# THE INFLUENCE OF EMOTION ON REMEMBERING AND FORGETTING

**Layla S. Akacem**

Thesis submitted to the University of Nottingham for the degree of

**Doctor of Philosophy**

School of Psychology
University of Nottingham

August 2020

## Abstract

How emotion influences the ability to control what we remember and forget remains unclear. The objective for the work described in this thesis was to develop a better understanding of the links between emotion, memory and memory control. This was achieved by acquiring behavioral and ERP data in a series of directed forgetting (DF) experiments. The DF item-method was used throughout, with different retrieval task requirements. These different retrieval requirements permitted an analysis of how emotion links to memorability, to directed forgetting, and to two different retrieval processes – recollection and familiarity.

A key goal in this work was to exert tight control over stimulus parameters and maintain that control consistently over studies. The intention was to permit contrasts between outcomes across the set of experiments with fewer degrees of freedom than is the case when comparing outcomes among published experiments from different researchers. This is an important consideration because the relevant literature contains many contradictory findings for which there are several competing explanations. Alongside several parameters over which control is commonly deployed, in the experiments described here semantic relatedness was also controlled for. This has not been done consistently in studies of the links between emotion and memory, and the outcomes here and elsewhere suggest that controlling for this factor, which likely co-varies with emotion, is important.

Key behavioral findings in these experiments were: mixed evidence for superior memory for emotional materials (with this term being used here to refer to positive and negative valence words); no evidence that DF varies according to valence, a consistent liberal response criterion for emotional materials, with the exception of Experiment 1 and when valence is manipulated at study only (Experiment 5); and mixed findings for the influence of emotion on recollection and familiarity. In addition, and for the first time, the data in the final experiment

suggest that the liberal response criterion associated with emotional materials can be attributed to processes that occur at the time of retrieval rather than at the time of encoding.

In three event-related potential (ERP) experiments the test phase data provided no new insights over and above what can be inferred from the behavioral data described above. There were, however, informative outcomes from study phase data where ERPs were acquired time-locked to the cues to remember/forget that followed the critical items. Neural activity elicited by remember and forget cues varied according to emotion, implicating valence in the strategies people use to remember and forget material.

In summary, across a series of experiments there was no behavioral evidence that memory control (operationalized as control over remembering and forgetting) varied with valence, but the ERP data indicate that remembering and forgetting does vary with emotion. Moreover, in what is perhaps the other substantive new insight in this thesis, the data in the final experiment suggest that the liberal response criterion associated with emotional materials can be attributed to processes that occur at the time of retrieval rather than at the time of encoding.

# ACKNOWLEDGEMENTS

First and foremost, I am profoundly grateful to my supervisors Professor Ed Wilding and Dr. Chris Madan who gave me invaluable support and guidance throughout my PhD. Their patience and encouragement have been of essence in the successful completion of this thesis.

I would like to thank my fellow office members Persefoni, Alexander, Louise, Nandini, Sara, Barbara and Tami for their friendship, support and being my family away from home.

Finally, my heartfelt gratitude to my parents, Professor Kada and Suzanne, for their moral support and encouragement throughout my PhD. Their love and friendship have been invaluable in preparing me for my future. Thank you, dad, for being my inspiration and leading by example. Mom, thank you for your selfless love.

I would also like to thank my brothers, Fayssal and Salim, and sister, Aynur, for their love and support, as well as my mother-in-law, Julia (Akua), for her love and support. To my fiancé Dennis (ftp), I would like to express my gratitude for his unconditional love and trust. His support has been my backbone and motivation.

## Table of Contents

## List of Tables

## List of Figures

# CHAPTER 1

# 1. GENERAL INTRODUCTION

## 1.1 Remembering and Forgetting Emotional Material

The focus of the research described in this thesis is on the links between emotion and memory control, the ability to exert control over what we remember and what we forget . This is motivated by theoretical and practical perspectives. It is uncontroversial to state that the links between emotion and memory control have been investigated and are not thoroughly understood (Blaney, 1986). One reason for this is the variability in findings in the literature focused on emotion and memory control. Although it is generally accepted that emotion enhances memory, sometimes referred to as EEM (the Emotion Enhanced Memory effect; Bradley, Greenwald, Petry, & Lang, 1992), whether emotional information is more difficult to forget than neutral material remains unclear. A general assumption is that emotional material is more resistant to forgetting compared to neutral material (Kensinger & Kark, 2018), however this is not always supported by empirical findings in the literature (Brandt et al., 2013; Harris et al., 2010; Yang et al., 2012).

The ability to forget is considered to be essential in service of efficient memory processing and goal-directed behaviour (Engen & Anderson, 2018). In the context of emotion, emphasis has been placed on the forgetting of emotionally negative information as a motivation to preserve our emotional state or to protect the sense of self (Anderson & Hanslmayr, 2014). When forgetting, or memory control, fails, then emotional memory intrusions can disrupt daily functioning. For example, in post-traumatic stress disorder (PTSD) an impairment of voluntary memory retrieval has been implicated (Brewin, 2018).

A key motivation for the work in this thesis, however, is the variability in findings in published work. This variability means that it is challenging to draw strong conclusions over how remembering and forgetting of emotional information may come about. To anticipate the fuller articulation in later sections, some studies have found emotional material to be more resistant to forgetting compared to neutral material (Bailey & Chapman, 2012; Blix & Brennen, 2012; Liu et al., 2017; Minnema & Knowlton, 2008), while others have found no differences between emotional and neutral material in the ability to forget (Barnier et al., 2004; Gallant & Yang, 2014; Wessel & Merckelbach, 2006).

### 1.1.1 Issues Concerning the Variability in Findings

#### 1.1.1.1 Stimulus Properties

In order to improve understanding of the links between emotion and memory control, and the variability in published findings, several issues have been addressed in the experiments discussed below. Briefly, care was taken to control for specific stimulus properties in order to avoid potential confounds. This was done consistently across all experiments, recognising that a likely explanation for disparate findings across published studies is inconsistency in control over stimulus properties.

From the perspective of the work in this thesis, foremost among these properties is semantic relatedness. The degree of semantic relatedness facilitates memory retrieval (Tulving & Pearlstone, 1966). Emotional words tend to be higher in semantic relatedness than neutral words, thus if this factor is not controlled for, then effects due to emotion and relatedness may be confounded. Control over relatedness has not been exercised consistently in the literature (for discussion, see Minnema & Knowlton, 2008). Previous studies have found no differences in memory performance when semantic relatedness did not differ between emotional and neutral material (Dougal & Rotello, 2007; Maratos et al., 2000; Talmi & Moscovitch, 2004). Based on these findings, the prediction in this thesis is that there will be no effect of emotion on memory control when comparable control over stimulus properties is exerted (as it is in all experiments reports in this thesis).

#### 1.1.1.2 Response Criterion

When examining memory for emotional materials it is important to consider response criterion as well as sensitivity (Dougal & Rotello, 2007; Snodgrass & Corwin, 1988; Windmann et al., 2002; Yonelinas, 1994). When criterion changes across conditions, measures of memory sensitivity (*d'*) are difficult to interpret if single-point measures are used (for a more detailed discussion, see section 1.2.1) (Dougal & Rotello, 2007; Kapucu et al., 2008; Macmillan & Creelman, 1990). There are several studies of the links between emotion and memory where single-point measures have been employed and where criterion has varied (Bailey & Chapman, 2012; Dougal & Rotello, 2007; Hauswald et al., 2011; Kapucu et al., 2008). Dougal and Rotello (2007) have argued that differences ascribed to memory may commonly reflect differences in criterion instead, which challenges the EEM assumption. This will be tested rigorously in this thesis.

One way to disentangle criterion and sensitivity is by using Receiver Operating Characteristics (ROCs) (Yonelinas & Parks, 2007). A commonly used method to plot ROCs is by acquiring multiple data points using confidence judgements. This was done in two of the experiments presented in this thesis; in one by collecting confidence judgements and in the other one by using a combination of confidence judgements and the Remember/Know procedure in a recognition memory test.

The third, overlapping, question that has received somewhat less attention in the memory, emotion and memory control literature (although see Liu et al., 2017) is whether changes with emotion (referring here to both criterion and sensitivity) are a result of processes operating during encoding and/or retrieval. Attributes of a retrieval cue may influence what information is retrieved (Keele, 1972; Moscovitch, 1992; Schacter et al., 1978). Thus, emotional attributes of the retrieval cue may influence the information to be retrieved.

Alongside this, however, several theories linking emotion and memory focus on operations occurring at the time of encoding. In fact, the predominant focus has been on operations engaged at encoding (Brandt et al., 2013; Gallant & Dyson, 2016; MacLeod, 1975; Van Hooff et al., 2009). Given the fact, however, that retrieval operations clearly influence measures of memorability (Liu et al., 2017; Maratos & Rugg, 2001; Smith et al., 2004; Taylor, Cutmore, et al., 2018), Experiment 5 in this thesis tests the possibility that operations at the time of retrieval influence measures of memory for emotional content.

### 1.1.2 Experimental Approach

This section covers the primary task employed in the experimental work in this thesis, and the main outcome measures as well. There are several tasks and variants that have been employed to investigate memory control (Anderson & Hanslmayr, 2014; Anderson & Spellman, 1995; Levy & Anderson, 2002). The directed forgetting paradigm is the task of choice in this thesis (MacLeod, 1975, 1999; Muther, 1965) (the rationale for this choice is described in section 1.4.3). Briefly, the task has been widely employed, it is one where there are inconsistencies in the published emotion/memory literature, and it is widely accepted that it can assess retrieval control, because it requires people to attempt to remember and forget memoranda they are presented with.

In terms of outcome measures, a combination of behavioural and electrophysiological measures was used in this thesis. Some of the behavioural measures have been anticipated in earlier sections: sensitivity and criterion, receiver operating characteristics (ROCs), and variants of the remember/know (R/K) paradigm. These comprise relatively standard tools for measuring memory and the contributions of specific processes. The preceding section described the links between emotion, control and memory only in terms of outcomes and not in terms of the retrieval processes that might be involved in those outcomes. Two processes that are likely implicated are recollection and familiarity. The use of ROCs and the R/K paradigm was motivated in order to ask specific questions about how emotion and memory control are linked. Recollection, but not familiarity, is considered to be a process over which control can be exerted. This motivates the general two-part hypothesis tested in this thesis, which is that, while any general influences of emotion on memory can influence recollection and familiarity, control will only influence recollection.

In specific experiments in this thesis, behavioural measures are complemented by event-related potentials (ERPs). Their use was motivated by observations in prior literature of ERPs being sensitive to changes in recollection and familiarity (Friedman & Johnson, 2000; Rugg & Curran, 2007; Sanquist et al., 1980; Vilberg & Rugg, 2008). Implementing the use of ROC curves not only created the opportunity to disentangle measures of sensitivity and criterion (see section 1.1.1.2 above and section 1.3.2 below), but also allowed the extraction of measures of recollection and familiarity. In this thesis this was also achieved in a follow up study using the Remember/Know (R/K) procedure (Tulving, 1985b). The objective of measuring recollection and familiarity using the R/K procedure was to investigate the generality of the findings obtained using ROCs. Although built on somewhat different assumptions, according to Yonelinas (2002) the R/K measures generally correspond well with the estimates derived from ROCs.

The rationale for using a combination of electrophysiological and behavioural measures is because when using ROCs and the R/K procedure, the extent to which participants rely on recollection and familiarity may differ compared to when using a simple old/new recognition task (Tulving, 1985b). Changes in the amplitudes of ERP old/new effects can be used to assess whether familiarity or recollection changes, and these can be employed in an old/new recognition memory task. The ERPs provide, therefore, a means of assessing whether estimates gained via ROCs and the R/K paradigm generalise to standard recognition paradigms. It is generally assumed that estimates acquired via these different tasks do not influence the contributions the processes make. On this basis, the prediction tested in this thesis is that the ERP markers of recollection and familiarity will permit similar conclusions about the links between these processes, emotion and memory control as the outcomes of the ROC and R/K experiments.

Briefly, the left-parietal old/new effect has been found to be a correlate of recollection and the mid-frontal old/new effect as a correlate of familiarity. If memory control and emotion have some influence on these processes, there should be some divergences in these old/new effects between conditions. The detailed justification for using ERPs to predict process contributions is provided in section 1.6.

### 1.1.3 Overview Introductory Sections

The preceding sections identify the key motivations and questions that are pursued in this thesis. The detailed elaboration of each of these is contained primarily in the General Introduction that immediately follows this section. The Introduction has several components and provides an overview of key literature and an account of how literature leads to the empirical work that is reported in subsequent chapters.

This starts with a focus on general memory process and systems, and the key separation between encoding and retrieval, as well as the recognition of the dependencies between them. The subsequent focus on recognition memory and the processes and measures commonly used in combination with this task reflects the importance of this task and variants for the empirical work in this thesis. This task has been employed – with changes in instructions – to investigate control over retrieval as well, and the next phase of the Introduction elaborates on that content, as well as providing the background and justification for the use of the directed forgetting manipulation in much of the empirical work in this thesis. Several other means of investigating memory control are also accounted for.

This set of sections provides the building blocks for then considering the links between memory systems and processes and their interactions with emotion, and then an extension into the links between memory control and emotion. Finally, the summary and reiteration of key predictions at the end of the Introduction is preceded by a detailed account of how ERPs can be employed in memory tasks to complement conclusions and recollection and familiarity that are derived from behavioural data. These sections also include (i) methodological and theoretical considerations germane to the use of the cognitive electrophysiology in this way, and (ii) a specific focus on the directed forgetting literature in which ERPs have been employed alongside behavioural measures. In later chapters, the use of ERPs to also make inferences about the processes engaged at the time of memory encoding and control is returned to. The somewhat smaller literature on encoding, emotion and memory relative to retrieval explains this organisational decision. Finally, as already anticipated the General Introduction concludes with a summary of the key issues to be addressed in this thesis.

## 1.2 Memory Processes and Systems

It is broadly acknowledged that human memory can be separated into several systems and processes (Squire, 2004; Tulving, 1972, 1985a). One of the most common separations is between episodic and semantic memory. Episodic memory supports memory for specific personally experienced events (Tulving, 1972), and has been described as memory that allows recovery of information about ‘when’, ‘where’ and ‘what’ occurred (Tulving, 1972; Clayton & Dickinson, 1998; but also see Easton & Eacott, 2008). Semantic memory supports general knowledge about word meanings and other verbal symbols and is necessary for the use of language (Gardiner, 2001; Tulving, 1972, 1985b).

The focus in this thesis is on episodic memory and specifically how emotion is linked to control over episodic memory. Episodic memory – as is the case for all types of memory - can be conceived of as having three stages: encoding, storage and retrieval. First, it is necessary for received information to be encoded, and this can happen with varying degrees of efficiency (Craik & Lockhart, 1972). Storage refers to the creation of a stable and possibly permanent record of encoded information, which is also referred to as consolidation (McGaugh, 2000). Retrieval involves both reactivating stored information and using it in service of task demands (Eustache et al., 1999; Schacter, Norman, & Koutstaal, 1998). The processes of encoding and retrieval, as well as the links between them, are of specific interest in this thesis.

### 1.2.1 Encoding Frameworks

There are several accounts of the relationship between memory encoding and memory retrieval (Buchanan, 2007; Tulving & Thomson, 1973). The encoding specificity principle is one such account (Tulving & Thomson, 1973). The key premise of this account is that how information is encoded determines what retrieval cues will be most successful in retrieving stored information. Tulving and Thompson (1973) conducted a set of three experiments using a cued-recall task and interpreted the results as support for this principle. In all three experiments, participants studied 2 lists of weak cue-target word pairs (e.g. *ground* - *cold)*. They were instructed to study the target words and pay attention to the cues. After each list participants were presented with the weak cues and instructed to report any target words they recalled from the encoding phase. A third list of cue-target word pairs was then presented, and in each experiment the instructions and the tasks differed. For example, in one experiment participants were presented with strong cues (e.g. *hot)*, replacing the weak cues presented during the encoding phase. These strong cues were not presented earlier during study. During the test phase, participants were instructed to generate free association responses to half of the strong cues and indicate which of these generated words they recognized as target words. They were then presented with the weak cues (previously presented during the encoding phase) in a cued-recall test and instructed to report the target words of the cue-word pairs. There was higher recall of target words when participants were presented with the weak cues than when presented with strong cues (Tulving & Thomson, 1973). This is consistent with the encoding specificity principle because strong cues should be more likely to elicit related targets, but the specific way in which pairs were encoded resulted in an outcome that is counter to this prediction (for similar findings see Higham, 2002). In other words, the circumstances under which information is encoded and stored, determines what retrieval cues are most effective in gaining access to the stored information (Tulving & Thomson, 1973).

According to a second principle - the transfer-appropriate processing principle - the likelihood of successful retrieval is higher when information is processed in a similar manner during encoding and retrieval (Morris et al., 1977) than when there is not the same degree of correspondence between encoding and retrieval operations. Morris et al. (1977) reported the outcomes of experiments in which participants studied some words in a semantic encoding condition, and others in a rhyming condition. When they were tested in a retrieval task focusing on rhyming information their memory was better for the information encoded in the rhyming condition. The reverse was true when the testing condition focused on semantic information.

### 1.2.2 Retrieval Frameworks

Several models have been proposed to explain how the process of retrieval works (Baddeley, Lewis, Eldridge, & Thomson, 1984; Gillund & Shiffrin, 1984; Schacter, Eich, & Tulving, 1978; Tulving & Thomson, 1973), in which retrieval cues play a crucial role (Keele, 1972; Moscovitch, 1992). A retrieval cue can be generated by an external stimulus (e.g. a picture or word during a memory task) which then generates an internal retrieval cue - a mental representation of the stimulus. This cue might activate a memory trace (cue-trace interaction), depending upon the quality and form of the cue, the qualities and forms of memory traces, and the overlap between the cue and memory traces (Morris et al., 1977; Tulving & Thomson, 1973). While as stated cues can be generated by external events, they can also be generated by internal thought processes (Gillund & Shiffrin, 1984).

