the observed effect. Without detailed descriptive statistics of the stimuli used in the experiments provided, the ability to analyze potential confounds was not feasible for the current thesis, although future work investigating these confounds is planned.

Conclusion

In conclusion, this meta-analysis has found supports the argument that emotional items are harder to intentionally forget in the context of item method directed forgetting. While the results provided insight into some of the potential variables affecting this discrepancy (e.g., arousal), the possibility of other contributing factors cannot be ruled out due to the uncontrolled nature of many of the stimuli sets used. In broader terms of directed forgetting, we found there were larger magnitude of directed forgetting in studies that a) used words as stimuli, b) used buffer items c) had recall testing prior to recognition, and d) used a lower number of items. Future research should aim to further investigate the role of arousal in the discrepancy between directed forgetting of emotional and neutral items.

References

Adelman, J. S., & Estes, Z. (2013). Emotion and memory: a recognition advantage for positive and negative words independent of arousal. *Cognition*, 129(3), 530–535. https://doi.org/10.1016/j.cognition.2013.08.014

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). https://doi.org/10.1176/appi.books.9780890425596

Ahmad, F. N., Tan, P., & Hockley, W. E. (2019). Directed forgetting for categorised pictures: recognition memory for perceptual details versus gist. *Memory*, 27(7), 894-903. https://doi.org/10.1080/09658211.2019.1591456

Anderson, M C, Bjork, R A, & Bjork, E L. (1994). Remembering can cause forgetting: Retrieval dynamics in long-term memory. *Journal of Experimental Psychology*. *Learning, Memory, and Cognition*, 20(5), 1063-1087. https://doi.org/10.1037/0278-7393.20.5.1063

Anderson, M. C., & C. Green. Suppressing unwanted memories by executive control. (2001). *Nature* 410.6826: 366-369. https://doi.org/10.1038/35066572

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

Anderson, M. C., & Levy, B. J. (2009). Suppressing Unwanted Memories. *Current Directions in Psychological Science*, 18(4), 189–194. https://doi.org/10.1111/j.1467-8721.2009.01634.x

Augusti, E., & Melinder, A. (2013) The effect of neutral and negative colour photographs on children's item directed forgetting, *European Journal of Developmental Psychology*, 10:3, 378-391, https://doi.org/10.1080/17405629.2012.686741

*Bailey, K., & Chapman, P. (2012). When can we choose to forget? An ERP study into item-method directed forgetting of emotional words. *Brain and Cognition*, *78*(2), 133–147. https://doi.org/10.1016/j.bandc.2011.11.004

Basden B, H., Basden D, R. Directed forgetting: further comparisons of the item and list Methods. *Memory*. 1996;4(6):633-653. doi:10.1080/741941000

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

Basden, B. H., Basden, D. R., & Gargano, G. J. (1993). Directed forgetting in implicit and explicit memory tests: A comparison of methods. *Journal of Experimental Psychology: Learning, Memory, & Cognition*,19, 603–616. https://doi.org/10.1037/0278-7393.19.3.603

Baumann, M., Zwissler, B., Schalinski, I., Ruf, M., Schauer, M., & Kissler, J. M. (2013). Directed forgetting in post-traumatic-stress-disorder: a study of refugee immigrants in Germany. *Frontiers in behavioral neuroscience*, 7, 94. https://doi.org/10.3389/fnbeh.2013.00094

*Berger, N., Crossman, M., & Brandt, K. (2018). No evidence for age-related differences in item-method directed forgetting of emotional words. *Quarterly Journal of Experimental Psychology*, 71(3), 595-604. https://doi.org/10.1080/17470218.2016.1264433

Borenstein, M., Hedges, L. V., Higgins, J. P., & Rothstein, H. R. (2010). A basic introduction to fixed-effect and random-effects models for meta-analysis. *Research synthesis methods*, 1(2), 97-111. https://doi.org/10.1002/jrsm.12

Bowen H, J., Kark S, M., Kensinger E, A. NEVER forget: negative emotional valence enhances recapitulation. *Psychon Bull Rev*. 2018;25(3):870-891. doi:10.3758/s13423-017-1313-9

Bireta, T. J., Gabel, A. J., Lamkin, R. M., Neath, I., & Surprenant, A. M. (2018). Distinctiveness and serial position functions in implicit memory. *Journal of Cognitive Psychology*, 30(2), 222-229. https://doi.org/10.1080/20445911.2017.1415344

Bjork, R. (1970). Positive forgetting: The noninterference of items intentionally forgotten. *Journal of Verbal Learning and Verbal Behavior*, *9,* 255–268. https://doi.org/10.1016/S0022-5371(70)80059-7

Bjork, R. A. (1972). Theoretical implications of directed forgetting. In A. W. Melton & E. Martin (Eds.), *Coding processes in human memory* (pp. 217-235). Washington, D.C.: Winston.

Bowen, H. J., Kark, S. M. & Kensinger, E. A. (2018). NEVER forget: negative emotional valence enhances recapitulation. Psychonomic Bulletin & Review, 25, 870–891. doi:10.3758/s13423-017-1313-9

Bradley, M. M., Sabatinelli, D., Lang, P. J., Fitzsimmons, J. R., King, W., & Desai, P. (2003). Activation of the visual cortex in motivated attention. *Behavioral Neuroscience*, *117*(2), 369–380. https://doi.org/10.1037/0735-7044.117.2.369

*Brandt, K., Nielsen, M., & Holmes, A. (2013). Forgetting emotional and neutral words: An ERP study. *Brain Research*, *1501,* 21–31. https://doi.org/ 10.1016/j.brainres.2013.01.019

Brosch, T., Sander, D., Pourtois, G., & Scherer, K. R. (2008). Beyond fear: Rapid spatial orienting toward positive emotional stimuli. *Psychological science*, 19(4), 362-370. https://doi.org/ 0.1111/j.1467-9280.2008.02094.x.

Bürkner, P. (2017). brms: An R package for Bayesian multilevel models using Stan.