Semon's theory of memory provides one starting point for considering how retrieval processes operate (Schacter et al., 1978; Semon, 1904). According to Semon, information is processed in fragments that are linked together. Which fragments are attended to influences which information is most likely to be retrieved, and only a fragment of information needs to be retrieved in order to enable retrieval of linked fragments. In his theory, Semon (1904) also describes two different laws that are fundamental for memory. The first, the law of ecphory, explains that an interaction between a retrieval cue and the stored fragments of information is necessary for the retrieval of memory. This retrieval of memory enhances or creates new memory traces (engrams). The second law, the law of engraphy (memory storage), explains that every time information is stored in memory, some retrieval is involved (Schacter et al., 1978). Anderson, Bjork and Bjork (1994) refer to this phenomenon as the retrieval-based learning assumption – that is, the process of retrieval enhances the likelihood of recall of that specific information at a later point in time. According to Anderson et al. (1994), the combination of the retrieval-based learning assumption, the competition assumption (memories linked with the same cue compete when retrieval processes are engaged) and the strength-dependence assumption (the strength of memory traces increases when retrieval occurs and at the same time the strength decreases for other information linked to the same cue), explains how retrieval of information will strengthen the memory traces of that information, sometimes at the expense of other related information (Anderson et al., 1994).

### 1.2.3 Neural Structures Supporting Memory

Besides behavioural research examining memory systems and processes, there has also been extensive research into the neural structures supporting encoding and retrieval. This research spans lesion and neuroimaging studies in humans, rats and nonhuman primates (Aggleton, 2012; Eichenbaum et al., 2007; Yonelinas & Ritchey, 2015). The standard neural model of episodic memory identifies the medial temporal lobe (MTL) as a critical structure (Aggleton, 2014; Eichenbaum et al., 1994; Squire & Wixted, 2011; Yonelinas, 2002; Yonelinas et al., 2010). In the context of this thesis, the role of the hippocampus, a structure within the MTL, is potentially of particular importance.

The hippocampus plays a critical role in binding fragments of item and context information of an event into a stable memory trace (Eichenbaum et al., 1994). This is assumed to be the basis for recollection. Familiarity, by contrast, has been proposed to be supported by the perirhinal cortex (Yonelinas, 2002), supporting the processing of item information only. While the specific contributions of components of the temporal lobe to recollection and familiarity remain contested, the strong link between hippocampus and recollection is established.

For the formation of emotional memories, another important brain structure involved is the amygdala (Kensinger & Schacter, 2006; Phelps & Sharot, 2008; Ritchey et al., 2008). An influential proposal is that for emotional information an interaction between the amygdala and hippocampus stabilizes hippocampus-dependent binding resulting in stronger memory traces (Kensinger & Schacter, 2006; Yonelinas & Ritchey, 2015). This can be regarded as the neural basis of the EEM, and could explain why emotional memories might be more resistant to forgetting in comparison to neutral memories that are solely dependent on hippocampus binding (Yonelinas & Ritchey, 2015). The specific links between amygdala and hippocampus, and the importance of the latter for recollection, is another reason that the focus for many questions about emotion and memory focus on the process of recollection rather than familiarity. Moreover, as the hippocampus is involved in reactivating memory traces during retrieval, this would suggest that this region is also critical at the time of retrieval.

The measures used in this thesis do not permit any claims about the roles of specific brain systems in emotion, control and memory, none the less it is reasonable to assume that modulation of hippocampal activity, at either encoding or retrieval, likely in tandem with the amygdala will be central to how this is enacted neurally. The mid-frontal old/new effect may reflect activity in the perirhinal cortex, either indirectly or directly. Patients with hippocampal damage show decrements in the left-parietal ERP old/new effect, but the scalp distribution of this effect suggests generators in parietal cortex, and it has been proposed that the effect might act as an indirect index of hippocampal integrity (Düzel et al., 2001).

The reason that these assumptions about how retrieval works, and about how encoding and retrieval are linked, are relevant here is because they emphasise the ways in which remembering and forgetting are associated, which is a key question driving this thesis. Remembering and forgetting can be linked via relatively automatic processes – such as those involved in the competition assumption - as well as those that involve conscious attempts to exert control over our memories. One possibility is that memory traces are increased in memory strength for emotional memories and because of this are prioritised during retrieval under the competition assumption. This would lead to increased difficulty in forgetting emotional memories. It may also be the case, however, that enhanced and prioritised processing during encoding for emotional information (Bowen et al., 2018) compared to neutral information makes it more challenging to forget (and easier to remember). In the experiments in this thesis, the primary manipulations that might be considered to involve exerting control are engaged at the encoding stage. The effect on memory is assessed via how people complete recognition memory tasks and variants. Hence, in the following section a consideration of the processes engaged during recognition memory is provided.

## 1.3 Recognition Memory Models

Several models have been proposed to explain the processes of recognition memory and these will be discussed in more depth below. Although the focus on memory control in this thesis is primarily during encoding, what happens during retrieval may also be crucial in memory control for emotional material.

Recognition memory tests are a common method for measuring memory performance. In the standard form a recognition memory test involves viewing old items (presented during a previous encoding phase) and new items (not presented at encoding). Participants are instructed to respond whether they have seen the item in the preceding encoding phase or not (Snodgrass & Corwin, 1988).

There are various models based on different assumptions that attempt to account for processes that are responsible for recognition memory, and linked methods that have been used to measure these processes. Prominent accounts are discussed below.

### 1.3.1 The Signal Detection Theory (SDT) Account of Recognition Memory

The Signal Detection Theory (SDT) account of recognition memory assumes that memory test items lie on a continuum of memory strength. Old items vary in memory strength, as do new items, but new items have a lower memory strength mean compared to old items

(Snodgrass & Corwin, 1988). The assumption is that the greater mean memory strength for old items is a consequence of their presentation in a preceding study phase.

Sensitivity, in recognition memory tests, is usually measured by calculating the distance (difference) between the means of correct and incorrect old and new responses (memory accuracy) for studied and unstudied items, which is referred to as d-prime (*d*') (Dougal & Rotello, 2007; Macmillan & Creelman, 2005; Snodgrass & Corwin, 1988), and is calculated by subtracting the z-scores for false alarms (incorrect judgements to unstudied items) from the z-scores for hits (correct judgements to studied items):

$$d' = z_H - z_{FA}.$$

The SDT assumes that participants adopt a response criterion (c; a level of strength) that an item must exceed to attract an old judgement. When this criterion is not exceeded the item is judged to be new. According to the SDT, a participant's criterion can be measured by the following formula:

$$c = 0.5\ (z_H + z_{FA}).$$

The measure of criterion locates the criterion relative to the intersection of the old and new distributions (Macmillan & Creelman, 2005; Snodgrass & Corwin, 1988). Three descriptors of criterion are commonly employed: conservative, neutral and liberal criterion. A liberal criterion is a higher tendency to respond old, regardless of whether the item has been shown before or not, and a conservative criterion is a lower tendency to respond old (Snodgrass & Corwin, 1988).

The standard SDT model assumes that distributions of old and new items overlap (at least when performance is not at ceiling) and that the distributions have the same shape (see Figure 1.1A). However, this is not the case in recognition memory tasks. Empirically, it has been shown that old and new item distributions do not have equal variance, as the distribution for old items is commonly more variable (see Figure 1.1B; Yonelinas, 2001a; Yonelinas &

Parks, 2007; Wixted, 2007; for a review of such findings see Yonelinas, 1994). As a result, when the equal variance assumption is violated conclusions based on memory sensitivity alone can be inaccurate (Dougal & Rotello, 2007; Snodgrass & Corwin, 1988; Yonelinas, 1994). Under these circumstances, when response criterion changes across experimental conditions, *d'* and response criterion become confounded and *d'* cannot be meaningfully interpreted (Dougal & Rotello, 2007; Kapucu et al., 2008; Macmillan & Creelman, 1990). This is an important consideration for the experiments that are described in this thesis, and it will be returned to in subsequent chapters and discussions.

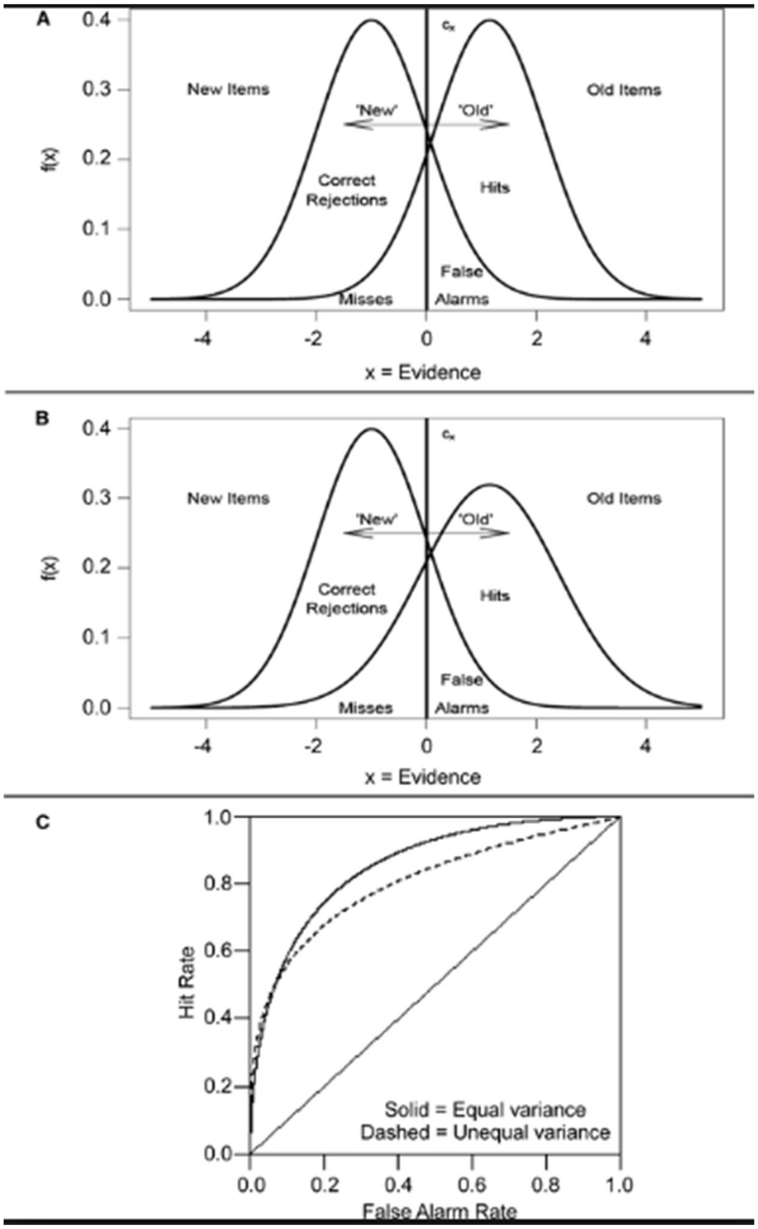



*Figure 1.1.* Examples of a distribution of equal variance (A) and unequal variance (B) for old and new responses in a recognition test and the resulting Receiver Operating Characteristic curves (C). f(x) = the frequency of evidence = x; cx = criterion for accepting items as "old". Taken from Pazzaglia, Dube and Rotello (2013).

### 1.3.2 Receiver Operating Characteristics (ROCs)

Receiver Operating Characteristic (ROC) curves are a set of measures where sensitivities at different response criteria can be inferred. In memory studies, the cumulative proportions of correctly recognized old items (hits) are plotted against the proportions of incorrectly recognized new items (false alarms) separated according to the level of confidence in the judgements (Yonelinas & Parks, 2007). Asking participants to indicate their confidence in their judgements is one way of plotting a ROC. Different levels of confidence can be regarded as different criteria for yes/no decisions. As already noted, ROCs are commonly plotted as a cumulative function, where the left-most point is the highest confidence hit against the highest confidence false alarm. The second point is the sum of these outcomes with the data for the next level of confidence, and so on.

Regarding the interpretation of ROCs, the curves represent levels of memory sensitivity (see Figure 1.1C). Curves that fall more towards the upper left in the ROC space indicate superior sensitivity (Dougal & Rotello, 2007; Yonelinas & Parks, 2007). The points that are located on a curve represent a range of hit and false alarm pairs for a level of memory sensitivity as a function of response criterion. The closer the points are to (1, 1) the more it reflects a liberal criterion and a decrease in confidence (see Figure 1.2).

The use of ROCs has been influential in investigating the processes that underlie recognition memory, and whether single process or dual process models are the best fit to explain recognition memory (Parks & Yonelinas, 2007; Stretch & Wixted, 1998; Wixted, 2007; Yonelinas, 1994, 2002; Yonelinas et al., 1996; Yonelinas & Parks, 2007). The reason for this is in large part because of how the shapes of ROC curves should vary according to different accounts.

According to the SDT model (that assumes equal variances for old and new item distributions), a typical ROC has a characteristic concave shape (see Figure 1.1C). The SDT model also predicts ROCs will be linear when plotted in z-scores, with a slope of 1.0 (Snodgrass & Corwin, 1988; Wixted, 2007). However, the typical shape of ROCs in recognition memory tasks is asymmetrical and usually appears in the upper left of the ROC space (i.e. it does not drop below the chance diagonal; see Figure 1.3) (Yonelinas, 1994). This is because old and new distributions, in general, do not have equal variances, hence the slope of zROCs is commonly less than 1.0 (Wixted, 2007). There are different process models that can explain this ROC shape.

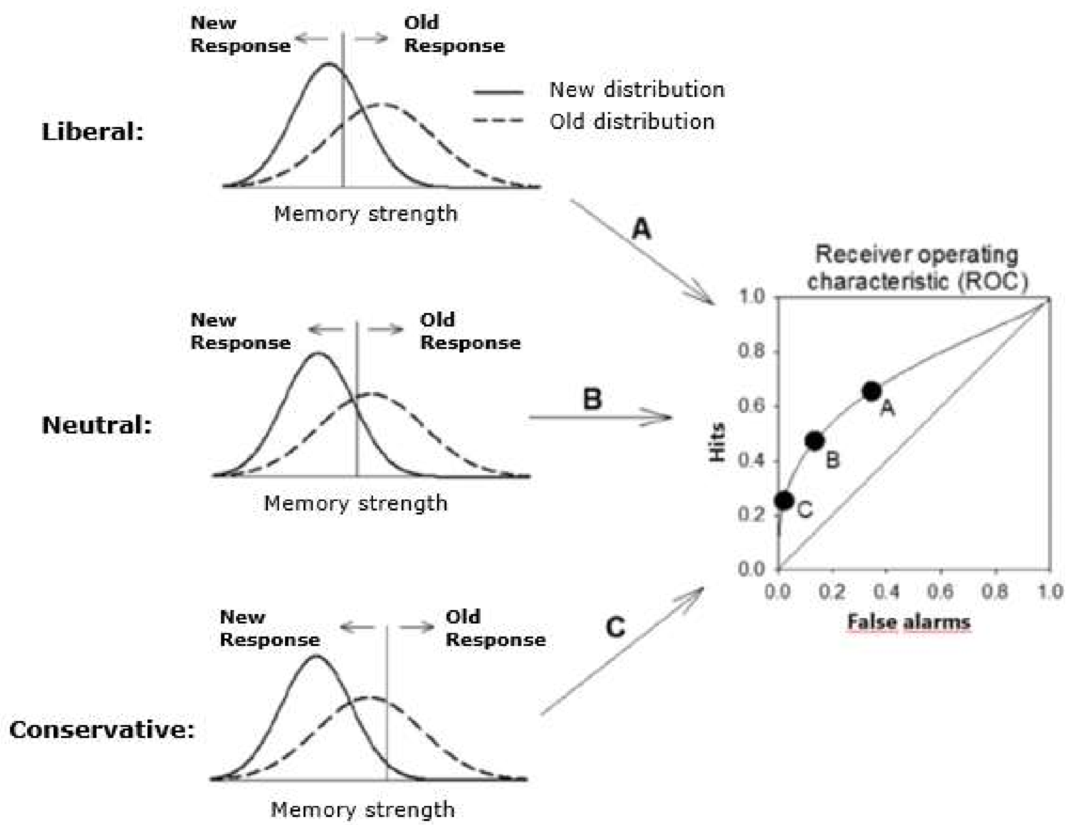



*Figure 1.2.* An illustration of three different levels of response criterion in a standard Unequal Variance Signal Detection (UVSD) model. On the right the ROC curve illustrates the ROC points that correspond with the response criterions, point A is a liberal criterion, point B a neutral and point C a conservative criterion. Taken from Wixted and Mickes (2014).

One such model is the Unequal Variance Signal Detection (UVSD) model (Snodgrass & Corwin, 1988; Wixted, 2007; Yonelinas & Parks, 2007). The UVSD model recognizes that the variability of the old item distribution can be greater than the variability of the new item distribution. This may arise because of variability during encoding (Yonelinas & Parks, 2007), thus memory strength may vary between studied items more than between unstudied items (Wixted, 2007). As a result of these assumptions the model can produce asymmetrical ROCs (Yonelinas & Parks, 2007), and produce linear zROCs with a slope less than 1.0 (Wixted, 2007).

### 1.3.3 The Dual Process Signal Detection (DPSD) Account of Recognition Memory.

The Dual Process Signal Detection (DPSD) model also assumes, similar to the UVSD model, that the variance for old and new item distribution may vary. It differs, however, by assuming that recognition memory is based on two distinctive processes. One of these processes, in common with standard SDT models, is an index of the strength of an item and is commonly referred to as familiarity (sometimes implemented as *d*') (Heathcote, Raymond, & Dunn, 2006; Rugg & Curran, 2007; Yonelinas, 1994, 2002). Familiarity has been described as a feeling of knowing without the presence of contextual details (Voss et al., 2012), but perhaps the critical component is that it operates as a graded strength signal (Yonelinas, 1994). The second process that is involved in recognition memory is memory judgements that are made on the basis of retrieval of contextual (qualitative) information, which is referred to as recollection (R) (Mandler, 1980; Rugg & Curran, 2007; Voss et al., 2012; Yonelinas, 1994, 2002). Recollection has been formalised by some as a discrete process with an all-or-none outcome (i.e. a threshold process), meaning that the processes of recollection can either succeed, when exceeding a threshold, or fail (Harlow & Donaldson, 2013; Murray, Howie, & Donaldson, 2015; Parks & Yonelinas, 2007). In dual-process accounts, familiarity is the basis for old judgements only if recollection has failed (Heathcote et al., 2006; Wixted, 2007; Yonelinas, 1994).

According to the DPSD model, recollection can increase high confidence hits without influencing the false alarms (the leftmost points on a ROC curve), which will result in an asymmetrical ROC (Yonelinas, 1994; Yonelinas & Parks, 2007). Changes in response criterion are, therefore, assumed to result in changes in familiarity based judgements (Yonelinas, 1994). The DPSD model predicts that ROC curves will increase in intercept and decrease in slope when only recollection differs across conditions (see Figure 1.3 (top ROCs)). When both recollection and familiarity differ across conditions, the DPSD predicts the slope of the ROC curves will remain constant while the intercept will increase (Yonelinas, 1994; see Figure 1.3 (bottom ROCs)). Furthermore, because, according to the DPSD, recollection is an all-or-none process, zROCs can produce a slight U-shape. This is commonly found when recognition relies heavily on recollection (Yonelinas & Parks, 2007), whereas when recognition relies heavily on familiarity zROCs are more linear (Parks & Yonelinas, 2007).

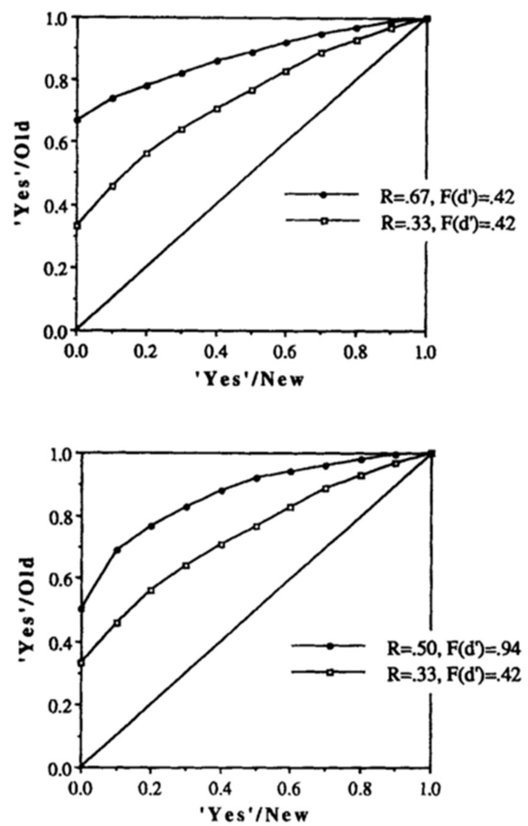



*Figure 1.3.* Examples of ROC curves generated by the Dual Process Signal Detection (DPSD) model taken from Yonelinas (1994). The top ROCs illustrate ROC curves when only recollection differs across conditions and familiarity remains static. The bottom ROCs illustrate when both recollection and familiarity change across conditions.

Both the UVSD and the DPSD models can account for the majority of the empirical evidence provided by data from ROCs in recognition memory tasks (Parks & Yonelinas, 2007; Wixted, 2007). Although the UVSD model describes only the memory strength variable as responsible for recognition memory, it has also been proposed that there may be two processes, recollection and familiarity, underlying this strength component (Rotello, Macmillan, & Reeder, 2004; Wixted, 2007; Wixted & Stretch, 2004). The UVSD model assumes that these two processes are combined whenever making a memory judgement to provide a single strength signal (Rotello et al., 2004; Wixted, 2007; Wixted & Stretch, 2004), whereas the

DPSD model describes recollection and familiarity as two memory processes with different properties that act independently (Yonelinas, 1994).

Wixted (2007) argued that even though the DPSD model fits well with many empirical findings, the UVSD model is a better fit based on direct comparisons of both models. For example, Glanzer, Hilford, Kim and Adams (1999), tested the curvilinear zROC assumption of the DPSD model and found that this assumption does not hold in old/new recognition memory tests, which has been confirmed by other studies as well (for a review of findings see Wixted, 2007). This observation has been linked with the argument that recollection is not a threshold process, which is an assumption of the DPSD model. When recollection is treated as a graded process, akin to familiarity, then the DPSD model provides a better fit to the data (Wixted, 2007). A demonstration of this is provided by Rotello, Macmillan, Hicks and Hautus (2006) who used a Remember/Know procedure combined with confidence judgements, allowing recollection judgements to be based on different levels of response confidence, which is in contrast with the assumption that recollection is only represented by high confidence judgements. They observed that this 'extended' DPSD model (i.e. allowing recollection to be a graded process) provided a better fit to the data than the standard DPSD model, even though the UVSD model still provided the best fit. According to Parks and Yonelinas (2007), however, the assumption that recollection is best described as a threshold process has been misinterpreted. They argue that some have interpreted this assumption as being that recollection is all-or-none, whereas they conceive of the process of recollection as being graded but the experience of recollection might occur only when an activation threshold is reached (Parks & Yonelinas, 2007; Yonelinas et al., 1996).

Parks and Yonelinas (2007) have argued that in order to explore which model is superior, it is necessary to examine data that relies heavily on one of two processes. This can be done using relational recognition tasks such as source or associative recognition tests, which rely mostly on recollection (even though familiarity can still contribute to judgements to a degree in some circumstance (Yonelinas, 1999, 2001b; Yonelinas et al., 1999)). If the data provides a more U shape zROC when memory relies heavily on recollection, this would provide evidence for the DPSD model. By contrast, if the zROCs remain linear this would provide evidence for the UVSD model. Several studies have provided evidence for U-shape zROCs in source and associative recognition memory tests (for a review of findings see Parks & Yonelinas, 2007), thereby supporting the predictions of the DPSD model. Yonelinas and Parks (2007) have argued that the DPSD equally explains the empirical data in item recognition relative to the UVSD model, and additionally provides a better fit for relational recognition memory, and so should be preferred on that basis. In light of these outcomes, in some of the experiments in this thesis estimates of recollection and familiarity based on the assumptions of the DPSD are employed to investigate the links between emotion and memory control.

### 1.3.4 The Remember/Know (R/K) Paradigm

During the Remember/Know (R/K) paradigm, participants, whenever an item is identified as 'old', are asked to identify whether they recognize the item based on conscious recollection of contextual information (via a 'Remember' response) or based on a feeling of having seen the item before in absence of any recollection of contextual information (via a 'Know' response) (Gardiner, 1988; Tulving, 1985b; Yonelinas & Parks, 2007). Contextual information commonly includes any recollected information about what happened or what was experienced at the time the item was studied (Yonelinas & Jacoby, 1995). Findings from early studies using the R/K paradigm support the idea that Remember (R) and Know (K) responses are based on distinctive processes, mapping on to recollection and familiarity respectively (Dunn, 2004; Gardiner, 1988; Rajaram, 1993). The most common outcome in experiments designed to assess how Remember and Know responses are sensitive to different manipulations is a change in Remember responses alongside no change in Know responses (Dunn, 2001, 2004; Yonelinas et al., 1998; Yonelinas & Jacoby, 1995).

Assumptions about the relationship between the processes of recollection and familiarity are key for answering the question of how recollection and familiarity can be estimated in the R/K paradigm. The earliest assumption underlying the R/K procedure was that recollection and familiarity are mutually exclusive (Gardiner, 1988; Rajaram, 1993; Tulving, 1985b; Yonelinas & Jacoby, 1995). Under this view the probability of recollection is the probability of a Remember response, while the probability of familiarity is the probability of a Know response: recollection and familiarity can never co-occur (Yonelinas & Jacoby, 1995). However, based on empirical data, Yonelinas and Jacoby (1995) argued against the exclusivity model as an explanation of the relationship between recollection and familiarity. They suggested that recollection and familiarity are not mutually exclusive, rather, the two processes are independent. The independence assumption means that although Remember responses should in theory provide a pure measure of recollection, Know responses will not provide a pure measure of familiarity. This is because, under independence, a proportion of items can be both recollected and familiar and, in this case, will receive a Remember response (if it is assumed that when both processes occur a Remember response will result). Consequently, the level of familiarity based responses is underestimated by the proportion of Know responses, because some items that elicited a Remember response would have attracted a Know response if recollection failed (Yonelinas & Jacoby, 1995). Therefore, according to the independence assumption, Remember responses can be taken as a measure of recollection (R) and familiarity (F) is measured by dividing the proportion of Know responses (K) by the opportunity the subject has to make a Know response (1 – R):

$$F = \frac{K}{1 - R}$$

Another assumption that might in principle explain the relationship between recollection and familiarity is the redundancy assumption, which is often best described in a source memory test (Yonelinas & Jacoby, 1995). In a source memory test, participants are instructed to indicate whether an item is old or new, a basic item recognition test, followed by whether they recognize the source of the recognized items. In such a task, participants must recognize the item in order to be able to recollect its source. Therefore, recollection (source memory test) is always accompanied with familiarity and responses based on recollection cannot increase beyond the responses based on familiarity (Yonelinas et al., 1996; Yonelinas & Jacoby, 1995).

Moreover, the redundancy assumption is consistent with the view that recognition memory is based on a single memory process, similar to the standard signal-detection models (Gillund & Shiffrin, 1984; Hintzman, 1986; Murdock, 1982; Yonelinas et al., 1996). As illustrated above, however, the empirical evidence in recognition memory tests does not support strongly a single process model. Moreover, the redundancy assumption does not account for the ROC shapes generally found in recognition memory tests (Yonelinas & Jacoby, 1995). The independence assumption, however, does seem to account for the general findings in recognition memory using the R/K paradigm and ROC curves. For this reason, the independence model is used in this thesis for estimating processes of recollection and familiarity in the R/K procedure.

### 1.3.5 Recognition Memory and Memory Control

Based on prior evidence supporting the DPSD model (Parks & Yonelinas, 2007; Yonelinas, 1994, 2002), recollection and familiarity have been measured based on this model throughout this thesis. As noted earlier in the introduction, it may be that memory control can only be exerted over recollection. An argument supporting this proposition from an emotion perspective is that retrieving contextual information is what specifically needs to be controlled, in particular to avoid re-experiencing negative memories. Breakdowns in this process might lead to pathological intrusions and disruptions to memory. These observations reinforce the general prediction tested in this thesis: that control over memory for emotional material will be evident on measures of recollection only.

## 1.4 Control of Memory

This section provides an overview of proposed mechanisms and frameworks that explain how memory control works. This is followed by a description of paradigms commonly used to investigate memory control.

It is widely held that the operation of the human memory system can only be explained via recourse to control processes that operate at multiple stages, including encoding and retrieval (Anderson & Hanslmayr, 2014; Moscovitch, 1992; Schacter et al., 1998). In order to recover past experiences appropriately, it is commonly assumed that memory depends on constructive processes during retrieval, which are prone to errors. For example, retrieval cues may match with information other than what needs to be retrieved, resulting in retrieval of irrelevant information that might be mistakenly assumed to be relevant (Schacter et al., 1998).

As already highlighted, for successful retrieval a cue-trace interaction is central (Schacter et al., 1978; Semon, 1904). Anderson et al. (1994) argued for three different mechanisms that can influence the cue-trace interaction: cue bias, target bias and attention bias. Cue bias involves processes that influence internal representations so that they are more likely to interact with some traces rather than others, or to fit with specific traces (Mecklinger, 2010). Target bias refers to processes that increase the accessibility of memory representations (Mecklinger, 2010). For example, the activation levels of some representations might increase or decrease, making them more or less accessible. A demonstration of target bias may also be found in the preparation processes in perception. When preparing to attend to colour for example, baseline levels change in neural activity in the V4 brain region (linked with the perception of colour) even before an object has been perceived (Driver & Frith, 2000). Meaning that some representation of colour has been increased making the perception of colour more accessible.

For both cue and target bias, the outcome – retrieval of specific information – would be the same. Attention bias refers to allocating attention towards the outcomes of specific relevant cue-target interactions. This attention allocation increases the likelihood of reporting relevant information. So, whilst a retrieval cue may activate several competing targets or memory traces, attention bias enables selection of task-relevant information amongst task-irrelevant information. Anderson et al. (1994) have acknowledged that some or all these biases may operate in different circumstances in parallel or separately. This framework provides a broad explanation as to how selective memory may come about, and, if this framework is accurate, then a remaining challenge is to consider which biases operate in which contexts. It may also be the case, however, that biases work at only some of these stages, and if this were to be the case then a linked question is whether the outcomes of the work in this thesis can speak to the accuracy of the framework. In this regard the opportunities afforded by the work in this thesis are limited. In ERP studies of memory and memory control it has been proposed that the time-course of processes during retrieval can be used to separate processes indexing retrieval and post-retrieval processing (Doidge et al., 2017). This offers a means of distinguishing attention bias (post-retrieval) from other forms of bias. These post-retrieval ERP indices, however, are not commonly observed in studies where only old/new judgments are required, as is the case in the ERP studies in this thesis. As already described above, the best evidence for the existence of target bias processes has also come from a combination of brain imaging and behavioural data.

For these reasons, the data to be presented here do not have much leverage for assessing the accuracy of this framework. The focus on the encoding/retrieval separation in Experiment 5, however, is relevant. In that case the question at issue is whether any processes operating at retrieval are influential in respect of the links between memory control and emotion. To anticipate, the findings in that experiment indicate the importance of processes engaged at the time of retrieval for the response criterion that accompanies memory judgements for different emotional materials. This outcome does not, however, align the behavioural outcomes in that experiment with any particular kind of bias.

As already alluded to in the preceding paragraph, one way of describing the bias framework is that it recognises the importance for memory retrieval of what processes precede and what processes follow cue-trace interactions. This separation between pre- and post-retrieval processing is also included in other process descriptions of memory retrieval, and in this context, one particularly influential proposal was provided by Tulving (1972). According to Tulving (1972), in order for retrieval to be successful it is necessary to enter a cognitive set, the purpose of which is to ensure that subsequent stimuli will be treated as retrieval cues. This cognitive set is referred to as retrieval mode, which is a tonically maintained state which ensures that processes are activated in service of retrieval when a retrieval cue is presented (Tulving, 1972; Wheeler, Stuss, & Tulving, 1997; Wilding & Nobre, 2001). Rugg and Wilding (2000) argued for a second cognitive set similar to retrieval mode that may be necessary for successful retrieval: retrieval orientation. While retrieval orientation is considered to be a tonically maintained state as well (Herron & Wilding, 2004), it can be described as a strategic source monitoring process (Wilding, 1999). The difference between these two cognitive sets is that, while retrieval mode will be engaged whenever there is a requirement for episodic retrieval, orientations will vary according to what kind of episodic demands are imposed: for example, on recall versus recognition tasks, or on tasks where different memory contents are the focus for retrieval judgments.

Adopting a retrieval set (which could be mode or orientation) is assumed to have positive consequences for memory judgments – these could be more accurate judgments, or they could be judgments made more quickly (Rugg & Wilding, 2000; Tulving, 1972). How does this characterisation of sets link to the proposals about bias discussed previously? Sets might in principle influence cue-, target- and/or attention bias, and for retrieval orientations this might vary in different circumstances. Preparing for episodic retrieval might influence the way cues are processed, as described by Tulving (1972), which would fit with the idea of cue bias (Mecklinger, 2010). It is also possible, however, that the function of retrieval sets is – either in whole or part – to enable selective attention to some of the products of retrieval rather than others (Dzulkifli et al., 2004), which is of course a re-description of attention bias. Similarly, increasing accessibility of specific (target) information may facilitate the initiation of a task specific retrieval orientation (Mecklinger, 2010), which illustrates the links between target bias and retrieval sets.

Another framework that considers the cue-trace interaction to be central to the likelihood of retrieval is the constructive memory model (Schacter et al., 1998). The model emphasises that for successful retrieval it is necessary for stored information to be integrated into a unified representation during encoding, and form a comprehensive memory representation (Schacter et al., 1998). When this integration of information fails, the retrieval of such information may only be partially successful in that some critical information may not be retrieved. Hence, in order for successful memory retrieval, what happens during encoding is important (Schacter et al., 1998). The model also assumes that because retrieval cues can activate memory traces that involve information other than the to-be-retrieved information and interfere with memory retrieval, it is necessary to select relevant information from among the activated memory traces. This process is referred to by Schacter et al. (1998) as focusing, a key requirement for the constructive memory model, and this set of processes has been described elsewhere as post-retrieval processing (Rugg & Wilding, 2000), and is of course also related to the earlier description of attention bias (Anderson et al., 1994). According to this model, irrelevant information may compete with relevant information and when the process of focusing fails, irrelevant information may come to mind (Schacter et al., 1998). Similarly, based on a study in which participants were asked to comment on their thoughts while simultaneously recovering autobiographical memories, Burgess and Shallice (1996) proposed a set of processes that are responsible for veridical memory retrieval. Editor processes are necessary to check whether the output from memory storage is relevant to the retrieval requirements, which is then controlled by mediator processes to ensure that the relevant information is being retrieved. These processes are comparable with the process of focusing and are in place to encourage retrieval of relevant information.

In summary, the previous descriptions of various forms of bias that can affect memory, of processes that might be engaged before retrieval occurs, and processes that operate downstream of memory retrieval, emphasise that control over memory can be exerted at multiple stages. In addition, it has also been emphasised that what happens at the time of encoding is important (see section 1.1.1). In the next section the focus moves to a description of paradigms that have been used to investigate the ability to control remembering and forgetting. They include paradigms where key manipulations are at either the retrieval phase, the encoding phase, or both.

The predominant focus in this thesis is on the Directed Forgetting (DF) paradigm, but other paradigms, namely the Retrieval-Practice Paradigm (examining Retrieval Induced Forgetting (RIF)) (Anderson et al., 1994), and the Think/No-Think (TNT) paradigm (Anderson

& Green, 2001), are also described here. Common to each of these is that they include measures of remembering and forgetting, and there is some overlap between the processes that are thought to be responsible for them (Anderson & Hanslmayr, 2014; Levy & Anderson, 2002). One of these common processes is inhibition, which can operate during encoding and/or retrieval (Anderson & Hanslmayr, 2014).