Journal of Statistical Software, 80(1), 1-28. https://doi.org/10.18637/jss.v080.i01

Bürkner, P. (2018). Advanced Bayesian multilevel modeling with the R package brms. *The R Journal, 10*(1), 395-411. https://doi.org/10.32614/RJ-2018-017

Butler, A., & James, J. (2010). The neural correlates of attempting to suppress negative versus neutral memories. *Cognitive, Affective, & Behavioral Neuroscience, 10*(2), 182-194. https://doi.org/10.3758/CABN.10.2.182

Cortese, M. J., Watson, J. M., Wang, J., & Fugett, A. (2004). Relating distinctive orthographic and phonological processes to episodic memory performance. *Memory & Cognition, 32(4), 632–639.* doi:10.3758/bf03195854

Craik, F. (1970). The fate of primary memory items in free recall. *Journal of Verbal Learning and Verbal Behavior*, 9(2), 143-148. https://doi.org/10.1016/S0022-5371(70)80042-1

Crowley, R., Bendor, D., & Javadi, A. H. (2019). A review of neurobiological factors underlying the selective enhancement of memory at encoding, consolidation, and retrieval. *Progress in neurobiology*. https://doi.org/10.1016/j.pneurobio.2019.04.004

Cuthbert, B., Schupp, H., Bradley, M., Birbaumer, N., & Lang, P. (2000). Brain potentials in affective picture processing: Covariation with autonomic arousal and affective report. *Biological Psychology*, *52*(2), 95-111. https://doi.org/10.1016/S0301-0511(99)00044-7

Dehli, L., & Brennen, T. (2009). Does retrieval-induced forgetting occur for emotional stimuli? Cognition and Emotion, 23(6), 1056-1068. https://doi.org/10.1080/02699930802285221

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. https://doi.org/10.1111/j.1467-9280.2006.01725.x

Detre, G., Natarajan, A., Gershman, S., & Norman, K. (2013). Moderate levels of activation lead to forgetting in the think/no-think paradigm. *Neuropsychologia*, 51(12), 2371-2388. https://doi.org/10.1016/j.pneurobio.2019.04.004

Dietvorst, B., & Simonsohn, U. (2019). Intentionally "biased": People purposely use to-be-ignored information, but can be persuaded not to. *Journal of Experimental Psychology. General*, 148(7), 1228-1238. https://doi.org/10.1037/xge0000541

Dolcos, F., LaBar, K. S., & Cabeza, R. (2004). Dissociable effects of arousal and valence on prefrontal activity indexing emotional evaluation and subsequent memory: an event-related fMRI study. *Neuroimage*, 23(1), 64-74. https://doi.org/10.1016/j.neuroimage.2004.05.015

Ducrocq, E., Wilson, M., Vine, S., & Derakshan, N. (2016). Training attentional control improves cognitive and motor task performance. *Journal of Sport & Exercise Psychology*, 38(5), 521-533. DOI: 10.1123/jsep.2016-0052

Ensor, T. M., Surprenant, A. M., & Neath, I. (2019). Increasing word distinctiveness eliminates the picture superiority effect in recognition: Evidence for the physical-distinctiveness account. *Memory & cognition*, 47(1), 182-193. https://doi.org/10.3758/s13421-018-0858-9

Fawcett, E. J., Fairbrother, N., Cox, M. L., White, I. R., & Fawcett, J. M. (2019). The prevalence of anxiety disorders during pregnancy and the postpartum period: A multivariate bayesian meta-analysis. The *Journal of Clinical Psychiatry*,*80*(4), 18r12527. https://doi.org/10.4088/JCP.18r12527

Fawcett, J. M., & Hulbert, J. C. (2020). The many faces of forgetting: toward a constructive view of forgetting in everyday life. *Journal of Applied Research in Memory and Cognition*, 1/2020. *9*(1), 1-18. https://doi.org/10.1016/j.jarmac.2019.11.002

Fawcett, J. M., Lawrence, M. A., & Taylor, T. L. (2016). The representational consequences of intentional forgetting: Impairments to both the probability and fidelity of long-term memory. *Journal of Experimental Psychology*: *General, 145,* 56–81. https://doi.org/10.1037/xge0000128

Fawcett, E. J., & Power, H., & Fawcett, J. M. (in press). Women are at greater risk of OCD than men: A meta-analytic review of OCD prevalence worldwide. *Journal of Clinical Psychiatry.*

Fawcett, J. M., & Ozubko, J. D. (2016). Familiarity, but not recollection, supports the between-subject production effect in recognition memory. *Canadian Journal of Experimental Psychology*, *70*(2), 99-115. http://dx.doi.org/10.1037/cep0000089

Fawcett, J. M., & Taylor, T. L. (2008). Forgetting is effortful: Evidence from reaction time probes in an item-method directed forgetting task. *Memory & Cognition, 6,* 1168–1181. doi:10.3758/MC.36.6.1168

Fawcett, J. M., Taylor, T., & Nadel, L. (2013b). Event-method directed forgetting: Forgetting a video segment is more effortful than remembering it. *Acta Psychologica*, 144(2), 332-343. https://doi.org/10.1016/j.actpsy.2013.07.005

Feldmann-Wüstefeld, T., Schmidt-Daffy, M., & Schubö, A. (2010). Neural evidence for the threat detection advantage: Differential attention allocation to angry and happy faces. *Psychophysiology*, 48(5), 697–707. doi:10.1111/j.1469-8986.2010.01130.x

Ferré, P., Haro, J., & Hinojosa, J. A. (2017). Be aware of the rifle but do not forget the stench: differential effects of fear and disgust on lexical processing and memory. *Cognition and Emotion*, 32(4), 796–811. doi:10.1080/02699931.2017.1356700

Fliessbach, K., Weis, S., Klaver, P., Elger, C. E., & Weber, B. (2006). The effect of word concreteness on recognition memory. *NeuroImage, 32*(3), 1413-1421. doi:http://dx.doi.org.qe2a-proxy.mun.ca/10.1016/j.neuroimage.2006.06.00

*Gallant, S. N. (2013). Effects of aging and emotional valence on item directed forgetting and source attributions (Doctoral dissertation, Ryerson University).

*Gallant, S., & Dyson, B. (2016). Neural modulation of directed forgetting by valence and arousal: An event-related potential study. *Brain Research*, 1648(Pt A), 306-316. doi: 10.1016/j.brainres.2016.08.009.

*Gallant, S. N., Pun, C., & Yang, L. (2018). Age differences in the neural correlates underlying control of emotional memory: An event-related potential study. *Brain research*, 1697, 83-92. https://doi.org/10.1016/j.brainres.2018.06.019

*Gamboa, O., Sung Lai Yuen, K., Von Wegner, F., Behrens, M., & Steinmetz, H. (2018). The challenge of forgetting: Neurobiological mechanisms of auditory directed forgetting. *Human Brain Mapping, 39*, 249-263. https://doi.org/10.1002/hbm.23840

Geraerts, E., & McNally, R. J. (2008). Forgetting unwanted memories: Directed forgetting and thought suppression methods. *Acta psychologica*, 127(3), 614-622. doi:10.1016/j.actpsy.2007.11.003

Glanc, G. A., & Greene, R. L. (2007). *Orthographic neighborhood size effects in recognition memory. Memory & Cognition, 35(2), 365–371.* doi:10.3758/bf03193457

Gelman, A., & Hill, J. (2007). *Data analysis using regression and multilevel/hierarchical models* (Vol. 1). New York, NY, USA: Cambridge University Press.