Control over memory can come about in different ways, and at different stages, for example by selecting from among retrieved information, or by stopping or inhibiting retrieval (Levy & Anderson, 2002). Selecting relevant information may be particularly necessary when information related to the to-be-retrieved information is also available. Inhibiting retrieval can achieve the same outcome, albeit via a different route by preventing retrieval of some contents from occurring.

During encoding, inhibition may disrupt the consolidation of information and subsequently prevent retention. When inhibition occurs during retrieval, by contrast, the information has already been encoded and stored, and inhibition processes might conceivably prevent retrieval cues from activating associated unwanted information, thereby decreasing their accessibility (Anderson & Hanslmayr, 2014; Levy & Anderson, 2002). Findings in the memory control paradigms discussed below have all been explained by appealing to inhibitory accounts that operate during either encoding (Directed Forgetting (DF)) or retrieval (DF, Retrieval-Practice and Think-No-Think (TNT)).

### 1.4.1 Retrieval-Practice Paradigm

Anderson et al. (1994) showed that retrieving certain items from memory decreased the likelihood of other, similar items being retrieved subsequently. The Retrieval-Practice paradigm, in which this was first characterised, consisted of a study phase, a retrieval-practice phase and a test phase. In the study phase participants were first instructed to study a set of category-exemplar pairs (e.g. fruit-orange, drinks-rum, fruit-banana). The study phase consisted of eight categories with six exemplars in each. Participants then engaged in retrieval practice on half of the exemplars from half of the categories. Each category was presented with the stem of an exemplar (e.g. fruit-or__), and participants were instructed to retrieve the exemplar that was shown with this in the study phase. After a distraction interval (20 minutes), a surprise cued-recall test followed. All the categories were presented, and participants were instructed to recall any exemplars they remembered from the study phase. The common finding in this paradigm is that retrieval practice improves recall for the practiced exemplars (orange) compared to the unpractised exemplars (banana) of the practiced categories and unpractised categories (drinks-rum). More important, however, is the finding that memory for the exemplars (rum) from the unpractised categories was improved compared to memory for the unpractised exemplars (banana) from practiced categories. Anderson et al.'s (1994) interpretation of these findings was that the increase of memory strength for certain information through retrieval causes a decrease in memory strength for competing information with a common retrieval cue – other items in the same category. This is referred to as retrieval-induced forgetting (RIF) (Anderson et al., 1994; Levy & Anderson, 2002).

This interpretation rests on three assumptions; the competition assumption, the strength-dependence assumption and the retrieval-based learning assumption (Anderson et al., 1994). The competition assumption holds that memories associated to a common cue compete for access to conscious recall when that cue is presented. The strength-dependence assumption holds that the cued recall of an item will decrease as a function of increases in the strengths of its competitors' associations to the cue. Finally, the retrieval-based learning assumption holds that the act of retrieval is a learning event in the sense that it enhances subsequent recall of the retrieved item (Anderson et al., 1994).

An alternative interpretation, however, is that inhibition during retrieval, as a resolution for the competition between active memory traces, is responsible for RIF (Anderson et al., 1994; Anderson & Spellman, 1995). Anderson et al. (1994) proposed that when participants are presented with a cue, all associated targets are activated. During the retrieval-practice phase, however, participants may inhibit targets that are not cued with the exemplar stem, causing a decrease in relative memory strength and subsequent recall of these items during a later memory test. This inhibition of unpracticed exemplars may be sufficient to cause RIF (Anderson et al., 1994; Anderson & Green, 2001; Dehli & Brennen, 2008).

In a set of experiments, Anderson and Spellman (1995) investigated the role of inhibition in RIF. They presented participants with category-exemplar pairs in which some exemplars (e.g. strawberry) could be paired with different categories (e.g. food and red). They argued that if a non-inhibitory account is responsible for RIF, memory for non-practiced exemplars (e.g. strawberry) from non-practiced categories (e.g. food) should not be impaired. However, if inhibition does play a role in RIF, memory for unpracticed, similar exemplars from an unpracticed category should be comparable to the memory performance for unpracticed exemplars in the practiced category, which is exactly what they found (Anderson & Spellman, 1995). In another study, Hellerstedt and Johansson (2014) presented participants with high and low category-exemplar associative strengths. They found that RIF was greater for exemplars with a high associative strength rather than a low associative strength. This finding is consistent with the view that competitive activation is necessary for the inhibitory account of RIF, because when the level of competition increases (high associative strength), so does the level of RIF. Presumably, this is a result of inhibition (Anderson et al., 1994; Hellerstedt & Johansson, 2014).

### 1.4.2 The Think/No-Think Paradigm (TNT)

Participants in the modal TNT paradigm study cue-target pairs, after which they are presented with some cues in the TNT phase. They are instructed either to retrieve targets (Think items), or to not retrieve targets (No-Think items). After the TNT phase, participants are commonly tested on all the studied items using a cued-recall test (Anderson & Green, 2001; Anderson et al., 2016). The key finding in this paradigm is that memory for no-Think items is impaired compared to baseline items (cue-target pairs shown only in the study phase). And, as expected, memory for Think items is increased compared to the no-Think and baseline items (Anderson et al., 2016; Anderson & Green, 2001; Lambert, Good, & Kirk, 2010; Murray, Anderson, & Kensinger, 2015; Vito & Fenske, 2017). One interpretation offered for this outcome is that being cued to not retrieve certain items (i.e. no-Think items) decreases the accessibility of this information, and one mechanism that could be responsible for this is inhibition (Anderson & Hanslmayr, 2014; Levy & Anderson, 2002).

Examining the inhibition account in the TNT paradigm, Anderson and Green (2001) instructed participants to study a set of cue-target pairs. In the TNT phase, the cue-target pairs were repeated 1, 8 or 16 times. They observed an impairment of No-Think items relative to baseline items on a subsequent recall task, which increased in magnitude along with the number of trials on which they were instructed to not be retrieved. Anderson and Green (2001) argued that this outcome is consistent with an inhibition account, because if inhibition is responsible for the decrease in memory for no-Think items relative to Think and baseline items, memory should decline even more when increasing the number of no-Think trials. This is what they observed. By this view, the suppressed information is progressively less accessible due to a weakening of the cue-target interaction (Levy & Anderson, 2002).

An alternative interpretation of these findings, however, is that, instead of inhibition, participants use a strategy in which they generate diversionary thoughts (i.e. thought substitution), creating new associates between the cue and substituted thoughts (Anderson & Green, 2001). These thoughts may interfere during later recall of the targets, thereby causing a decrease in memory for no-Think items. However, this strategy cannot explain a decrease in recall for no-Think targets even when participants were presented with a novel cue in a cued recall test (Anderson & Green, 2001). They argued that if inhibition impairs the memory for no-Think items itself, then this should be the case even when the no-Think items are cued with a novel cue during a cued recall test. In a second experiment, Anderson and Green (2001) using a cued recall test, cued participants with novel cues, semantically related to target items. Memory was decreased for no-Think items compared to Think-items and baseline items, consistent with an inhibition account.

Using event-related potentials (ERPs), Bergström, de Fockert and Richardson-Klavehn (2009) assessed the processes responsible for the memory outcomes in the TNT. In the TNT phase, participants were either instructed to suppress (i.e. not retrieve) or to substitute thoughts in the no-Think trials. ERPs were recorded during the TNT phase. Behaviorally, in both groups (suppression and thought substitution condition) memory was decreased for no-Think items relative to Think and baseline items. Electrophysiologically, they focused on a parietal positivity that is considered to be an index of recollection (Rugg, 1995a; Rugg & Curran, 2007). This effect was smaller in the no-think (i.e. suppression) than in the thought substitution condition. They interpreted this outcome as evidence that inhibition is involved in the suppression of memory retrieval, as previous research has shown this ERP effect to be reduced when attempting to stop retrieval (Bergström et al., 2009). Because, behaviorally in both groups, memory for no-Think items was decreased relative to Think and baseline items, their data are consistent with the view that there are two distinct mechanisms that can lead to forgetting in the TNT paradigm.

### 1.4.3 The Directed Forgetting Paradigm

Although several paradigms have been used to examine memory control, such as the Retrieval-Practice and TNT paradigms, the Directed Forgetting (DF) paradigm has been selected in this thesis to examine the links between emotion and memory control for several reasons. First, the DF paradigm is applied to investigate intentional forgetting. Linked with the motivation of this thesis is the practical importance of the ability to control memory for emotional challenging memories, which is presumably impaired for people suffering from PTSD or depression (Brewin, 2018; Hertel & Gerstle, 2003). In light of this consideration, the Retrieval-Practice paradigm would not be an appropriate method to use as this paradigm focuses on unintentional forgetting. Finally, another option would be to use the TNT paradigm. However, there have been extensive discussions over the reliability of the TNT effect (Bulevich et al., 2006). There is considerable variability in findings: some studies do not report TNT effects, suggesting that forgetting in the TNT paradigm is not robust. This has been reported without the manipulation of emotion (Bulevich et al., 2006). In order to examine the links between emotion and memory control it is critical to have some effect of forgetting as a baseline. In the absence of such a baseline, it would not be possible to make inferences about the effect of emotion on memory control. For these reasons, the choice has been made to use the DF paradigm to examine memory control.

There are two commonly used variations of the Directed Forgetting (DF) paradigm: the item-method and the list-method (MacLeod, 1975, 1999; Muther, 1965). In the item-method, participants are presented with items and are instructed to either forget or remember them on a trial-by-trial basis. The specific point in time at which participants are alerted to the requirement to remember or forget an item varies across studies. For example, some present the instruction simultaneously with the item (e.g. a word presented in colour which represents the instruction; Bailey & Chapman, 2012), but more often the instruction is presented after the item (MacLeod, 1999; McNally et al., 1998). Following the study phase, participants complete a test of memory for items they were directed to remember or forget.

In the list-method, participants are given two lists of items to study and most commonly are instructed to forget the first list of items after memorizing the list and to only memorize the second list. As for the item-method, this is followed by a test of memory for items they are directed to remember or forget. The common finding in the DF paradigm is superior memory (most often assessed by recognition or recall) for items followed by a remember cue (TBR) than items followed by a forget cue (TBF) (MacLeod, 1975; MacLeod, 1999; Chiu et al., 2010), which is referred to as the DF effect.

In the experiments described in this thesis the item-method is employed. Although emotion and memory control have been investigated using both methods, the majority of DF studies have used the item-method. One important goal of this thesis was to try and understand at least some of the variability in the literature. Hence, in order to do so, comparing results with previous studies is essential. Taken together, using the item-method would be a sensible way to start, and this was done consistently across all experiments in this thesis. A potential downfall of restricting the DF manipulation is that only one set of processes are investigated in these studies. The assumption is that a DF effect in both methods is a result from a different set of processes. In the item-method, processes during encoding are proposed to be responsible for a DF effect and in the list-method this has been linked with processes operating during retrieval. Of course, it is interesting to understand both sets of mechanisms as emotion may influence memory control at different stages. Particularly having in mind the challenges that people experience in certain pathological developments such as in PTSD. Therefore, investigating the list-method using a similar approach as taken in this thesis would be a sensible next step. However, implementing the item-method here may hopefully result in understanding the variability in findings, therefore the following description of processes supporting DF is restricted to that method.

#### 1.4.3.1 Processes Responsible for Directed Forgetting in the Item-Method

Several explanations have been offered for the DF effect. Bjork (1972) argued that it is mainly caused by the selective rehearsal of TBR items, similarly MacLeod (1975) argued that the encoding of TBR items is enhanced compared to TBF items. An alternative account is that the differences in memory strengths between TBR and TBF items are due to reduced motivation by participants to retrieve TBF items and not because of differential processing at encoding (MacLeod, 1999). This latter possibility was examined by using a reward system in the item-method (Macleod, 1999). Participants were instructed to try and remember all TBR and TBF items in an initial recall test. To determine whether there was a motivational factor in play, participants were offered a reward for every additional TBF item that was remembered in a second recall task. A directed forgetting effect remained evident in the second recall test, with no additional TBF words that were remembered relative to the first recall test. This suggests that differential processing of TBR and TBF items does contribute to the directed forgetting effect and motivational factors do not play a role (MacLeod, 1999).

MacLeod (1975) considered two different accounts for the DF effect in the item-method. The selective rehearsal account (Basden et al., 1993; Bjork, 1972), according to which all items are initially processed, but processing is discontinued when a forget instruction is presented. This results in TBR items being rehearsed and encoded more extensively than TBF items. The other account, the selective search account, proposes that TBR items are given a higher priority when searching for items (MacLeod, 1975). The former seems the most likely, since DF effects occur both in recall and recognition tasks in the item-method (MacLeod, 1975). The reason for this observation is the fact that there is not typically considered to be much opportunity for, or indeed engagement with, selective search in recognition memory tests (Basden et al., 1993). Moreover, the selective search account assumes that all items have already been encoded and stored before receipt of DF instructions, which is not the case in the item-method (MacLeod, 1975). If it is the case that the selective search account is applicable, a DF effect would be absent under a recognition memory test.

Arguably, if the selective rehearsal account is correct then it should be more effortful to remember than to forget, because TBF items should not be attended to or encoded as extensively as TBR items. However, in studies examining this, slower reaction times during forgetting were observed compared to remembering (e.g. Fawcett & Taylor, 2008), suggesting that forgetting is more effortful than remembering (Anderson & Hanslmayr, 2014). This outcome suggests that different processes are applied to TBR and TBF items, and this possibility is supported by evidence provided in brain imaging studies (Anderson & Hanslmayr, 2014). In these studies, greater activity in the right superior and middle frontal gyrus and right inferior parietal lobe was observed for TBF items (that were subsequently forgotten) compared with TBR items that were subsequently forgotten (e.g. Nowicka, Marchewka, Jednoróg, Tacikowski, & Brechmann, 2011; Rizio & Dennis, 2013). Taken together, these findings suggest that processes during intentional forgetting are different from processes during incidental forgetting. Anderson and Hanslmayr (2014) concluded that a combination of processes operate upon TBR and TBF items during encoding, and that selective rehearsal and active inhibition, respectively, are responsible for the item-method DF effect. The evidence for inhibition being the process involved is indirect, in so far as it relies on the brain regions engaged in response to a TBF cue, but it is none the less an interesting proposal.

## 1.5 Emotion and Memory

The main research question that drives the empirical work in this thesis is the links between emotion, memory and memory control; more specifically, how does emotion influence the ability to exert control over what we remember and what we forget? A common finding is that memory for emotional material is enhanced compared to neutral material (Buchanan, 2007; Hamann, 2001; Reisberg & Heuer, 1992; for reviews see Kensinger & Kark, 2018; Levine & Edelstein, 2009). Memory enhancement for emotional material, it has been argued, may function to direct attention to threatening stimuli to facilitate rapid and appropriate responses (Bowen, Spaniol, Patel, & Voss, 2016; Dolcos & Cabeza, 2002; Yang, Lei, & Anderson, 2016). Linking this with the earlier descriptions of memory processes in section 1.4, one possibility is that attention bias works in favor of emotional information which as a result leads to enhanced memory. It may be that this is not a sufficient explanation in and of itself, however, and pre-retrieval processes (that is, those that precede cue-trace interactions) are in play as well. This provides a mechanistic link to occasions when emotional memories disrupt daily functioning and become intrusive or traumatic (Barnier et al., 2004; Eftekhari et al., 2009). This can be conceived as a failure to exert control over emotional memories. Therefore, the ability to forget certain memories is crucial and may help prevent to avoid such intrusive or traumatic memories. To reiterate, links between emotion and memory are well-documented, although not thoroughly understood, and this is at least equally true for the links between emotion and control over memory.

Emotional material can be recovered and reported more frequently than neutral material (Kensinger & Corkin, 2003). This has been referred to as the emotion-enhanced memory effect (EEM) (Bradley, Greenwald, Petry, & Lang, 1992; Cahill & McGaugh, 1995; Kensinger & Corkin, 2003; Ochsner, 2000; Talmi, Anderson, Riggs, Caplan, & Moscovitch, 2008; Talmi, Schimmack, Paterson, & Moscovitch, 2007; for reviews see Levine & Edelstein, 2009), and might be due, in whole or in part, to better encoding and consolidation of emotional material. One explanation for the EEM is that emotional information reminds participants more often of personal experiences compared to neutral information. By associating this information with personal experiences, the consequence is enhanced memory for emotional material (Kensinger, & Corkin, 2004).

### 1.5.1 Effect of Emotion on Encoding and Consolidation Processes

Consolidation is a post-encoding process by which memories become more permanent and more resistant to loss (McGaugh, 2000). According to Bradley et al. (1992), there may be increased post-encoding elaboration for emotional material during consolidation. By this account, when arousal levels (a component of emotional material) increase, the likelihood that this information is rehearsed and elaborated upon increases as well. Moreover, because consolidation is considered to be a process that requires time, the enhanced effect of emotion on the consolidation process is thought to increase with time, making these memories more memorable than neutral ones (Hamann, 2001). However, often an immediate effect of emotion on memory is also observed in studies (e.g. Dougal & Rotello, 2007; Talmi et al., 2008). Hence, processes during consolidation alone might not be responsible for EEM. Rather a combination of processes during encoding and consolidation may offer a more accurate account.

The attention mediation hypothesis (Hamann, 2001) states that the enhanced allocation of attention towards emotional materials at the time of encoding underlies the EEM (Dolcos & Cabeza, 2002). Linking this to earlier observations in this thesis as well as the literature, Semon (1904) suggested that information is processed in fragments and the likelihood of retrieval depends upon which fragments we attend to (Schacter et al., 1978; Semon, 1904). The links between fragments of emotional material may be enhanced compared to neutral material, which offers a mechanistic explanation for why emotional materials are better remembered (Bowen et al., 2016; Dolcos & Cabeza, 2002), or more resistant to forgetting (Bailey & Chapman, 2012; Lambert et al., 2010). This enhanced allocation of attention towards emotional material has been argued to be modulated by increased levels of arousal (Bradley et al., 1992), based on a fairly common assumption that the arousal level of emotional information has an effect on memory (a detailed account on the effects of arousal on memory is described below). Emotional material has also been considered to be prioritized by processes that operate upon the information during encoding, at the expense of neutral material (Bowen et al., 2018). Taking these together, emotional material benefits from enhanced processing compared to neutral material, which consequently results in better memory.

### 1.5.2 Effect of Emotion on Retrieval Processes

The discussion in section 1.2.2 of this thesis, however, also emphasizes that processes operating at the time of retrieval can be important with respect to what is retrieved from memory. The influence of emotion on retrieval processes may occur in at least two different ways; either at the level of the item or at the level of the state of the person who is attempting to remember.

At the level of the item, some kinds of memory advantage for emotional content can be explained by the encoding specificity principle (Tulving & Thomson, 1973) or the principle of transfer-appropriate processing (Morris et al., 1977). In both cases, what happens during encoding is critical for whether retrieval cues are successful subsequently for retrieving specific information (Levine & Edelstein, 2009; Maratos & Rugg, 2001). In addition, it has been proposed that retrieval cues themselves can elicit the affective states that are comparable with those experienced during encoding, which may then facilitate retrieval of emotional material (Buchanan, 2007). It might also be the case, however, that the processing of emotional material at retrieval results in the engagement of additional processes, such as greater allocation of attention or prioritizing of resources in working memory such as rehearsing, which influences retrieval (Levine & Edelstein, 2009). In combination, these considerations offer several process-level possibilities for how encoding specificity or transfer-appropriate processing might be recruited to explain the EEM.

At the level of the participant, mood state might also contribute to what is more likely to be retrieved, and more specifically to what kinds of emotional information are retrieved. Key concepts here are mood-dependent memory and mood-congruent memory (Bower & Mayer, 1989; Eich, 1995). Mood-dependent memory refers to memory advantages when the mood-state is similar during retrieval and encoding (e.g. retrieving information that has been encoded in a positive mood, during the experience of a positive mood). Bower et al. (1978) used hypnosis to induce a happy or a sad mood. Participants were instructed to learn a list of words in one of these mood states, after which they learned a second list either under the same mood or the opposite mood state. During a final recall test, participants were asked to recall words from both lists while being in a happy or sad mood. They observed better memory for words that were encoded and retrieved under the same mood state (Bower et al., 1978).

Mood-congruent memory refers to memory advantages when the emotional component of information is similar to the mood state during retrieval (e.g. retrieving positive information in a positive mood) (Bower & Mayer, 1989; Buchanan, 2007). For example, Teasdale and Fogarty (1979) instructed participants to study a list of positive, negative and neutral words, and during the test phase they manipulated the mood state. Positive mood facilitated the recall of positive words, and negative mood the recall of negative words.

In the experiments in this thesis, the primary focus is on item-level influences on emotion, remembering and forgetting. Notable exceptions are the additional individual difference measures acquired in Experiment 1, and the manipulation of emotion via context in Experiment 5.

### 1.5.3 Response Criterion in Emotional Memory

Despite the potential influence of response criterion on memory performance, response criterion has often been regarded as no more than a nuisance variable, which when determining memory performance should be corrected for. However, developments in the emotional memory literature have indicated this not to be the case (Windmann et al., 2002). A common finding in studies of memory and emotion when recognition memory is used as the test measure is a relatively more liberal criterion for emotional material compared to neutral material (Bailey & Chapman, 2012; Dougal & Rotello, 2007; Kapucu et al., 2008; Ochsner, 2000). As a result, in some instances it has been argued that apparent differences in memory for emotional material are in fact solely due to changes in response criterion (Kapucu et al., 2008; Windmann & Kutas, 2001). When response criterion changes across conditions, differences in memory are difficult to interpret (Dougal & Rotello, 2007). The reason for this difficulty has already been explained in section 1.3.1, and to anticipate, there will be a strong focus on changes in criterion and changes in sensitivity (memory performance) in this thesis, because many inconsistencies among published findings might be related to the links between these measures.

One explanation for differences in criterion with emotion is that emotional materials are likely more prone to be associated with personal experiences, which can enhance the ‘sense of familiarity’ (Kensinger & Corkin, 2003). This increased sense of familiarity, relative to neutral material, would result in a more lenient response criterion (Dougal & Rotello, 2007; Kapucu et al., 2008; Kensinger & Kark, 2018). By manipulating the level of categorical similarities within lists between negative and neutral words, White et al. (2015) examined the effects of categorical similarities on response criterion. The categorical level of words is the degree to which words share similar features (White et al., 2015). A liberal response criterion for emotional words was only observed when lists contained a high number of words that fell into similar categories (e.g. death, pain, failure). They proposed that a liberal response criterion for emotional material may be a consequence of categorical effects within an emotional category: when a large number of words that share features are studied, old words as well as new words that appear during a test phase that share these features are more likely to be judged as having seen before (White et al., 2015). If this is the case, the categorical level of words within a list is what influences response criterion rather than the emotionality of a word, and in so far as these category influences promote a sense of familiarity or 'gist' (Koutstaal & Schacter, 1997; Schacter et al., 1998), they align changes in criterion with changes in familiarity.

### 1.5.4 Valence and Arousal Effects on Memory

Valence and arousal are two dimensions along which emotions can be considered (Russell, 1980). Valence is the pleasure or displeasure property of an item (Kensinger & Corkin, 2003), and arousal is the state of being physiologically alert, awake and attentive, and varies from calm to excitement (Dolcos & Denkova, 2008; Thayer, 1978). It has been argued that only arousal plays a role in enhancing memory for emotional material (Buchanan et al., 2006; Ochsner, 2000). However, differences along the valence dimension have been shown to elicit memory enhancement, beyond the effects of arousal (Kensinger & Kark, 2018). Opinion is divided: while some have argued that EEM is a result of increased levels of arousal alone (Hamann, 2001; Mather, 2007), others emphasize that EEM may also be explained by effects of valence (Dolcos et al., 2004; Dolcos & Cabeza, 2002; Kensinger & Corkin, 2003, 2004).

In a study that examined whether arousal and valence predicted recognition memory, Adelman and Estes (2013) reanalyzed data from Cortese, Khanna and Hacker (2010). Memory was equivalent for positive and negative words and superior for both compared to neutral words. Assuming that arousal was equivalent for the emotional words, this suggests that differences in memory for emotional materials are due to arousal, because valence had no influence. Differences in memory, however, have also been observed between negative and positive material (Kensinger & Kark, 2018). For example, details of negative memories tend to be better remembered than details of positive memories (Kensinger & Schacter, 2006), and positive emotions are often associated with a feeling of familiarity, whereas recollection has been shown to be enhanced for negative emotions (Ochsner, 2000).

Despite the evidence for differential effects of negative and positive emotions on memory, the majority of studies investigating the effect of emotion on memory (and memory control) have focused on differences between negative and neutral material (Brandt et al., 2013; Dolcos & Cabeza, 2002; Kensinger & Corkin, 2003; Xie et al., 2018; Yang et al., 2016; Yonelinas & Ritchey, 2015). This may be explained by the general finding of enhanced memory for negative material relative to positive and neutral material (Minnema & Knowlton, 2008; Ochsner, 2000; Otani et al., 2012). From a clinical perspective as well, examining the effects of negative emotions on memory may further the understanding of emotional disorders, such as PTSD (Engen & Anderson, 2018; McNally et al., 2004; McNally et al., 1998).

There are two dominant explanations for why negative material is often associated with better memory than positive material (Ochsner, 2000). First, it may be that attentional biases are more pronounced for negative than positive information for the purpose of attending to threatening stimuli (Kensinger & Corkin, 2004; Yang et al., 2016), which may be survival-relevant information (Bradley et al., 1992; Ochsner, 2000). Perhaps for the same reason, negative information may also benefit from prioritization or facilitated processing (Kensinger, & Corkin, 2004), leading to superior memory. This has been referred to as natural selective attention (Dolan et al., 1999). Using a Stroop task, Pratto and John (1991) examined differences in attentional bias between negative and positive personality traits. They observed longer reaction times for negative traits, which they interpreted as negative valence receiving more attention relative to positive valence.

A second explanation is that negative material elicits more rumination and extensive elaboration (Engen & Anderson, 2018; Ochsner, 2000). Xie et al. (2018) examined the ability to remember and forget emotional material in participants with depressive tendencies, who experienced an increased difficulty to forget negative material. This provides some support for the rumination argument, because rumination is linked with depression (Garnefski & Kraaij, 2007; Joormann & Gotlib, 2010; Nørby, 2018).

Turning to arousal, this physiological factor is considered to elicit an autonomic response and typically attracts attention (Dolan et al., 1999; Kensinger & Kark, 2018), which can result in enhanced encoding and consolidation (Madan, Shafer, et al., 2017; Maratos et al., 2000). Examining the links between arousal and attention in EEM, Talmi et al. (2007) compared memory under a full attention condition and divided attention condition. In the full attention condition, participants commenced with an auditory discrimination task, followed by an encoding phase in which participants were presented with negative, positive and neutral images. After this, they were tested on their memory using a recall test. In the divided attention condition, the auditory discrimination task and the encoding phase of the images occurred simultaneously. During this phase participants were instructed to pay attention to the auditory discrimination task. When attention was divided, performance on the discrimination task was worse when participants viewed emotional images. Based on these findings, Talmi et al. (2007) concluded that arousal modifies attentional levels. In addition, mediator analyses indicated that attention mediated the effect of arousal on memory. In other words, superior memory is a result of enhanced allocation of attention due to increased levels of arousal.

The level of arousal tends to be higher in emotional material compared to neutral material (Buchanan et al., 2006; Kensinger & Corkin, 2003). Although other stimulus properties may influence the memorability of emotional material as well, arousal is considered to be one of the main factors explaining enhanced memory for emotional material (Buchanan et al., 2006; Cahill & McGaugh, 1995; Ochsner, 2000; Phelps & Sharot, 2008; Sharot et al., 2004).

### 1.5.5 Item Properties of Emotional and Neutral Material

The preceding section contained a detailed account of how arousal and valence might play a role in memory for emotional materials. There are also several other factors to consider when investigating factors that might influence memorability. Word properties and the properties of word sets are important considerations (Buchanan et al., 2006; Glanzer & Bowles, 1976; MacLeod & Kampe, 1996; Madan et al., 2017; Maratos et al., 2000; Rugg & Doyle, 1992; Talmi & Moscovitch, 2004).

In the experiments described in this thesis, four word properties are controlled for: word length, word frequency, semantic relatedness, and arousal. Word length is the measure of the number of letters a word consists of. Word frequency is the measure of the probability of occurrence of a word (Madan et al., 2012) and semantic relatedness is the level of inter-item associations within a set of words (Buchanan et al., 2006).

#### 1.5.5.1 Word Length

The word length effect is the finding that short words are easier to remember than long words (Baddeley, Thomson, & Buchanan, 1975). This is most commonly observed when assessed via short-term memory tasks (Baddeley et al., 1975), and in long-term recall tasks (Russo & Grammatopoulou, 2003; Tehan & Tolan, 2007). However, the reverse of this word length effect has also been reported in lists where short and long words were intermixed (Katkov et al., 2014). This reversed word length effect has also been reported in recognition tasks (Tehan & Tolan, 2007). Hendry and Tehan (2005) argued that this reversal in recognition tasks is the result of long words taking longer to be processed and being more elaboratively processed in such a way that subsequent memory is better for long relative to short words. In studies of memory and emotion, word length is an important property to control for because emotional words tend to be longer relative to neutral words (Madan et al., 2017).

#### 1.5.5.2 Word Frequency

The same reversal applies for word frequency. In recall tests, high frequency (HF) words are better recalled, whereas in recognition memory tests, memory is enhanced for low frequency (LF) words (Glanzer & Bowles, 1976; MacLeod & Kampe, 1996; Rugg & Doyle, 1992). There are several accounts that attempt to explain this effect for LF words. One is that because people encounter LF words less than HF, they are relatively less familiar and receive more attention (Mandler, 1980). Another is that the degree of automaticity of processing a word decreases for LF words (MacLeod & Kampe, 1996), which ultimately leads to increased recognition of these words relative to HF words (Jacoby & Dallas, 1981). A common assumption is that encoding processes for LF words are more extensive and elaborate relative to HF words, resulting in increased memory strength for LF words compared to HF words (Glanzer & Bowles, 1976; Kinsbourne & George, 1974; MacLeod & Kampe, 1996; Rugg & Doyle, 1992). Because emotional words occur less frequently relative to neutral words (Madan et al., 2017) it is important to remove this confound when investigating the influence of emotion on memory.

#### 1.5.5.3 Semantic Relatedness

According to Tulving and Pearlstone (1966), because semantically related words are more accessible in memory, search parameters during retrieval are reduced. Hence, the degree of semantic relatedness between items within a word set can influence memorability. The degree of semantic relatedness for emotional words tends to be elevated relative to the semantic relatedness of neutral words (Buchanan et al., 2006; Kensinger & Kark, 2018), which in turn may facilitate memory for emotional words (Kensinger & Kark, 2018; Maratos et al., 2000; Talmi & Moscovitch, 2004). To test the effect of relatedness on memory, Talmi and Moscovitch (2004) tested memory for emotional, random neutral and categorized neutral words. The categorized neutral words were words that were increased in the degree of relatedness (e.g. category kitchen). The degree of relatedness was equated between the emotional (negative) and categorized words and differed from the random neutral words. Memory for emotional words was enhanced relative to random neutral words, however this memory benefit was absent between emotional and categorized neutral words. Thus, equating the level of relatedness eliminated the memory benefits for emotional words. Talmi and Moscovitch (2004) concluded that enhanced memory for emotional words is linked to semantic relatedness. Dougal and Rotello (2007), conducted two experiments using the Remember/Know procedure, looking at memory for emotional and neutral words. The first experiment showed enhanced memory for negative words compared to positive and neutral words when semantic relatedness was not controlled for. In the second experiment, where semantic relatedness was controlled for between emotional and neutral words, no enhanced memory for emotional words was observed. Similar to other word properties, these patterns of data suggest that, when the degree of semantic relatedness is not controlled between emotional and neutral words, it can contribute to any differential memory outcomes that are observed.

### 1.5.6 EEM in Memory Control

Besides the effects of emotion on memory, the effects of emotion on memory control have also been examined. Emotional material is considered to be more resistant to forgetting compared to neutral material (for a review see Kensinger & Kark, 2018). It remains unclear, however, why this is the case (Yonelinas & Ritchey, 2015). One way of explaining this is perhaps a combination of enhanced encoding and consolidation processes that increase memory strength for emotional material and consequently make it more difficult to forget. Memory control paradigms (including DF, RIF and TNT) have been used to investigate whether, and if so how, memory for emotional materials can be controlled. Findings for the links between emotion and memory control in RIF (Barnier et al., 2004; Dehli & Brennen, 2008; Harris et al., 2010; Hauer & Wessel, 2006; Moulds & Kandris, 2006) and TNT (Chen et al., 2012; Lambert et al., 2010) are mixed. In RIF for example, Dehli and Brennen (2008) used emotional and neutral category-exemplar pairs and instructed participants to make old and new judgments in a recognition memory test after a retrieval-practice phase. They observed only a RIF effect for neutral items, indicating that emotional words were resistant to retrieval-induced forgetting (for similar findings see Blix & Brennen, 2012; Moulds & Kandris, 2006). Other researchers have found inconsistent differences in RIF between positive and negative valence stimuli. While some have found a RIF effect for negative material compared to positive material (Harris et al., 2010), others have found the reverse (Hauer & Wessel, 2006). There are also reports of no differences in the RIF effect between emotional and neutral material (Barnier et al., 2004). These mixed findings in the effects of emotion on RIF have been explained by the use of different types of measurements (Barber & Mather, 2012). For example, Blix and Brennen (2012) highlighted that different memory tests have been used in these studies which might explain the inconsistencies in the effect of emotion on RIF (recognition memory: Blix

& Brennen, 2012; Dehli & Brennen, 2008; recall memory: Harris et al., 2010; Hauer & Wessel, 2006).

In the TNT literature, the general findings are that emotional words are more resistant to attempts to forget ('not thinking') than are neutral words (Chen et al., 2012; Lambert et al., 2010). Chen et al. (2012) used face-picture pairs as cue-target pairs in which the faces had neutral expressions and images were manipulated in valence (negative and neutral valence). After the TNT phase, participants were presented with faces from the study phase and were instructed to recall the images that were previously paired with these faces in a cued-recall test. They observed that negative images were resistant to forgetting, suggesting that participants were less successful in 'not thinking' about the negative images compared to neutral images (Chen et al., 2012).

#### 1.5.6.1 EEM in Directed Forgetting

Examining the effect of emotion on DF has been done using the two widely used DF methods; the item-method and the list-method. The materials used in these studies have included words and/or images (Bailey & Chapman, 2012; Brandt et al., 2013; Gallant, Pun, & Yang, 2018; Gallant & Yang, 2014; Hauswald, Schulz, Iordanov, & Kissler, 2011; Liu, Chen, & Cheng, 2017; Marchewka et al., 2016; Minnema & Knowlton, 2008; Nowicka et al., 2011; Otani et al., 2012; Payne & Corrigan, 2007; Wessel & Merckelbach, 2006; Xie et al., 2018; Yang et al., 2016).

Table 1.1 provides a summary of outcomes in critical studies. The table shows there is variability in findings: there have been studies observing similar DF effects for emotional material relative to neutral material, which indicates that, at least in those cases, emotion did not affect memory control (Gallant & Yang, 2014; Wessel & Merckelbach, 2006). Other studies, however, have found a reduced DF effect for emotional material relative to neutral ones, suggesting that emotional material is more resistant to intentional forgetting (Bailey & Chapman, 2012; Hauswald et al., 2011; Liu et al., 2017; Minnema & Knowlton, 2008; Nowicka et al., 2011; Otani et al., 2012; Payne & Corrigan, 2007; Yang et al., 2016). These studies differ in the type of stimuli (e.g. words or images) that were used and to what extent the stimuli used had controls for critical factors (e.g. word frequency). The differences in findings are unlikely to be due completely to the factors controlled for in these studies, as inconsistent findings have been observed in studies where similar control has been exerted. In addition to providing an overview of the inconsistent findings in DF studies investigating the links between emotion and memory control, the table gives a full description of which types of stimuli were used and which factors were controlled for.