Glenberg, A. M., Bradley, M. M., Stevenson, J. A., Kraus, T. A., Tkachuk, M. J., Gretz, A. L., Fish, J. H., & Turpin, B. M. (1980). A two-process account of long-term serial position effects. Journal of Experimental Psychology: *Human Learning and Memory*, 6(4), 355–369. https://doi.org/10.1037/0278-7393.6.4.355

Grühn, D. (2016). *An English Word Database of EMOtional TErms (EMOTE). Psychological Reports, 119(1), 290–308.* doi:10.1177/0033294116658474

Hamann, S. (2001). Cognitive and neural mechanisms of emotional memory. Trends in *Cognitive Sciences*, 5(9), 394-400. https://doi.org/10.1016/S1364-6613(00)01707-1

Hanczakowski, M., Pasek, T., & Zawadzka, K. (2012). Context-dependent impairment of recollection in list-method directed forgetting. *Memory*, 20(7), 758-770. https://doi.org/10.1080/09658211.2012.702774

Hauswald, A., & Kissler, J. (2008). Directed forgetting of complex pictures in an item method paradigm. *Memory*, *16*, 797–809. https://doi.org/10.1080/09658210802169087

*Hauswald, A., Schulz, H., Iordanov, T., & Kissler, J. (2010). ERP dynamics underlying successful directed forgetting of neutral but not negative pictures. *Social Cognitive and Affective Neuroscience*, *6*, 450–459. https://doi.org/10.1093/scan/nsq061

Hicks J, L., Marsh R, L., Cook G, I. An observation on the role of context variability in free recall. *J Exp Psychol Learn Mem Cogn*. 2005;31(5):1160-1164. doi:10.1037/0278-7393.31.5.1160

Hindi Attar, C., Müller, M. (2012). Selective attention to task-irrelevant emotional distractors is unaffected by the perceptual load associated with a foreground task (selective attention to emotional distractors). *PLoS ONE*, 7(5), E37186. https://doi.org/10.1371/journal.pone.0037186

Hoffman P, Lambon Ralph MA, Rogers TT. Semantic diversity: a measure of semantic ambiguity based on variability in the contextual usage of words. *Behav Res Methods*. 2013;45(3):718-730. doi:10.3758/s13428-012-0278-x

Hourihan, K.L., & Macleod, C. M. (2008). Directed forgetting meets the production effect: Distinctive processing is resistant to intentional forgetting. *Canadian Journal of Experimental Psychology = Revue Canadienne De Psychologie Experimentale*, *62*(4), 242-246. https://doi.org/10.1037/1196-1961.62.4.242

Hourihan, K. L., Ozubko, J. D., & MacLeod, C. M. (2009). Directed forgetting of visual symbols: Evidence for nonverbal selective rehearsal. *Memory & Cognition*, 37(8), 1059-1068. DOI: 10.3758/MC.37.8.1059

Inaba, M., Nomura, M., & Ohira, H. (2005). Neural evidence of effects of emotional valence on word recognition. *International Journal of Psychophysiology*, 57(3), 165-173. https://doi.org/10.1016/j.ijpsycho.2005.01.002

IntHout, J., Ioannidis, J. P., Rovers, M. M., & Goeman, J. J. (2016). Plea for routinely presenting prediction intervals in meta-analysis. *BMJ open, 6*(7), e010247. https://doi.org/10.1136/bmjopen-2015-010247

Johnson, H. M. (1994). Processes of successful intentional forgetting. *Psychological Bulletin*, 116(2), 274. https://doi.org/10.1037/0033-2909.116.2.274

Jing, J., Qi, M., Gao, H., & Zhang, Q. (2019). The role of forgetting cues in directed forgetting: Ceasing maintenance rehearsal. *Acta Psychologica, 199*, 102922. https://doi.org/10.1016/j.actpsy.2019.102922

Kensinger E. A. (2009). Remembering the details: Effects of emotion. Emotion review : *journal of the International Society for Research on Emotion*, 1(2), 99–113. https://doi.org/10.1177/1754073908100432

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*, 101(9), 3310-3315. https://doi.org/10.1073/pnas.0306408101

Kensinger, E. A., & Schacter, D. L. (2006). Processing emotional pictures and words: Effects of valence and arousal. Cognitive, *Affective, & Behavioral Neuroscience*, 6(2), 110-126. https://doi.org/10.3758/CABN.6.2.110

Keresztes, A., & Racsmány, M. (2013). Interference resolution in retrieval-induced forgetting: Behavioral evidence for a nonmonotonic relationship between interference and forgetting. *Memory & Cognition*, 41(4), 511-518. https://doi.org/10.3758/s13421-012-0276-3

Kuehl, L. K., Wolf, O. T., Driessen, M., Schlosser, N., Fernando, S. C., & Wingenfeld, K. (2017). Effects of cortisol on the memory bias for emotional words? A study in patients with depression and healthy participants using the Directed Forgetting task. *Journal of psychiatric research*, 92, 191-198. https://doi.org/10.1016/j.jpsychires.2017.04.012

Kim, H., Shimojo, S., & O'Doherty, J. (2011). Overlapping responses for the expectation of juice and money rewards in human ventromedial prefrontal cortex. *Cerebral Cortex, 21*(4), 769-776. https://doi.org/10.1093/cercor/bhq145

Kissler J, Herbert C. Emotion, Etmnooi, or Emitoon?--Faster lexical access to emotional than to neutral words during reading. *Biol Psychol*. 2013;92(3):464-479. doi:10.1016/j.biopsycho.2012.09.004

*Korfine, L., & Hooley, J. M. (2000). Directed forgetting of emotional stimuli in borderline personality disorder. *Journal of Abnormal Psychology*, 109(2), 214. https://doi.org/10.1037/0021-843X.109.2.214

Kruschke, J. K., & Liddell, T. M. (2018). The Bayesian New Statistics: Hypothesis testing, estimation, meta-analysis, and power analysis from a Bayesian perspective. *Psychonomic Bulletin & Review*, 25(1), 178-206. https://doi.org/10.31234/osf.io/wntsa

Lang, P. J., Greenwald, M. K., Bradley, M. M., & Hamm, A. O. (1993). Looking at pictures: Affective, facial, visceral, and behavioral reactions. *Psychophysiology*, 30, 261-273. https://doi.org/10.1111/j.1469-8986.1993.tb03352.x