Table 1.1

*An Overview of Previous Findings on the Effects of Emotion on Directed Forgetting, Memory Sensitivity and Response Criterion*

| Study | Stimulus | Valence | Stimulus control | Findings | | |
|---|---|---|---|---|---|---|
| | | | | DF effects | Memory sensitivity | Response criterion |
| Wessel & Merckelbach, (2006) | Words | Negative and neutral | | Similar DF effects for negative and neutral words | Equal memory for negative and neutral words | |
| Payne & Corrigan, (2007) | Images | Negative, positive and neutral | Arousal (negative and positive) | Smaller DF for emotional relative to neutral images | Superior memory for emotional relative to neutral images | |
| Minnema & Knowlton, (2008) | Words | Negative, positive and neutral | Arousal (negative and positive), length, frequency, syllables and semantic relatedness | Smaller DF for negative relative to positive and neutral words | Superior memory for negative relative to positive and neutral words | |
| Nowicka et al., (2011) | Images | Negative (fear & disgust) and neutral | | Smaller DF for negative relative to neutral images | Superior memory for negative relative to neutral images | |
| Bailey & Chapman, (2012) | Words | Negative, positive and neutral | Arousal (negative and positive) | Smaller DF for emotional relative to neutral words | Superior memory for neutral relative to emotional words | Liberal criterion for emotional words |
| Hauswald et al.,( 2011) | Images | Negative and neutral | | Smaller DF for negative relative to neutral images | Equal memory for negative and neutral images | Liberal criterion for negative images |
| Otani et al., (2012) | Images | Negative, positive and neutral | Arousal (negative and positive) | Smaller DF for negative relative to positive and neutral images | Superior memory for negative relative to positive and neutral images | |
| Yang et al., (2012) | Images | Negative and neutral | Arousal | Similar DF effects for negative and neutral images | Equal memory for negative and neutral images | Liberal criterion for negative images |
| Brandt et al., (2013) | Words | Negative and neutral | Frequency | Greater DF for negative relative to neutral words | Superior memory for negative relative to neutral words | |
| Gallant & Yang, (2014) | Words | Negative, positive and neutral | Arousal, length and frequency | Similar DF effects for emotional and neutral words | Superior memory for emotional relative to neutral words | |
| Yang et al., (2016) | Words | Negative and neutral | Frequency and familiarity | Smaller DF for negative relative to neutral words | Superior memory for negative relative to neutral words | |
| Marchewka et al., (2016) | Images | Negative (fear, disgust & sadness) and neutral | | Similar DF effects for negative and neutral images | Superior memory for negative relative to neutral images | Liberal criterion for negative images |
| Li, Wang, & Han, (2017) | Phrases | Negative, positive and neutral | Arousal (negative and positive) | Smaller DF effects for negative relative to positive and neutral phrases | Superior memory for emotional relative to neutral phrases | |

| | | | | | | |
|---|---|---|---|---|---|---|
| Berger, Crossman, & Brandt, (2018) | Words | Negative, positive and neutral | Arousal (negative and positive) and word frequency | Similar DF effects for emotional and neutral words | Superior memory for neutral relative to emotional words | Liberal criterion for emotional words |
| Gallant et al., (2018) | Words | Negative, positive and neutral | Arousal, length and frequency | Similar DF effects for emotional and neutral words | Equal memory for emotional and neutral words | Liberal criterion for emotional words |
| Taylor, Cutmore, & Pries, (2018) | Images | Negative, positive and neutral | | Similar DF effects for emotional and neutral images | Superior memory for emotional relative to neutral images | |

The degree of semantic relatedness between words in a set influences the memorability of the words (Buchanan et al., 2006). It may also be the case that when relatedness is matched between emotional and neutral words, this manipulation eliminates effects of emotion on DF. However, Minnema and Knowlton (2008) conducted a list-method DF study using negative, positive and neutral words and observed the opposite. They used free association norms for each word to control for any words that were strongly related within a list. The association norms measure the likelihood that a word generates other words, which is presumably increased when words are elevated in the degree of relatedness (Nelson et al., 2004). Words that had some degree of association (i.e. words that generated existing words in a list) were placed in different lists. In addition, words were matched for word frequency, length, and number of syllables, as well as for arousal between negative and positive words. Recall for negative words was superior compared to positive and neutral words, and there was a diminished DF effect for negative words. In a second experiment, to examine the effects of arousal on DF, arousal levels were elevated for half of the participants. Again, they found a reduced DF effect for negative words relative to positive and neutral words. Minnema and Knowlton (2008) argued that the diminished DF effect for negative words is caused by a higher arousal level of negative words relative to neutral words. However, this does not explain why negative words were better remembered compared to positive words since arousal levels were controlled for between negative and positive words.

The most common approach in the item-method DF paradigm, is to present the instruction cues (remember/forget) after each stimulus. For emotional material, the period between stimulus and cue presentation may result in extensive encoding relative to neutral material, which perhaps explains the diminished DF effects for emotional material (Bailey & Chapman, 2012). To test this possibility, Bailey and Chapman (2012) reduced the encoding opportunity by presenting the stimuli (words) and instructions simultaneously. The colour of the words served as the remember or forget cue. They used negative, positive and neutral words and controlled only for arousal levels between negative and positive words. Participants were later tested using recall and recognition tests. A diminished DF effect was observed for both negative and positive words relative to neutral words in both tests. However, they found memory sensitivity to be superior for neutral words compared to emotional words. Thus, the extended encoding of emotional material due to a gap between stimulus and cue presentation does not explain a diminished DF effect for emotional material, although they did not directly compare both approaches in their experiment.

Other item-method studies in which the remember and forget cues were presented after the stimulus, using emotional and neutral images, have revealed similar DF findings when using a recognition test (e.g. Nowicka, Jednoróg, Wypych, & Marchewka, 2009), and when using a recall test (e.g. Otani et al., 2012). They differ, however, in finding better memory sensitivity for emotional relative to neutral material. One reason for these discrepancies might be that participants in Bailey and Chapman's (2012) study were more focused on the instruction because these were presented simultaneously with the study material and diminished the emotional effect on memory, which was not the case in the other studies (Nowicka et al., 2009; Otani et al., 2012).

In contrast to the above findings for DF, in some other studies there has been no effect of emotion on DF effects. Using the list-method, Wessel and Merckelbach (2006) presented participants with either two negative word lists or two neutral word lists. Which word properties were controlled for was not reported. After the study phase participants were tested using both a recall and a recognition memory test. They observed equivalent memory and DF effects for both negative and neutral words in the recall test. In the recognition test, there was no DF effect, which replicates findings in other comparisons of the item- and list-method where no DF effect was visible when applying a recognition test in the list-method  (Basden et al., 1993; MacLeod, 1999). Moreover, memory did not vary according to emotion. Similarly, Gallant and Yang (2014), using the item-method, observed no effects of emotion on the DF effect but did observe better memory for emotional words after a recognition memory test. These findings were replicated by Gallant et al. (2018), who observed no effects of emotion on DF. Also, they did not find any differences in memory between emotional and neutral words. In both studies the authors controlled the three valence categories (negative, positive and neutral) for word length, frequency, and arousal. Gallant et al. (2018) suggested that because the words were matched for arousal, no enhanced memory was observed for emotional words, and emotion did not have an effect on the DF effect, which has also been argued by Minnema and Knowlton (2008). However, this still may not be enough to explain the inconsistencies in the DF literature for emotional material. Brandt et al. (2013), using negative and neutral words that were matched on frequency but differed on arousal levels, found a larger DF effect for negative relative to neutral words.

Wessel and Merckelbach (2006) argued that the inconsistent findings described above are a product of timing and stimulus presentation differences across DF paradigms. They argued that studies observing memory benefits and diminished DF effects for emotional material have used a minimal time to present words (e.g. 180ms), thereby offering little time for elaborative encoding of stimuli and creating an advantage for emotional over neutral items. However, findings from various DF studies suggest otherwise. The presentation time of the stimulus varies within studies with diminished DF effects for emotional material (ranging from 500-5000ms) and in studies without these effects (ranging from 1000-3000ms). This set of outcomes suggests that variations in presentation time are not a good predictor of whether an effect of emotion on DF will be observed. Furthermore, Wessel and Merckelbach (2006) proposed that the method of instruction presentation may also affect the encoding processes of the item and therefore affect memory sensitivity between the valence categories. For example, presenting the remember/forget instruction simultaneously with the stimuli, as in Bailey and Chapman's (2012) study, may require different cognitive operations to those associated with DF when the cue is delayed. In the case of simultaneous presentation, participants may choose to simply ignore items they were instructed to forget instead of making an attempt to forget these items after they have been encoded. According to Brandt et al., (2013) a temporal separation between the study material and the remember/forget instruction is critical to investigate processes responsible for DF effects. However, the observation of a DF effect and diminished DF effect for emotional words in Bailey and Chapman's (2012) study suggests otherwise.

Adding further complexity, there has not been consistency in terms of stimulus control throughout these studies, as well as reporting omissions in some cases (Liu et al., 2017; Marchewka et al., 2016; Nowicka et al., 2011; Wessel & Merckelbach, 2006). In most studies there has been control over word length (Gallant & Yang, 2014; Minnema & Knowlton, 2008; Myers et al., 1998), word frequency (Brandt et al., 2013; Gallant & Yang, 2014; McNally et al., 1998; Minnema & Knowlton, 2008; Moulds & Bryant, 2002; Myers et al., 1998; Yang et al., 2016), and arousal between negative and positive material (Bailey & Chapman, 2012; Minnema & Knowlton, 2008; Otani et al., 2012; Payne & Corrigan, 2007). After controlling for all of these item properties, Minnema and Knowlton (2008) also measured the association norms within lists to reduce the likelihood of words within a list leading to the generation of words existing in the same list. However, most of these studies mentioned above did not control for these item properties altogether. This makes it challenging to detect an overall pattern which might explain the inconsistencies in the DF literature for emotional material.

The inconsistencies in the existing literature in terms of outcomes, properties that are controlled for, and how criterion is accounted for, have guided the set of experiments in this thesis. At issue are questions about the links between emotion and memory control, differential effects of positive and negative valence, and the links between changes in sensitivity (accuracy) and changes in criterion. The variability in published findings makes it challenging to make strong predictions, but based on the mechanisms and theories discussed in earlier sections some cautious predictions can be made. Although the work in this thesis is not designed to provide an answer to the question 'What processes are responsible for memory control?', it is expected that memory control will be more difficult for emotional compared to neutral material. This will be manifest in lower rates of forgetting for emotional as compared to neutral contents. Second, because recollection is considered to be a controlled process while familiarity is not, then in the experiments measuring recollection and familiarity, a directed forgetting effect will be observed only for recollection.

With respect to emotion and memory, when controlling for confounding factors, in this thesis it is expected to find no differences in memory sensitivity between emotional and neutral materials. These confounding factors relate to the control of stimulus properties, most notably semantic relatedness (Dougal & Rotello, 2007). This consistency of control is important, because a consistent set of findings across this series of studies would strongly argue that the variability in the literature is (at least partially) due to the different control over possible confounding factors in published work to date.

Key outcomes in all experiments reported here are behavioural measures, as already indicated. Moreover, as already alluded to, in some experiments electrophysiological measures are also employed. For this reason, the following section describes fundamental aspects of the use of real time measures of brain activity to inform questions about human memory.

## 1.6 Electroencephalography (EEG) Technology

### 1.6.1 Fundamentals of EEG

Berger (1924; as cited in Teplan, 2002) showed that electric currents generated in the brain can be recorded in real time via electrodes placed on the outside of the skull. This was the beginning of the use of electroencephalography (EEG) to study brain function. EEG now commonly involves the recording of electrical brain activity using electrodes that are placed on the scalp.

Electrical activity that is recordable at the surface of the skull reflects large populations of active neurons firing simultaneously (Picton et al., 2000; Teplan, 2002). However, not all electrical brain activity is detectable with EEG. Failures to detect activity occur when neurons do not fire synchronously and/or are not properly arranged for the activity that is produced to propagate to the scalp. What this means is that non-significant findings when contrasting brain activity in two conditions do not allow the claim that the same processing occurred in those conditions (Coles & Rugg, 1995; Rugg, 1995b).

One way to measure EEG is to assess activity in different frequency bands. For example, brain oscillations can be characterised into at least five bands: Gamma (>30 Hz), Beta (13-30 Hz), Alpha (8-13 Hz), Theta (4-8 Hz) and Delta (0.5-4 Hz) (Jackson & Bolger, 2014; Luck, 2014; Teplan, 2002). Activity in these bands varies according to detectable changes such as alertness, sleep status and the cognitive processes that are engaged while recording is taking place. Another way of measuring EEG is to assess activity in the time domain, plotting voltage changes over time. This is most commonly done by recording EEG time-locked to events of interest. Perhaps the most common form of time-locking is to stimuli that require a response from a participant. Other kinds of time-locking can, however, be employed, for example, response-locking (Luck, 2014). These event-locked elements of the EEG are referred to as event-related potentials (ERPs; Coles & Rugg, 1995; Rugg, 1995b).

The ERPs typically subjected to analysis are most commonly averages of numerous trials of the same condition. This is because the signal:noise ratio on individual trials makes analysis of them difficult (Luck, 2014). Averaging over numerous trials should retain the signal (assuming it is consistent across trials) while reducing the noise (assuming it is random across trials) (Luck, 2014). There are, however, several caveats here. First, if there is variability in the timing of modulations of interest across trials this will influence the averaged ERP that is obtained: with greater variability (often referred to as latency jitter) the averaged ERP will be smoothed out and the maximum amplitude will be reduced (Coles & Rugg, 1995; Luck, 2014; Picton et al., 2000). So, in principle differences in amplitude across conditions might reflect only differences in jitter. Furthermore, in some cases the assumption of consistency of signal across trials is acknowledged as not being met. For example, in experiments where forced choice decisions are made it is reasonable to assume that in at least some cases the neural activity on some trials will differ from that on others. For example, the neural activity underpinning a guess, or a decision based on low quality information, will in all likelihood differ from that underpinning a highly confident response.

### 1.6.2 Models of ERP Generation

The classical ‘evoked’ model, holds that ERPs are transient, fixed latency and fixed polarity responses to a stimulus and/or event (Hanslmayr et al., 2007). However, Hanslmayr et al. (2007) argue that the evoked model does not fully account for the generation of ERPs. They report evidence for a different ‘phase-reset’ model. This model holds that mechanisms underlying ERP generation can be partially explained via the reorganization of oscillations (Sayers et al., 1974). The phase-reset model assumes that oscillations undergo a reset that generates evoked components in response to a stimulus and/or event. To clarify, each single trial as a response to a stimulus of a certain frequency is reset. These trials are averaged, with the result bringing an ERP after the phase reset (Sauseng & Klimesch, 2008).

The accuracy of these competing accounts does not impact directly on the use of ERPs to make inferences about cognitive operations, in so far as the signal of interest is the averaged neural signature associated with different experiment conditions. In cognitive electrophysiology, the main purpose of employing ERPs is to use them to understand human information processing. The key to this approach is the assumption that using ERPs alongside behavioural measures can provide insights that are not available via behavioural measures alone, and/or can provide converging forms of evidence (Rugg, 1995b; Rugg, & Coles, 1995). ERPs can be used in this way to permit inferences about what processes were engaged, their time-courses and their properties.

### 1.6.3 Interpreting ERP Data

One way of characterising ERPs is by delineating the series of peaks and troughs, establishing the functional significance of each, and then observing how they vary according to conditions of interest (Donchin, 1966; Picton et al., 2000). For example, how does P3b vary with smoking status or arousal? Another approach is to identify ERP modulations via a contrast between experiment conditions of interest, for example, between items attracting old and new judgments on a recognition memory task. The underlying assumption here is that the contrast will allow a meaningful functional interpretation of any divergences between the associated ERPs, and via this approach the modulations that are identified might span one or more peaks and troughs as described above. For any given ERP signature of a cognitive process, the latency of the relevant ERP peak is used to infer when the cognitive process is maximally active. To infer how strongly a process is engaged, the amplitude is used (Luck, 2014; Kutas, & Dale, 1997).

### 1.6.4 Methods of ERP Acquisition

In order to collect EEG data, most commonly an elasticised cap containing electrodes is placed on the scalp (see Figure 1.4A). The number of electrodes required depends on the experimental design and pre-existing knowledge of where effects might be largest (Luck, 2014). Electrodes that are spaced sufficiently close together, increasing the number of electrodes, allows for precise measurement of the distribution of neural activity over the scalp. When electrodes are placed too far apart, one risk is that information is missed (Luck, 2014). A typical approach is to use an array that covers the scalp, and as one widely used visualisation tool is of scalp distribution, 64 or more electrodes should be used in order to ensure that these visualisations (often based on spherical spline interpolations) are accurate (Luck, 2014). When analyses, however, are targeted to specific modulations, or broadly distributed modulations, fewer than 64 electrodes could be sufficient.

Capture of the signal propagating to the scalp is commonly facilitated by applying a gel that contains conducting electrolytes and which sits between the electrodes and the scalp. Electrodes are commonly placed according to the international 10/20 system (see Figure 1.4B; Klem, Luders, Jasper, & Elger, 1958). There are other systems (Luck, 2014), but the 10/20 system is described here because it is used in the experiments described in this thesis. In this system, the electrodes are labelled by a letter and a number. The letter corresponds broadly to the cortex area underlying the electrode (e.g. F corresponds to the frontal cortex and T corresponds to the temporal cortex). Electrodes placed between two cortical regions are designated by two letters (e.g. FC refers to frontal-central). The number indicates on which side and how far from the midline the electrode is placed. Odd numbers correspond to the left side of the scalp and even numbers to the right side, low numbers being more medial locations than higher numbers. The letter z refers to the midline from the scalp which goes from the front to the back of the scalp (e.g. POz refers to the electrode in the parietal-occipital cortical region on the midline).

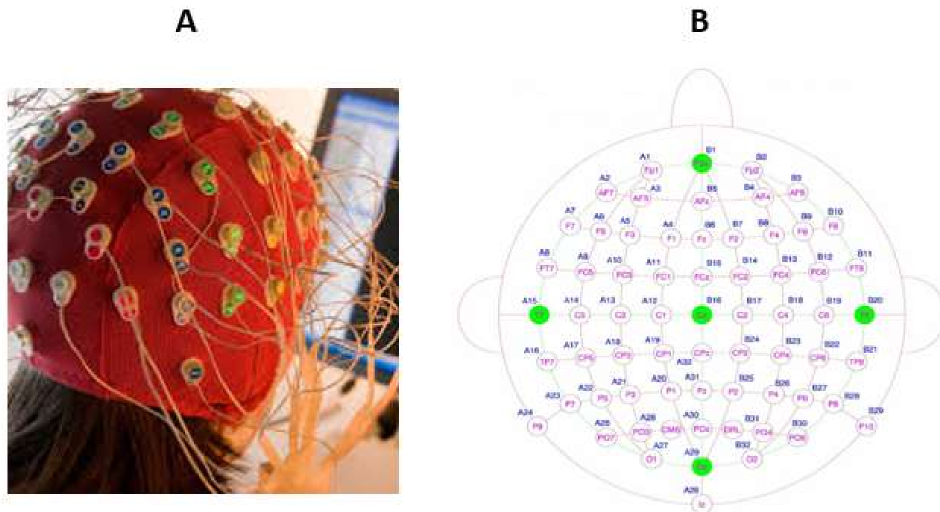



*Figure 1.4*. (A) An elasticised cap that contains the electrodes. (B) Electrode locations of the 10-20 system, as standardized by the American Electroencephalographic Society (Klem et al., 1958).

Alongside the active electrodes (those from which a signal of interest will be acquired), EEG recordings also require a ground electrode and a reference electrode (Luck, 2014). The ground electrode assists in reducing electrical noise from power lines and other sources. EEG is a relative measure and signals at active electrodes represent the difference between activity there and at the reference electrode. For example, a nose tip or vertex reference might be used. Other common references include the average of the signal at the two mastoids (as is the case in the ERP experiments in this thesis) or an average reference (Luck, 2014). The preference for using the two mastoids as reference in this thesis is because this is a method that is widely used in the memory literature, which facilitates comparison of results between different studies.

The perfect situation to record electrical brain activity would be when the recorded activity would only contain relevant brain activities that represent cognitive processes. Unfortunately, this is not the case and besides these cognitive processes, EEG recordings contain artefacts, generated by events such as muscle activities (including eye blinks) and sweating. In addition to physically (participant) related artefacts, the EEG will also pick up technical electrical noise (Luck, 2014; Teplan, 2002). In order to identify relevant and meaningful components in EEG data it is necessary to remove artefacts. This can be done during online acquisition of the EEG data and/or offline after data has been acquired. For physical artefacts, placing extra electrodes that record these physical artefacts is helpful in order to clean the data. For example, electrodes that record eye blinking will help in identifying non-relevant activity and to subsequently remove this artefact from the data. Eye movements can be recorded by using an electro-oculogram (EOG) with electrodes placed on the supra- and suborbital ridges of the eyes (Luck, 2014; Teplan, 2002). In addition, providing participants with specific instructions also helps to reduce the number of artefacts in the EEG data. For example, asking participants to reduce eye blinking or to limit eye blinking to a specific cue on the screen, although this can also introduce a dual-task requirement that might be of concern when testing specific groups, for example older participants, in certain kinds of paradigms.

### 1.6.5 ERPs in Recognition Memory

The most studied ERP correlates in recognition memory are 'old/new effects', which are differences between the neural activities elicited by old and new test items attracting correct judgements (Warren, 1980). The rationale for this contrast is that it may reveal processes that support memory judgements (Friedman & Johnson, 2000; Sanquist et al., 1980). Separable old/new effects for recollection and familiarity have been identified (Friedman & Johnson, 2000; Rugg & Curran, 2007; Vilberg & Rugg, 2008).

#### 1.6.5.1 Recollection - The Left-Parietal Old/New Effect

The ERP correlate that has been linked with recollection is the left-parietal old/new effect. It onsets around 400-500ms post-stimulus with a left-sided posterior maximum and lasts for 400-500ms in recognition memory tasks (Rugg & Curran, 2007; Yonelinas, 2002). In an early study by Sanquist et al. (1980) participants were instructed to judge whether two sets of words were similar or different based on three conditions: (i) orthographic (upper or lower case), (ii) phonemic (rhyming) and (iii) semantic (synonyms). For ERPs recorded in a subsequent recognition memory test they found increased positivity over parietal sites for judgements in the phonemic and semantic conditions. Furthermore, a greater positivity was found over parietal sites for words that were correctly recognized as old compared to words that were correctly recognized as new: a typical left-parietal old/new effect (Sanquist et al., 1980). This outcome links the parietal effect to recollection because it was larger (more positive-going) in the conditions associated with deeper processing (phonemic and semantic) that are likely to give rise to greater recollection (e.g. Rugg et al., 1998).

According to Rugg and Curran (2007), the strongest evidence that the left-parietal old/new effect indexes recollection is how the effect behaves in tasks requiring Remember/Know judgements or source judgements. The effect is larger (more positive-going) for Remember (reflecting recollection processes) than for Know judgements (Smith, 1993), and larger when source judgements are correct than when they are incorrect (Wilding et al., 1995; Wilding & Rugg, 1996). It is also larger when two rather than one source judgements are correct (Vilberg & Rugg, 2008; Wilding, 2000). For example, in a study by Wilding and Rugg (1996), participants were asked to judge whether words judged as old were previously presented in a male or female voice. The typical left-parietal old/new effect increased in magnitude for correct source (voice) judgements compared to incorrect source judgements.

#### 1.6.5.2 Familiarity - The Mid-Frontal Old/New Effect (FN400)

The mid-frontal old/new effect, also referred to as the FN400, has been linked with the process of familiarity. While there are some early studies where, on inspection, a FN400 might be evident (e.g. Düzel, Yonelinas, Mangun, Heinze, & Tulving, 1997; Sanquist et al., 1980; Smith, 1993; Wilding & Rugg, 1996), the first strong proposal was made by Rugg and colleagues (1998). They gave participants instructions to encode items deeply (sentence generation) or shallowly (alphabetic judgement). This encoding task was followed by a recognition memory test. In addition to a left-parietal old/new effect, they observed a second effect that was of the same magnitude for deeply and shallowly studied items. The effect comprised a greater relative positivity for old items than new items from around 300-500ms with a frontal scalp distribution. It was evident only for old items judged correctly to be old. While this effect was not the focus of their paper, Rugg et al. (1998) suggested that this was a candidate for an index of familiarity because behavioural data indicate that depth of processing does not impact markedly on the process of familiarity (for reviews and discussions, see Curran, 2000; Paller, Voss, & Boehm, 2007; Rugg et al., 1998).

In a study by Curran (2000), participants were presented with studied words, words that were similar to the studied words and new words in a recognition memory test. Besides a typical left-parietal old/new effect, there was a greater mid-frontal old/new effect for similar words relative to new words and there was no difference in the old/new effect between studied and similar words study. Studied and similar words should be more familiar than new words, thus these findings are consistent with a view of the mid-frontal old/new effect as an index of familiarity. These findings have also been replicated using images instead of words (Curran & Cleary, 2003). In addition, another study that provides evidence for the case of the mid-frontal old/new effect as an index of familiarity used a Remember/Know paradigm combined with confidence judgements (Woodruff et al., 2006). If familiarity is a graded process, then neural activity should covary with differences in confidence for Know judgements. Woodruff et al. (2006) found the mid-frontal old/new effect to be largest for highly confident responses relative to less confident responses, linking the mid-frontal old/new effect to familiarity strength (also see Curran, 2004).

Although these results are consistent with a familiarity account of the mid-frontal old/new effect (i.e. FN400), whether the FN400 is an index of familiarity is a matter of ongoing debate (Paller et al., 2007; Voss et al., 2012). It has been argued that the FN400 is not functionally distinct from the N400, which is apparent mostly over centro-parietal sites and is linked with conceptual priming (Bridger et al., 2012; Paller et al., 2007; Voss et al., 2012). Consequently, the FN400 has been interpreted as an index for conceptual priming, a form of implicit memory, rather than familiarity (Paller et al., 2007; Rugg & Curran, 2007; for a review of results see Voss et al., 2012). Bridger et al. (2012), however, demonstrated that the N400 and the FN400 differ functionally and topographically from each other. They achieved this by separating neural activity that could be elicited by conceptual priming (in a study phase) from activity that could be elicited by familiarity-based judgements (in a test phase). Participants were presented with primed and un-primed targets to which they were asked to make valence ratings. This allowed them to look into neural activity that was elicited by conceptual priming separately from any memory judgements. In a surprise recognition test, participants were presented only with un-primed targets and new words and instructed to make old/new judgements. Hence, by only presenting un-primed targets, if a FN400 was found this could be linked with familiarity and not with conceptual priming. They found a greater N400 for un-primed targets compared to primed targets in the study phase, which was distributed over centro-parietal sites. For the recognition phase, they found a typical mid-frontal old/new effect that peaked around 300-500ms and was distributed over frontal sites. Bridger et al. (2012) interpreted these findings as evidence for functional differences between the N400, indexing conceptual priming, and the FN400, indexing familiarity-based recognition memory. In addition, Mecklinger, Frings and Rosburg (2012) argued that the arguments made against the

FN400 as an index for familiarity, are inconsistent with both electrophysiology and behavioral data. For example, the FN400 varies where conceptual priming is held constant during a recognition test (e.g. Groh-Bordin, Zimmer, & Mecklinger, 2005). Furthermore, behavioural data indicating changes in conceptual priming have not always correlated with variations in the FN400 signal (Johansson et al., 2004).

#### 1.6.5.3 ERP Correlates of Emotional Material

Neural activity has also been found to vary for emotional material. In general, an enhanced early posterior negativity (EPN) within the time region of 200-300ms after stimulus onset has been found for emotional material compared to neutral material (Herbert et al., 2008; Schupp et al., 2003). This enhanced EPN has been proposed to reflect selective sensory encoding (Kok, 1997). Later in the processing of emotional material, around 400-600ms over parietal sites, an enhanced late positive potential (LPP) has been observed (Cuthbert et al., 2000). The same enhanced LPP is observed when participants are instructed to attend to and/or remember specific stimuli, however, with emotional material these processes may work automatically (Hauswald et al., 2011). In summary, both the EPN and LPP have been interpreted as indexing enhanced processing afforded emotional material, perhaps due to the importance of perceiving and evaluating information relevant to threat or well-being (Bailey & Chapman, 2012; Bradley et al., 1992; Ochsner, 2000).

These modulations are a helpful tool for assessing how emotion influences cognitive processing, and some of these changes may be relevant for memory encoding or retrieval. In and of themselves, however, changes in these modulations cannot be linked directly to memory processes.

In addition to the ways of exploring retrieval in the discussion of old/new effects, another means of assessing memory with ERPs is via subsequent memory effects (also called

Dm: Difference due to Memory) (Paller et al., 1987). These are obtained by contrasting neural activity during encoding for items that receive correct or incorrect judgements on a subsequent memory task. They are assumed to reflect processes that support successful encoding. While there are examples of the use of Dm effects to study encoding of emotional materials (for example, Yick, Buratto and Schaefer (2015) reported larger Dm effects for high rather than low arousal images), it is arguable whether new insights into emotion and memory have been achieved by this approach to date. These effects are returned to in more detail in section 3.3.

#### 1.6.5.4 ERPs and the Directed-Forgetting Paradigm

A common finding for ERPs in the Directed Forgetting paradigm is differences between neural activities for TBR and TBF items (Bailey & Chapman, 2012; Hauswald et al., 2011; Liu et al., 2017; Paller, 1990). In general, an enhanced positivity over parietal scalp regions during 300-500ms after the cue onset for TBR compared to TBF items has been reported. This enhanced positivity strongly resembles the P300 effect, which is thought to reflect attentional processes and has been linked with memory encoding (Kok, 1997). Moreover, this ERP effect resembles the LPP effect, which is linked with processes engaged when people are instructed to remember stimuli (Hauswald et al., 2011). This enhanced positivity is thought to indicate the initiation of encoding processes for TBR items, and in this sense the data are consistent with the selective rehearsal account (Hauswald et al., 2011; Liu et al., 2017). In addition, Hauswald et al. (2011) observed an enhanced frontal positivity from 500-700ms after cue onset for TBF items compared to TBR items, which they have linked with active inhibition (also see Liu et al., 2017). They argued that both the effective encoding processes of TBR items and the active inhibition of TBF items are responsible for the directed forgetting effect. These ERP findings correspond with the two accounts, selective rehearsal and active inhibition, that are proposed as being responsible for the directed forgetting effect based on behavioural findings (for a review of such findings see Basden et al., 1993), as discussed in a previous section (see section 1.3.3.1). The absence of direct evidence for the association of the latter modulation with inhibition, however, means that the basis for this inference is not well-established.

Although the main focus in this thesis is on the ERPs measured during retrieval, to complement behavioural insights in how memory control and recognition memory are linked, ERPs during encoding will also been examined. This provides the opportunity to link memory control and encoding. Based on the findings described above, it is predicted that there will be differential neural activity elicited by cues to remember or to forget. In keeping with prior literature, this would comprise of an enhanced LPP for TBR items, and an enhanced frontal positivity for TBF items (Bailey & Chapman, 2012; Hauswald et al., 2011; Liu et al., 2017; Paller, 1990).

Other insights into the directed forgetting effect have been provided by studies measuring ERPs during recognition memory test phases. Nowicka et al. (2009) found a typical old/new effect over central and parietal regions for correctly recognised TBR items, which was absent for correctly recognized TBF items. These data support the prediction that control operates over recollection, because this old/new effect has been linked with this process. In addition, ERPs were more negative-going for forgotten TBF items compared to correctly judged new items over central and parietal sites. The researchers interpreted the first of these effects as greater recollection for TBR than TBF items. The second, negative-going effect is intriguing, as it suggests that there are active encoding (possibly inhibition) processes engaged for TBF items (Nowicka et al., 2009).

# CHAPTER 2

# 2. EXAMINING EMOTION IN DIRECTED FORGETTING USING BEHAVIORAL MEASURES

## 2.1 Experiment 1 – Emotion and Directed Forgetting: a Baseline Experiment

As has been emphasised already, links between emotion and memory are well-documented, although not thoroughly understood, and this is at least equally true for the links between emotion and control over memory. In some DF studies there is evidence for the emotion-enhanced memory effect (Minnema & Knowlton, 2008; Otani et al., 2012) and an effect of emotion on memory control (Bailey & Chapman, 2012; Minnema & Knowlton, 2008; Nowicka et al., 2011; Otani et al., 2012; Yang et al., 2016), while in others there is no effect of emotion on memory control (Gallant & Yang, 2014; Wessel & Merckelbach, 2006). Several factors may play a role in these inconsistencies, including the control exerted over stimulus sets, as well as a failure to consider response criterion adequately.

Differences between semantic relatedness among emotional words compared to neutral words may partially account for the finding that memory is enhanced for emotional words (Dougal & Rotello, 2007; Talmi & Moscovitch, 2004). The degree of relatedness has also been used to explain why emotional words are easier to remember even under directed forgetting instructions (Maratos et al., 2000). Only a small number of relevant studies, however, have controlled for relatedness between emotional and neutral words. In the directed forgetting literature, to the best of my knowledge, there is one study in which semantic relatedness was controlled between negative, positive and neutral words (Minnema & Knowlton, 2008). The researchers observed enhanced memory and a diminished DF effect for negative words relative to positive and neutral words, suggesting that differences among properties other than semantic relatedness contribute to memory for emotional materials. A consistent approach in this thesis is to control for semantic relatedness across emotional and neutral words. Semantic relatedness is measured using the online source tool ‘snaut’ (Mandera et al., 2017). This tool works on prediction models in which a word is predicted given the context (the associated words) it appears in. The level of relatedness is measured by calculating the semantic distance between word pairs, using a matrix function (Mandera et al., 2017). The matrix function allows comparison of the semantic distance for a large set of words simultaneously and provides semantic distance scores for each word pair.

Besides semantic relatedness, considerations related to response criterion are also relevant to questions about how emotion, memory and memory control interact. The common finding in studies of recognition memory and emotion, as well as the linked DF literature, is a more liberal criterion for emotional material relative to neutral material (Bailey & Chapman, 2012; Hauswald et al., 2011; Marchewka et al., 2016). Some researchers have questioned whether differences ascribed to emotion are in fact due to differences in response criterion and not due to changes in memorability itself (Kapucu et al., 2008; see section 1.4.3).

Another factor that may play a role in memory control for emotional material is different coping styles (emotion regulation). Several frameworks, as discussed in the introduction (see section 1.3), attempt to explain memory control via different memory processes and mechanisms. However, another way to examine the links between emotion and memory control is by investigating the effect of emotional coping styles. It may be that the ability to regulate emotions influences the ability to forget or remember (Engen & Anderson, 2018; Hertel & McDaniel, 2010; Myers et al., 1998; Richards & Gross, 2000). For example, Richards and Gross (2000) conducted three experiments to investigate whether two emotion regulation strategies (expressive suppression and cognitive reappraisal) influence memory performance. Instructing participants which regulation strategy to apply, they were initially divided in two groups: a watch condition, were participants had to watch a short film clip and listen carefully, and an expressive suppression condition, where participants were instructed to suppress their feelings so that nobody watching them could see what they were feeling. In a second experiment a third condition was added - the reappraisal condition - where participants were instructed to have a neutral and objective view when watching slides of injured people, varying from low to high emotionality. In both experiments, after watching the clip/slides, participants were given a recognition memory test and a cued-recall test. There was a decrease in memory performance on the cued-recall test and the recognition memory test in the expressive suppression condition. There were no differences in memory between the reappraisal condition and the watch condition. In addition, Richards and Gross (2000) replicated these findings in a field experiment. Based on these findings they concluded that the strategy used to regulate emotion does influence memory for emotional material, although whether the different conditions simply resulted in different levels of attention being paid to stimuli remains an open question.