Lee, K. Y., Lee, T. H., Yoon, S. J., Cho, Y. S., Choi, J.S., & Kim, H. T. (2010). Neural correlates of top–down processing in emotion perception: An ERP study of emotional faces in white noise versus noise-alone stimuli. *Brain Research*, 1337, 56-63. https://doi.org/10.1016/j.brainres.2010.03.094

Lee, Y. (2013). Costs and benefits in item-method directed forgetting: Differential effects of encoding and retrieval. *The Journal of General Psychology*, 140(3), 159-173. https://doi.org/10.1080/00221309.2012.750591

Lee, Y. (2018). Withdrawal of spatial overt attention following intentional forgetting: Evidence from eye movements. *Memory, 26*(4), 503-513. https://doi.org/10.1080/09658211.2017.1378360

Lee, Y. S., & Lee, H. M. (2011). Divided attention facilitates intentional forgetting: Evidence from item-method directed forgetting. *Consciousness and Cognition*, 20(3), 618-626. https://doi.org/10.1016/j.concog.2010.09.008

*Lee, Y., & Hsu, Y. (2013). How do we forget negative events? The role of attentional, cognitive, and metacognitive control. *Cognition and Emotion*, 27(3), 401-415. https://doi.org/10.1080/02699931.2012.713326

Leshikar, E. D., Leach, R. C., McCurdy, M. P., Trumbo, M. C., Sklenar, A. M., Frankenstein, A.

N., & Matzen, L. E. (2017). Transcranial direct current stimulation of dorsolateral prefrontal cortex during encoding improves recall but not recognition memory. *Neuropsychologia,* 106, 390-397. https://doi.org/10.1016/j.neuropsychologia.2017.10.022

*Li, G., Wang, L., & Han, Y. (2017). Directed forgetting of negative performed actions is difficult: A behavioural study. *The Quarterly Journal of Experimental Psychology*, 70(1), 53-61. https://doi.org/10.1080/17470218.2015.1120331

*Liu, T., Chen, N., & Cheng, S. (2017). Selective rehearsal is affected by the emotionality of the encoding context in item-method directed forgetting: An event-related potential study. *Biological Psychology*, 123, 15-24. https://doi.org/10.1016/j.biopsycho.2016.11.012

Liu, Y., Zhang, D., & Luo, Y. (2015). How disgust facilitates avoidance: an ERP study on attention modulation by threats. *Social Cognitive and Affective Neuroscience*, 10(4), 598-604. doi:10.1093/scan/nsu094

LoBue, V., & Matthews, K. (2014) The snake in the grass revisited: An experimental comparison of threat detection paradigms, *Cognition and Emotion*, 28:1, 22-35, DOI: 10.1080/02699931.2013.790783

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. https://doi.org/10.1037/1528-3542.8.2.199

MacLeod, C. M. (1975). Long-term recognition and recall following directed forgetting. *Journal of Experimental Psychology: Human Learning and Memory*, 1(3), 271. https://doi.org/10.1037/0278-7393.1.3.271

MacLeod, C. M. (1989). Directed forgetting affects both direct and indirect tests of memory.

Journal of Experimental Psychology: Learning, Memory, and Cognition, 15(1), 13 https://doi.org/10.1037/0278-7393.15.1.13

Macleod, C. M. (1999). The item and list methods of directed forgetting: Test differences and the role of demand characteristics. *Psychonomic Bulletin & Review*, 6(1), 123-129. https://doi.org/10.3758/BF03210819

MacLeod, C. M. (1998). Directed forgetting. In J. M. Golding & C. M. MacLeod (Eds.), *Intentional forgetting: Interdisciplinary approaches*. Mahwah: Lawrence Erlbaum. pp. 157.

MacLeod, C. M., & Kampe, K. E. (1996). Word frequency effects on recall, recognition, and word fragment completion tests. *Journal of Experimental Psychology*: Learning, Memory, and Cognition, 22(1), 132. https://doi.org/10.1037/0278-7393.22.1.132

*Marchewka, A., Wypych, M., Michałowski, J., Sińczuk, M., Wordecha, M., Jednoróg, K., & Nowicka, A. (2016). What is the effect of basic emotions on directed forgetting? Investigating the role of basic emotions in memory. *Frontiers in Human Neuroscience*, *10*, 378. https://doi.org/10.3389/fnhum.2016.00378

Marful A, Gómez-Ariza CJ, Barbón A, Bajo T (2016) Forgetting “Novel” but Not “Dragon”: The role of age of acquisition on intentional and incidental forgetting. PLoS ONE 11(5): e0155110. https://doi.org/10.1371/journal.pone.0155110

*McNally, R. J., Otto, M. W., Yap, L., Pollack, M. H., & Hornig, C. D. (1999). Is panic disorder linked to cognitive avoidance of threatening information?. *Journal of Anxiety Disorders*, 13(4), 335-348. https://doi.org/10.1016/S0887-6185(99)00008-0

*McNally, R. J., Metzger, L. J., Lasko, N. B., Clancy, S. A., & Pitman, R. K. (1998). Directed forgetting of trauma cues in adult survivors of childhood sexual abuse with and without posttraumatic stress disorder. *Journal of Abnormal Psychology*, 107(4), 596. https://doi.org/10.1037/0021-843X.107.4.596

Migita, M., Otani, H., Libkuman, T. M., & Sheffert, S. M. (2011). Preattentive processing, poststimulus elaboration, and memory for emotionally arousing stimuli. *The Journal of General Psychology*, 138(4), 260-280. https://doi.org/10.1080/00221309.2011.604364

Minor, G., & Herzmann, G. (2019). Effects of negative emotion on neural correlates of item and source memory during encoding and retrieval. *Brain Research*, 1718, 32-45. doi:10.1016/j.brainres.2019.05.001

*Moulds, M. L., & Bryant, R. A. (2008). Avoidant encoding in acute stress disorder: a prospective study. *Depression and Anxiety*, 25(12), E195-E198. https://doi.org/10.1002/da.20368

*Moulds, M. L., & Bryant, R. A. (2002). Directed forgetting in acute stress disorder. *Journal of Abnormal Psychology*, 111(1), 175. https://doi.org/10.1037/0021-843X.111.1.175

Neath, I. (1993a). Contextual and distinctive processes and the serial position function. *Journal of Memory and Language*, 32, 820–840. doi:10.1006/jmla.1993.1041

Neath, I. (1993b). Distinctiveness and serial position effects in recognition. Memory & *Cognition*, 21, 689–698. doi:10.3758/BF03197199

Nelson, D. L., Reed, V. S., & Walling, J. R. (1976). Pictorial superiority effect. *Journal of experimental psychology: Human learning and memory*, 2(5), 523. https://doi.org/10.1037/0278-7393.2.5.523

Newman, E., & Norman, K. (2010). Moderate excitation leads to weakening of perceptual representations. *Cerebral Cortex*, 20(11), 2760-2770. https://doi.org/10.1093/cercor/bhq021

Noel, M., Taylor, T. L., Quinlan, C. K., & Stewart, S. H. (2012). The impact of attention style on directed forgetting among high anxiety sensitive individuals. *Cognitive therapy and research*, 36(4), 375-389. DOI 10.1007/s10608-011-9366-y

*Nowicka, A., Marchewka, A., Jednorog, K, Tacikowski, P., & Brechmann, A. (2010). Forgetting of emotional information is hard: An fMRI study of directed forgetting. *Cerebral Cortex*, *21*, 539–549. https://doi.org/10.1093/cercor/bhq117

Ohman, A., Flykt, A., & Esteves, F. (2001). Emotion drives attention: Detecting the snake in the grass. *Journal of Experimental Psychology*, *130,* 466-478. https://doi.org/10.1037/0096-3445.130.3.466

*Otani, H., Libkuman, T., Goernert, P., Kato, K., Migita, M., Freehafer, S., & Landow, M. (2012). Emotion, directed forgetting, and source memory. *British Journal of Psychology*.103(3), 343-358. https://doi.org/10.1111/j.2044-8295.2011.02078.x

*Patrick, R. E., Kiang, M., & Christensen, B. K. (2015). Neurophysiological correlates of emotional directed-forgetting in persons with Schizophrenia: An event-related brain potential study. *International Journal of Psychophysiology*, 98(3), 612-623. https://doi.org/10.1016/j.ijpsycho.2015.01.006

Pastötter, B., Kliegl, O., & Bäuml, K. H. T. (2016). List-method directed forgetting: evidence for the reset-of-encoding hypothesis employing item-recognition testing. *Memory*, 24(1), 63-74. https://doi.org/10.1080/09658211.2014.985589

Payne, B., & Corrigan, E. (2007). Emotional constraints on intentional forgetting. *Journal* *of Experimental Social Psychology*, 43(5), 780-786. https://doi.org/10.1016/j.jesp.2006.07.005

*Pierguidi, L., Righi, S., Gronchi, G., Marzi, T., Caharel, S., Giovannelli, F., & Viggiano, M. (2016). Emotional contexts modulate intentional memory suppression of neutral faces: Insights from ERPs. *International Journal of Psychophysiology*, 106, 1-13. https://doi.org/10.1016/j.ijpsycho.2016.05.008

*Quinlan, C. K., & Taylor, T. L. (2014). "I never forget a face, but in your case I'll be glad to make an exception": Intentional forgetting of emotional faces. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, 68(3), 212–221. https://doi.org/10.1037/cep0000024

Quinlan, C., Taylor, T., & Fawcett, J. (2010). Directed forgetting: Comparing pictures and words. *Canadian Journal of Experimental Psychology = Revue Canadienne De Psychologie Experimentale, 64*(1), 41-46. https://doi.org/10.1037/a0016569

Racsmány, M., Conway, M., Keresztes, A., & Krajcsi, A. (2012). Inhibition and interference in the think/no-think task. *Memory & Cognition*, 40(2), 168-176. DOI: 10.3758/s13421-011-0144-6

R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Rauch, W., & Schmitt, K. (2009). Fatigue of cognitive control in the stroop-task. *In Proceedings of the Annual Meeting of the Cognitive Science Society* 31, 750-755.

Rebetez, M., Rochat, L., Billieux, J., Gay, P., & Van Der Linden, M. (2015). Do emotional stimuli interfere with two distinct components of inhibition? *Cognition and Emotion*, 29(3), 559-567. https://doi.org/10.1080/02699931.2014.922054

Richardson, M. P., Strange, B. A., & Dolan, R. J. (2004). Encoding of emotional memories depends on amygdala and hippocampus and their interactions. *Nature Neuroscience*, 7(3), 278-285. DOI: 10.1038/nn1190

Rizio, A. A., & Dennis, N. A. (2013). The Neural Correlates of Cognitive Control: Successful Remembering and Intentional Forgetting. *Journal of Cognitive Neuroscience*, 25(2), 297–312. doi:10.1162/jocn_a_00310

Roediger, H., & Butler, A. (2011). The critical role of retrieval practice in long-term retention. *Trends in Cognitive Sciences*, 15(1), 20-27. https://doi.org/10.1016/j.tics.2010.09.003

Sala, G., & Gobet, F. (2019). Cognitive training does not enhance general cognition. *Trends in Cognitive Sciences*, 23(1), 9-20. https://doi.org/10.1016/j.tics.2018.10.004

Sahakyan, L., & Foster, N. L. (2009). Intentional forgetting of actions: Comparison of list-method and item-method directed forgetting. *Journal of Memory and Language*, 61(1), 1 34-152. https://doi.org/10.1016/j.jml.2009.02.006

Sahakyan, L., & Kelley, C. (2002). A contextual change account of the directed forgetting effect. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 28(6), 1064-1072. https://doi.org/10.1037/0278-7393.28.6.1064

Scholz, S., & Dutke, S. (2019) Investigating intentional forgetting using pupillometry: no evidence for the ongoing allocation of cognitive resources during forgetting, *Journal of Cognitive Psychology*, *31*,4, 416-437. https://doi.org/10.1080/20445911.2019.1622548

Schupp, H., Junghoefer, M., Weike, A., & Hamm, A. (2003). Attention and emotion: An ERP analysis of facilitated emotional stimulus processing. *Neuroreport*, *14*(8), 1107-1110.

Scott, G.G., Keitel, A., Becirspahic, M. *et al.* The Glasgow Norms: Ratings of 5,500 words on nine scales. *Behav Res* 51, 1258–1270 (2019). https://doi.org/10.3758/s13428-018-1099-3

Segal, S. K., Stark, S. M., Kattan, D., Stark, C. E., & Yassa, M. A. (2012). Norepinephrine-mediated emotional arousal facilitates subsequent pattern separation. *Neurobiology of learning and memory*, 97(4), 465-469. https://doi.org/10.1016/j.nlm.2012.03.010

Song, S., Zilverstand, A., Song, H., Uquillas, F. D. O., Wang, Y., Xie, C., Cheng, Li., & Zou, Z. (2017). The influence of emotional interference on cognitive control: A meta-analysis of neuroimaging studies using the emotional Stroop task. *Scientific reports*, 7(1), 1-9. https://doi.org/10.1038/s41598-017-02266-2

Steblay, N., Hosch, H. M., Culhane, S. E., & McWethy, A. (2006). The impact on juror verdicts of judicial instruction to disregard inadmissible evidence: A meta-analysis. *Law and Human Behavior*, 30(4), 469-492. https://doi.org/10.1007/s10979-006-9039-7

Strange, B., Dolan, R., & McGaugh, J. (2004). β-Adrenergic modulation of emotional memory-evoked human amygdala and hippocampal responses. *Proceedings of the National Academy of Sciences of the United States of America*, 101(31), 11454-11458. DOI: 10.1073/pnas.0404282101

Sutton, T., & Lutz, C. (2019). Attentional capture for emotional words and images: The importance of valence and arousal. *Canadian Journal of Experimental Psychology = Revue Canadienne De Psychologie Experimentale*, 73(1), 47-54. https://doi.org/10.1037/cep0000154

Szőllősi, Á., & Racsmány, M. (2020). Enhanced mnemonic discrimination for emotional memories: the role of arousal in interference resolution. *Memory & cognition*, 10.3758/s13421-020-01035-3. Advance online publication. https://doi.org/10.3758/s13421-020-01035-3

*Taylor, T., Quinlan, L., & Vullings, C. (2018). Decomposing item-method directed forgetting of emotional pictures: Equivalent costs and no benefits. *Memory & Cognition*, *46*, 132-147. https://doi.org/10.3758/s13421-017-0751-y

*Tay, P. K., & Yang, H. (2017). Angry faces are more resistant to forgetting than are happy faces: directed forgetting effects on the identity of emotional faces. *Journal of Cognitive Psychology*, 29(7), 855-865. https://doi.org/10.1080/20445911.2017.1323907

*Tekcan, A. I., Taş, A. Ç., Topçuoğlu, V., & Yücel, B. (2008). Memory bias in anorexia nervosa: Evidence from directed forgetting. *Journal of Behavior Therapy and Experimental Psychiatry*, 39(3), 369-380. DOI: 10.1016/j.jbtep.2007.09.005

Thompson, K. M., Hamm, J. P., & Taylor, T. L. (2014). Effects of memory instruction on attention and information processing: Further investigation of inhibition of return in item-method directed forgetting. *Attention, Perception, & Psychophysics*, 76(2), 322-334. https://doi.org/10.3758/s13414-013-0584-0

Thompson, K. M., & Taylor, T. L. (2015). Memory instruction interacts with both visual and motoric inhibition of return. *Attention, Perception, & Psychophysics*, 77(3), 804-818. https://doi.org/10.3758/s13414-014-0820-2

Titz, C., & Verhaeghen, P. (2010). Aging and directed forgetting in episodic memory: A meta-analysis. *Psychology and Aging*, 25(2), 405-411. https://doi.org/10.1037/a0017225

Tomlinson, T. D., Huber, D. E., Rieth, C. A., & Davelaar, E. J. (2009). An interference account of cue-independent forgetting in the no-think paradigm. *Proceedings of the National Academy of Sciences*, 106(37), 15588–15593. doi:10.1073/pnas.0813370106

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. https://doi.org/10.1016/j.cognition.2015.05.025

Warriner, A.B., Kuperman, V. & Brysbaert, M. Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behav Res* 45**,** 1191–1207 (2013). https://doi.org/10.3758/s13428-012-0314-x

Weller, P., Anderson, M., Gómez-Ariza, C., & Bajo, M. (2013). On the status of cue independence as a criterion for memory inhibition: Evidence against the covert blocking hypothesis. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 39(4), 1232-1245. DOI: 10.1037/a0030335

*Wierzba, M., Riegel, M., Wypych, M., Jednoróg, K., Grabowska, A., & Marchewka, A. (2018). Cognitive control over memory–individual differences in memory performance for emotional and neutral material. *Scientific reports*, 8(1), 1-18.

Wilding, E. (2000). In what way does the parietal ERP old/new effect index recollection? *International Journal of Psychophysiology*, 35(1), 81-87. https://doi.org/10.1016/S0167-8760(99)00095-1

*Wilhelm, S., McNally, R., Baer, L., & Florin. I. (1996) Directed forgetting in obsessive-compulsive disorder. *Behaviour research and therapy* 34, 633-641. https://doi.org/10.1016/0005-7967(96)00040-X

Wingenfeld, K., Terfehr, K., Meyer, B., Löwe, B., & Spitzer, C. (2013). Memory bias for emotional and illness-related words in patients with depression, anxiety and somatization disorders: an investigation with the directed forgetting task. *Psychopathology*, 46(1), 22-27. https://doi.org/10.1159/000338609

Wiswede, D., Rüsseler, J., & Münte, T. (2007). Serial position effects in free memory recall—An ERP-study. *Biological Psychology*, 75(2), 185-193. https://doi.org/10.1016/j.biopsycho.2007.02.002

Woodward, A. E., & Bjork, R. A. (1971). Forgetting and remembering in free recall: Intentional and unintentional. *Journal of Experimental Psychology*, 89(1), 109. https://doi.org/10.1037/h0031188

Wylie, G., Foxe, J., & Taylor, T. (2008). Forgetting as an active process: An fMRI investigation of item-method directed forgetting. *Cerebral Cortex*, *18*, 670–682. DOI: 10.1093/cercor/bhm101

Xia, J., & Evans, L. H. (2020). neural evidence that disengaging memory retrieval is modulated by stimulus valence and rumination. *Scientific reports*, 10(1), 1-11. DOI: 10.1038/s41598-020-64404-7

*Yang, W., Chen, Q., Liu, P., Cheng, H., Cui, Q., Wei, D., Zhang, Q., Qiu, J. (2016). Abnormal brain activation during directed forgetting of negative memory in depressed patients. *Journal of Affective Disorders, 190*, 880-888. https://doi.org/10.1016/j.jad.2015.05.034

*Yang, W., Liu, P., Xiao, X., Li, X., Zeng, C., Qiu, J., & Zhang, Q. (2012). Different neural substrates underlying directed forgetting for negative and neutral images: An event-related potential study. *Brain Research*, 1441, 53-63. https://doi.org/10.1016/j.brainres.2011.10.042

Yang, W., Chen, Q., Liu, P., Cheng, H., Cui, Q., Wei, D., Zhang, Q., & Qiu, J. (2016). Abnormal brain activation during directed forgetting of negative memory in depressed patients. *Journal of Affective Disorders*, 190, 880-888. https://doi.org/10.1016/j.jad.2015.05.034

Zeelenberg, R., Wagenmakers, E. J., & Rotteveel, M. (2006). The impact of emotion on perception: Bias or enhanced processing?. *Psychological Science*, 17(4), 287-291. DOI: 10.1111/j.1467-9280.2006.01700.x

Zwissler, B., Hauswald, A., Koessler, S., Ertl, V., Pfeiffer, A., Wöhrmann, C., Winkler, N., & Kissler, J. (2012). Memory control in post-traumatic stress disorder: evidence from item method directed forgetting in civil war victims in Northern Uganda. *Psychological Medicine*, 42(6), 1283-1291. https://doi.org/ 10.1017/S0033291711002273

*Zwissler, B., Koessler, S., Engler, H., Schedlowski, M., & Kissler, J. (2011). Acute psycho-social stress does not disrupt item-method directed forgetting, emotional stimulus content does. *Neurobiology of Learning and Memory*, 95(3), 346-354. https://doi.org/10.1016/j.nlm.2011.01.007

Appendix A

Effect of Moderators on the Magnitude of Directed forgetting for Negative and Positive Items

Moderator	***k***	***M***	**Difference (%)**	***p***
Negative Items				
Stimulus				
Complex	14	10.3 [6.5, 14.0]	8.8 [3.9, 13.9]	1
Words	26	19.1 [16.0, 22.4]		
Task				
Recall	15	21.0 [16.1, 26.0]	-6.7 [-12.7, -0.6]	0.98
Recognition	25	14.3 [10.8, 17.8]		
Arousal	25	5.3 [-26.2, 35.7]		0.64
Valence	28	2.87 [-40.8, 47.5]		0.55
Recall First				
No Recall	28	12.1 [8.8, 15.4]	10.4 [2.8, 18.0]	0.99
Recall	12	22.5 [15.7, 29.3]		
List Size	40	-3.23 [-6.1, -0.3]		0.99
Buffer				
No buffers	24	12.5 [9.3, 15.7]	8.41 [3.0, 13.9]	1
Buffers	15	20.9 [16.5, 25.3]		
Activity Measure				
No measure	25	17.1 [13.5, 20.9]	-3.7 [-9.6, 1.9]	0.90
Measure	15	13.4 [8.8, 17.8]		
Positive Items				
Stimulus				
Complex	8	9.0 [2.4, 15.5]	8.5 [0.3, 16.6]	0.98
Words	18	17.6 [12.5, 22.5]		
Task				
Recall	11	16.0 [7.8, 22.8]	-0.7 [-10.5, 9.8]	0.57
Recognition	15	15.3 [8.4, 22.1]		
Arousal	17	-11.8 [-61.5, 37.5]		0.68
Valence	20	23.65 [-23.4, 70.1]		0.84
Recall First				
No Recall	16	8.9 [4.5, 13.5]	20.6 [11.2, 29.5]	1
Recall	10	29.5 [21.3, 37.1]		
List size	26	-2.49 [-7.4, 2.2]		0.85
Buffer				
No buffers	15	10.5 [5.5, 15.3]	10.3 [2.2, 18.4]	0.99
Buffers	10	20.8 [14.3, 27.2]		
Activity Measure				
No measure	21	15.6 [10.8, 20.3]	-6.2 [-17.0, 4.4]	0.88
Measure	5	9.3 [-0.5, 19.0]		

Note. k = the number of observations for each moderator or each level of the moderator; M = the mean estimate of directed forgetting for each level of the moderator, in the case of continuous moderators it indicates the slope, 95% confidence interval presented in brackets; Difference (%) = difference in magnitude of directed forgetting between levels of the moderator, 95%

Effect of Moderators on the Magnitude of Directed forgetting for Negative and Positive Items in d'

Moderator	*k*	*M*	Difference(%)	p
Negative Items				
Stimulus				
Complex	10	0.3 [0.1, 0.5]	0.3 [0.1, 0.5]	0.99
Words	12	0.6 [0.4, 0.7]		
Arousal	14	-0.19 [-1.1, 0.7]		0.66
Valence	15	0 [-0.9, 0.9]		0.5
Recall First				
No Recall	18	0.4 [0.2, 0.5]	0.2 [-0.1, 0.6]	0.92
Recall	4	0.6 [0.3, 0.9]		
List size	22	-0.08 [-0.2, 0.1]		0.87
Buffers				
No Buffers	17	0.4 [0.2, 0.5]	0.3 [-0.0, 0.6]	0.96
Buffers	5	0.6 [0.4, 0.9]		
Activity Measure				
No measure	10	0.5 [0.3, 0.7]	-0.2 [-0.4, 0.1]	0.90
Measure	12	0.4 [0.2, 0.5]		
Positive Items				
Stimulus				
Complex	7	0.2 [0.0, 0.5]	0.45 [0.12, 0.78]	0.99
Words	7	0.7 [0.4, 0.9]		
Arousal	9	-0.12 [-1.1, 0.8]		0.6
Valence	10	0.05 [-0.9, 1.0]		0.53
Recall First				
No recall	11	0.4 [0.1, 0.6]	0.41 [-0.0, 0.8]	0.97
Recall	3	0.8 [0.4, 1.1]		
List Size	14	0.01 [-0.2, 0.2]		0.52
Buffers				
No Buffers	11	0.3 [0.1, 0.5]	0.53 [0.1, 0.9]	0.99
Buffers	3	0.9 [0.5, 1.2]		
Activity Measure				
No measure	10	0.5 [0.2, 0.7]	-0.1 [-0.6, 0.3]	0.68
Measure	4	0.4 [0.0, 0.8]		

Note. *k* = the number of observations for each moderator or each level of the moderator; *M* = the mean estimate of directed forgetting for each level of the moderator, in the case of continuous moderators it indicates the slope, 95% confidence interval presented in brackets; Difference (%) = difference in magnitude of directed forgetting between levels of the moderator, 95% confidence interval presented in brackets p = Bayesian p-value reflecting confidence in the direction of the effect (e.g., p = .95 for a positive effect means 95% confidence that the effect is positive).

Effect of Moderators on the Difference in the Magnitude of Directed forgetting Between Emotion Conditions

Moderator	***K***	***M***	**Difference(%)**	***p***
Neu-Neg				
Stimulus				
Complex	13	5.3 [0.7, 9.9]	-2.4 [-8.3, 3.4]	0.8
Words	26	2.9 [-0.7, 6.4]		
Task				
Recall	15	2.3 [-1.9, 6.5]	2.6 [-2.2, 7.1]	0.87
Recognition	24	4.9 [1.6, 8.2]		
Arousal	24	-16.1 [-44.6, 14.9]		0.86
Valence	28	8.84 [-23.7, 41.2]		0.7
Recall First				
No recall	27	3.7 [0.6, 6.9]	4.6 [-2.5, 11.8]	0.9
Recall	12	8.3 [1.9, 14.8]		
List size	39	0.43 [-2.65, 3.5]		0.61
Buffers				
No Buffers	23	2.7 [-0.6, 6.1]	4.2 [-1.6, 10.0]	0.92
Buffers	15	6.9 [2.2, 11.5]		
Activity measure				
No measure	22	3.1 [-0.6, 6.9]	1.7 [-4.1, 7.4]	0.72
Measure	17	4.8 [0.5, 9.1]		
Neu-Pos				
Stimulus				
Complex	7	6.2 [2.0, 10.2]	-2.3 [-7.2, 3.1]	0.82
Words	18	3.9 [1.1, 7.1]		
Task				
Recall	11	3.7 [0.5, 7.2]	1.66 [-3.0, 6.1]	0.78
Recognition	14	5.4 [2.0, 8.8]		
Arousal	16	-10.2 [-40.1, 21.5]		0.75
Valence	20	2.79 [-32.2, 37.7]		0.56
Recall First				
No recall	15	4.5 [0.7, 8.5]	2.7 [-5.2, 10.4]	0.76
Recall	10	7.2 [0.3, 13.9]		
List size	25	0.1 [-2.5, 2.8]		0.53
Buffers				
No Buffers	14	3.7 [0.9, 6.7]	4.2 [-1.2, 9.3]	0.94
Buffers	10	7.9 [3.6, 12.2]		
Activity Measure				
No measure	20	4.0 [1.0, 7.1]	2.4 [-3.2, 8.2]	0.81
Measure	5	6.4 [1.8, 11.5]		

Note. *k* = the number of observations for each moderator or each level of the moderator*;* *M* = the mean estimate of directed forgetting for each level of the moderator, in the case of continuous moderators it indicates the slope, 95% confidence interval presented in brackets; Difference (%) = difference in magnitude of directed forgetting between levels of the moderator, 95% confidence interval presented in brackets p = Bayesian p-value reflecting confidence in the direction of the effect (e.g., p = .95 for a positive effect means 95% confidence that the effect is positive).

Moderators Influencing the Difference in the Magnitude of Directed Forgetting Between Emotion Conditions in d'

Moderator	***k***	***M***	**Difference(%)**	***p***
Neu-Neg				
Stimulus				
Complex	9	0.1 [-0.1, 0.3]	0.0 [-0.2, 0.2]	0.64
Words	12	0.1 [0.0, 0.3]		
Arousal	13	-0.1 [-0.5, 0.4]		0.63
Valence	15	0.1 [-0.4, 0.5]		0.58
Recall First				
No recall	17	0.1 [0.0, 0.2]	0.17 [-0.1, 0.4]	0.94
Recall	4	0.3 [0.1, 0.4]		
List size	21	0.03 [-0.1, 0.1]		0.69
Buffers				
No buffers	16	0.1 [0.0, 0.2]	0.16 [-0.1, 0.4]	0.94
Buffers	5	0.2 [0.1, 0.4]		
Activity Measure				
No measure	9	0.1 [0.0, 0.3]	0.0 [-0.2, 0.2]	0.5
Measure	12	0.1 [0.0, 0.2]		
Neu-Pos				
Stimulus				
Complex	6	0.3 [0.1, 0.4]	-0.1 [-0.3, 0.1]	0.87
Words	7	0.2 [0.0, 0.3]		
Arousal	8	-0.01 [-0.5, 0.5]		0.52
Valence	10	0.0 [-0.5, 0.5]		0.51
Recall First				
No recall	10	0.2 [0.1, 0.3]	0.01 [-0.2, 0.2]	0.71
Recall	3	0.2 [0.0, 0.4]		
List size	13	-0.02 [-0.1, 0.1]		0.63
Buffers				
No buffers		0.2 [0.1, 0.3]	0.1 [-0.2, 0.3]	0.68
Buffers		0.3 [0.1, 0.4]		
Activity Measure				
No measure	9	0.2 [0.1, 0.3]	0.0 [-0.2, 0.2]	0.65
Measure	4	0.2 [0.1, 0.4]		

Note. *k* = the number of observations for each moderator or each level of the moderator; *M* = the mean estimate of directed forgetting for each level of the moderator, in the case of continuous moderators it indicates the slope, 95% confidence interval presented in brackets; Difference (%) = difference in magnitude of directed forgetting between levels of the moderator, 95% confidence interval presented in brackets *p* = Bayesian p-value reflecting confidence in the direction of the effect (e.g., *p* = .95 for a positive effect means 95% confidence that the effect is positive).

Appendix B

Information Provided for Studies using Word Stimuli

Paper	**Wordlist Provided**	**Dominance Rating Provided**	**Frequency Rating Provided**	**Frequency Matched**	**Length Matched**	**Familiarity Matched**
Bailey and Chapman (2012, Experiment 1)	No	No	No	No	No	No
Bailey and Chapman (2012, Experiment 2)	No	No	No	No	No	No
Berger et al. (2018)	No	No	Yes	Yes	No	No
Brandt et al. (2013)	No	No	No	Yes	No	No
Gallant (2013)	Yes	No	No	Yes	No	No
Gallant and Dyson (2016)	Yes	No	Yes	Yes	Yes	No
Gallant et al. (2018)	No	No	Yes	Yes	Yes	No
Gamboa et al. (2017)	No	No	No	No	Yes	No
Korfine and Hooley (2000)	Yes	No	No	Yes	Yes	No
McNally et al. (1998)	Yes	No	No	Yes	No	No
McNally et al. (1999)	No	No	No	Yes	Yes	No
Moulds and Bryant (2002)	Yes	No	No	Yes	No	No
Moulds and Bryant (2008)	Yes	No	No	Yes	No	No
Patrick et al. (2015)	No	Yes	Yes	No	Yes	No
Teckan et al. (2008)	No	No	No	Yes	No	No
Wierzba et al. (2018)	No	No	No	No	No	No
Wilhelm et al. (1996)	Yes	No	No	Yes	Yes	No
Yang et al. (2016, Experiment 1)	No	No	No	Yes	No	Yes
Yang et al. (Experiment 2)	No	No	No	Yes	No	Yes