While Richards and Gross (2000) instructed participants which regulation strategy to apply, Myers et al. (1998) assessed which strategy participants adopted. Using the Manifest Anxiety Scale (MAS) to establish emotion regulation strategies, participants were divided into two groups: repressors and non-repressors. In a DF paradigm they observed poorer recall for negative TBF items in repressors relative to non-repressors, which they interpreted as repressors being better at limiting access to negative material that were instructed to forget (also see Xie et al., 2018). Motivated by these findings, measures of coping style were collected in the experiment described below.

This first experiment can be considered to be a baseline experiment. The sections below detail the materials and methods used in order to examine the influence of semantic relatedness, response criterion and emotion regulation strategies on memory control in a DF recognition memory paradigm with three valence categories: negative, positive and neutral. Despite some inconsistencies, as already described, the literature to date suggest similar memory sensitivity for emotional and neutral material while controlling for semantic relatedness (Buchanan et al., 2006; Dougal & Rotello, 2007; Talmi & Moscovitch, 2004). The prediction in the first experiment therefore is that memory sensitivity will not change between emotional and neutral material. Because of the inconsistencies among published work, it is less clear whether emotion will influence the DF effect (Bailey & Chapman, 2012; Brandt et al., 2013; Gallant et al., 2018; Minnema & Knowlton, 2008; Otani et al., 2012; Wessel & Merckelbach, 2006). However, if sensitivity does not change, comparable levels of directed forgetting might well accompany that. The second prediction is that there will be a relatively more liberal criterion for emotional material, in keeping with the general finding in the literature (Bailey & Chapman, 2012; Dougal & Rotello, 2007; Kapucu et al., 2008; Ochsner, 2000).

### 2.1.1 Methods

#### 2.1.1.1 Participants

There were 51 undergraduates (41 females, *M* = 20.66 years, *SD* = 2.97) from the University of Nottingham. After giving informed consent, participants completed the experiment and received either an inconvenience allowance of £2 or credits for their first-year courses. The experiment received ethical approval from the Ethics Committee in the School of Psychology at the University of Nottingham, as is the case for all experiments reported in this thesis.

#### 2.1.1.2 Design and Materials

A 2 (instruction: remember vs forget) x 3 (valence: negative vs positive vs neutral) within-subjects design was employed.

Three lists of 96 words (96 negative, 96 positive and 96 neutral) were selected from the Warriner, Kuperman and Brysbaert (2013) database. The three valence categories, negative, positive and neutral differed on the basis of valence (*M* = 2.92, *SD* = 0.63; *M* = 7.34, *SD* = 0.49; *M* = 5.45, *SD* = 0.42 respectively). The negative, positive and neutral words were matched for semantic relatedness, arousal (negative and positive words only), word length and word frequency. Semantic relatedness was measured using the open source tool 'snaut'. Word length and word frequency were measured using the open source tool SUBTLEX-UK (Van Heuven et al., 2014). A log transformation was used for the measures of word frequency in order to assess conformity to normality (but see Appendix D for differences in word frequency). At the level of the entire word sets independent sample t-tests were conducted to assess the equivalence of the semantic relatedness scores (range 0.80 – 0.98), arousal (range 2.25 – 6.90), word lengths (range 3 – 14 letters) and word frequencies (range 0.30 – 4.83) between the negative, positive and neutral words. The only significant differences were in arousal between emotional and neutral words (but see Appendix D for differences in word frequency). Full details of these outcomes can be seen in Appendix A (Table A1).

These three lists were then split into study (48 words) and test (96 words) lists. Two study lists were then formed, each containing 24 words of each valence type (72 words in total). The corresponding test lists each contained 144 words (48 words for each valence category, half of which were also on the study lists). A further set of lists was created in which the words designated as either TBR or TBF were alternated. Again, independent sample t-tests were conducted separately on the study and test lists. The outcomes are consistent with those described above (see Appendix A, Table A2).

The Emotion Regulation Questionnaire (ERQ) was used in order to measure emotion regulation strategies (Gross & John, 2003). The ERQ is a 10-item measure that assesses individual differences in two emotion regulation strategies. There are two subgroups, cognitive reappraisal ('*I control my emotions by changing the way I think about the situation I am in')* which has 6 items, and expressive suppression ('*I keep my emotions to myself')* which has 4 items. The items were scored on a 7-point Likert scale (1 = totally disagree to 7 = totally agree), with a higher score indicating more use of that strategy. Internal consistency on this measure was assessed by Gross and John (2003), where it was higher for the cognitive reappraisal subgroup than the expressive suppression subgroup.

#### 2.1.1.3 Procedure

Participants were tested individually, and the experiment lasted no more than 20 minutes. The experiment consisted of two study-test blocks. In the study phase, words were presented individually for 2000ms in the centre of a computer screen. After each word a blank screen appeared for 250ms followed by a remember cue (VVVVV in the colour green) or a forget cue (XXXXX in the colour red). Participants were instructed to attempt to remember the preceding word following a remember cue, and to forget the word when a forget cue followed. The instructions remained on the screen for 500ms and the order of remember and forget words was determined randomly for each participant. Following the cue, a blank screen appeared for 250ms before the next word was presented.

Test trials commenced with a fixation cross for 500ms, followed by a word. Participants were asked to make an old/new recognition judgment on the word, regardless of the TBR or TBF instruction given in the study phase, by pressing designated keys with their left and right index fingers on a keyboard. The hands used for old and new responses were counterbalanced across participants. Each word remained on the screen until participants pressed a response key. Once a response was made, the fixation cross was presented again for 500ms, followed by the next word.

After completing the recognition task on the computer, participants were asked to complete the ERQ questionnaire.

### 2.1.2 Results

Table 2.1 shows mean probabilities of correct judgements to old words (hits) and incorrect judgements to new words (false alarms) across instruction (remember, forget) and valence (negative, positive, neutral). Summary statistics across instruction and valence are presented in Table 2.2. As a guide, although the emphasis here is on differences in criterion across valence, a c-value of 0 indicates a neutral criterion, negative c-values indicate a liberal criterion and positive c-values indicate a conservative criterion.

Table 2.1
*Probabilities of Correct Judgements (Hits) to Old Items and Incorrect Judgements (False Alarms) to New Items across Instruction (Remember/Forget) and Valence (Negative, Positive and Neutral). SD = standard deviation*

| | | Valence | | |
|---|---|---|---|---|
| | | Negative<br>M (SD) | Positive<br>M (SD) | Neutral<br>M (SD) |
| Hits | Remember | 0.63 (0.15) | 0.62 (0.16) | 0.62 (0.16) |
| | Forget | 0.44 (0.17) | 0.48 (0.17) | 0.43 (0.17) |
| False alarms | | 0.18 (0.12) | 0.15 (0.10) | 0.21 (0.13) |

Table 2.2
*Summary Statistics for Each Valence Type across Instructions. d' is the Estimate of Memory Sensitivity and c is the Estimate of Criterion. SD = standard deviation*

| | | Valence | | |
|---|---|---|---|---|
| | | Negative<br>M (SD) | Positive<br>M (SD) | Neutral<br>M (SD) |
| *d'* | Remember | 1.39 (0.54) | 1.48 (0.49) | 1.24 (0.58) |
| | Forget | 0.86 (0.44) | 1.08 (0.40) | 0.70 (0.41) |
| c | Remember | 0.33 (0.37) | 0.41 (0.40) | 0.29 (0.41) |
| | Forget | 0.60 (0.44) | 0.61 (0.42) | 0.56 (0.47) |

A 2 x 3 repeated measures ANOVAs was conducted separately for sensitivity (*d'*) and criterion (c). In each case these analyses included the factors of instruction (remember and forget) and valence (negative, positive and neutral). Any violations of sphericity assumptions were adjusted by the Greenhouse-Geisser epsilon correction (this was done in every experiment reported in this thesis) (Greenhouse & Geisser, 1959). Identical ANOVA results for *d'* and criterion for the factor instruction and the interaction (see Table 2.3) are due to shared false alarm rates for TBR and TBF items (this is the case in every experiment reported in this thesis). Only main effects are elaborated on below, as there were no reliable interactions between factors.

#### 2.1.2.1 Effects of Instruction

There were main effects of instruction for both *d'* and c. In the case of sensitivity this reflects a directed forgetting effect: *d'* was superior for TBR than for TBF words. In the case of response criterion, there was a more liberal criterion for TBR words than for TBF words (see Table 2.3).

Table 2.3
*Summary of Repeated Measures ANOVA Results for Comparisons between Instruction and Valence*

| Measure (DV) | | *df* | *F* | *p* | $\eta_p^2$ |
|---|---|---|---|---|---|
| Instruction | *d'* | 1, 50 | 97.36 | < .001*** | .66 |
| | c | 1, 50 | 97.36 | < .001*** | .66 |
| Valence | *d'* | 2, 100 | 16.66 | < .001*** | .25 |
| | c | 2, 100 | 3.55 | .032* | .07 |
| Instruction x Valence Interactions | *d'* | 1.64, 82.09 | 1.32 | .269 | .03 |
| | c | 1.64, 82.09 | 1.32 | .269 | .03 |

*Notes.* * $p < .05$, *** $p < .001$.
df = degrees of freedom

#### 2.1.2.2 Effects of Valence

There were main effects of valence for both *d'* and c (see Table 2.3), and these were followed up via paired sample t-tests. Sensitivity was superior for positive ($M$ = 1.28, $SD$ = 0.38) and negative ($M$ = 1.12, $SD$ = 0.41) words relative to neutral ($M$ = 0.97, $SD$ = 0.42) words (positive vs neutral: $t(50) = 5.89$, $p < .001$, $d = .83$; negative vs neutral: $t(50) = 2.70$, $p = .009$, $d = .38$). Sensitivity was also superior for positive words relative to negative words ($t(50) = -3.02$, $p = .004$, $d = .43$).

Furthermore, there was a more liberal response criterion for neutral (*M* = 0.42, *SD* = 0.42) relative to positive (*M* = 0.51, *SD* = 0.39) words ($t(60) = 2.43$, $p = .019$, $d = .37$). Although not reliable, there is a trend in the same direction for negative words as well.

#### 2.1.2.3 Emotion Regulation

Bivariate correlations were conducted for the ERQ scale separately for reappraisal and suppression scores. These were plotted against sensitivity for each valence type, separated by directed forgetting instruction. The scatterplots revealed a small but significant positive relationship between the neutral forget condition and the expressive suppression condition ($r(49) = .29$, $p = .037$), see Figure 2.1 below for the scatterplot (see Appendix B for other scatterplots).

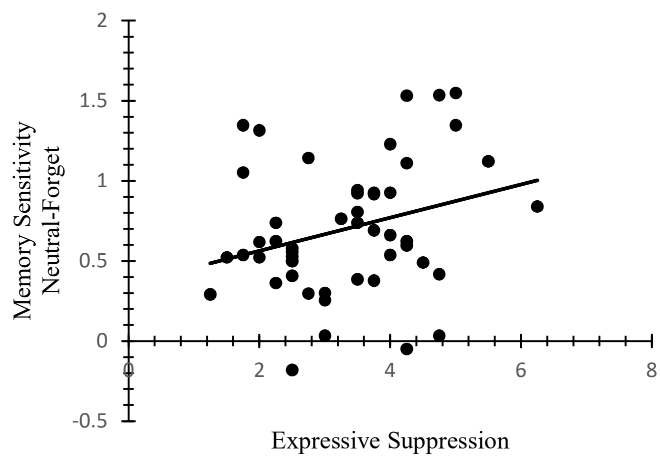



*Figure 2.1.* Scatterplot of the expressive suppression mean scores agains sensitivity in the forget condition for neutral items. This scatterplot is replotted in Figure B12 (see Appendix B) along with the scatterplots in the remaining conditions.

#### 2.1.2.4 Additional Study-Test Cycle Analyses

To ensure that participants were adhering to the directed forgetting instructions in the second study-test cycle (block 2), a 2 (instruction: remember vs forget) x 2 (block: block 1 vs block 2) repeated measures ANOVA was conducted. This was completed for correct old judgements (hits) and sensitivity measures by collapsing the TBR and TBF items across valence, separated by block. Table 2.4 shows mean probabilities of hits and *d'* measures across block (block 1, block 2) and instruction (remember, forget). The main effect of instruction was significant for both hits and *d'*, with higher hit rates and sensitivity for TBR compared to TBF items. There was also a main effect of block for hits, with superior hits in block 1 than block 2 (see Table 2.5 for the ANOVA results).

Table 2.4
*Averages of Hits and d' Measures for Remember and Forget Items Averaged across Valence*

| | | M (SD) | |
|---|---|---|---|
| | | Remember | Forget |
| Hits | Block 1 | 0.65 (0.02) | 0.49 (0.02) |
| | Block 2 | 0.60 (0.02) | 0.41 (0.03) |
| *d'* | Block 1 | 1.30 (0.07) | 0.91 (0.05) |
| | Block 2 | 1.23 (0.07) | 0.75 (0.06) |

Table 2.5
*Summary of Repeated Measures ANOVA Results for Comparisons between Instruction and Block*

| | Hits | | | | *d'* | | | |
|---|---|---|---|---|---|---|---|---|
| Measure (DV) | *df* | *F* | *p* | $\eta_p^2$ | *df* | *F* | *p* | $\eta_p^2$ |
| Instruction | 1, 50 | 102.59 | <.000*** | .67 | 1, 50 | 61.34 | <.000*** | .55 |
| Block | 1, 50 | 10.57 | .002** | .17 | 1, 50 | 2.88 | .096 | .05 |
| Instruction x Block Interactions | 1, 50 | 1.59 | .210 | .03 | 1, 50 | 1.62 | .209 | .03 |

*Notes.* ** p < .01, *** *p* < .001.
df = degrees of freedom

### 2.1.3 Discussion

The main objective in this baseline experiment was to investigate the links between emotion and memory control while controlling for semantic relatedness. The use of the ERQ also enabled an assessment of whether emotion regulation strategies affect memory performance. Sensitivity was superior for emotional words relative to neutral words, and participants were more conservative in their responding to positive words relative to neutral words as well. Moreover, there was no evidence that the directed forgetting effects obtained for negative, positive and neutral words were different.

The outcomes for measures of sensitivity are consistent with the view that memory for emotional material is superior to that for neutral material, as described in Chapter 1 (see section 1.4.2) (Kensinger & Corkin, 2003; Ochsner, 2000; Talmi et al., 2008; for reviews see Levine & Edelstein, 2009), and is counter to the prediction. This consistency, however, does not extend to measures of criterion. In those studies where criterion has been reported, which is a subset of those in which emotion and memory control have been examined, the consistent finding has been that emotional material was associated with a more liberal criterion (Bailey & Chapman, 2012; Hauswald et al., 2011; Marchewka et al., 2016). The findings in this experiment are exactly the opposite. This outcome will be considered in the context of findings in later experiments in this thesis.

The outcomes for the links between emotion and DF measurements are consistent with some previous work, but not other work. Outcomes in previous studies were shown in Table 1.1 above, and briefly, the outcomes in this experiment are consistent with those in the studies by Berger et al. (2018), Gallant et al. (2018), Gallant and Yang (2014), Marchewka et al. (2016), Taylor, Quinlan and Vullings (2018), Wessel and Merckelbach (2006) and Yang et al. (2012), who observed no differences in DF effects between emotional and neutral words. This outcome is in line with the prediction, although as already noted memory sensitivity differed, counter to predictions.

An important component of this experiment was the attempt to balance for semantic relatedness. In studies where relatedness has been controlled for and where the focus was only on the question of whether this factor influences estimates of memorability, the key findings have been that memory sensitivity is not different between emotional and neutral words (Dougal & Rotello, 2007; Kapucu et al., 2008). In two particularly relevant experiments, Dougal and Rotello (2007) tested memory for emotional and neutral words. In the first experiment, enhanced memory for emotional words was observed. However, when controlling for semantic relatedness in the second experiment, memory sensitivity was not different between emotional and neutral words. The outcomes described here are inconsistent with those of Dougal and Rotello (2007), because sensitivity advantages were still present despite controlling for relatedness.

It may be the case that these divergences are due in part to the different response requirements in the two experiments (R/K versus old/new discrimination), but it may also be the case that the variations are due to the fact that sensitivity measures may be inaccurate when criterion measures differ and only single point assessments are made available. Dougal and Rotello (2007) (among Kapucu et al., 2008; Windmann et al., 2002) have pointed out that measures of sensitivity and criterion are often confounded. What this means is that apparent differences in sensitivity might be misleading if they are accompanied by differences in criterion, and apparent similarities might be misleading for the same reason. This concern will also apply to measures of directed forgetting, since they are based on an assessment of sensitivity that cannot be separated from contributions due to criterion. One solution to these concerns is to acquire data points in a way that allows the two measures to be separated, and this will be described in more detail in the following section 2.2, following a brief consideration of emotion regulation and memory.

The ERQ was an exploratory measure used here to investigate the relationship between emotional regulation and control over emotional material. There was a positive relationship between memory for neutral TBF words and the expressive suppression coping style. This stands in contrast with results from Myers et al. (1998), who found that using a repressive coping style resulted in worse memory for only negative TBF items. Furthermore, in contrast with studies of emotion regulation discussed earlier (Hertel & McDaniel, 2010; Richards &

Gross, 2000), and with Myers et al. (1998), there was no effect of emotion regulation strategies on memory control for emotional material.

One explanation for the variability in findings lies in differences in how emotion regulation is measured and/or manipulated. Myers et al. (1998) used self-report questionnaires that measure levels of anxiety and defensiveness in order to measure to what extent participants relied on a repressive coping style. In this experiment a questionnaire that directly measures emotion regulation coping styles was used. As already described above, Richard and Gross (2000) instructed participants which strategies to apply (suppression or reappraisal) and showed that strategy use was related to memory for emotional material. Adopting an expressive suppression strategy resulted in a decrease in memory performance, whereas a reappraisal strategy did not. It may be that the absence of data consistent with this outcome in this experiment is because what was measured was the tendency to regulate emotion in particular ways, rather than a direct manipulation of it. This account does not explain, however, findings in studies where memory for emotional material has varied according to assessments of traits. For example, Ho, Cheng and Dai (2017) investigated the links between anxiety and memory control using item-method directed forgetting, and observed that higher levels of anxiety were associated with a decreased DF effect for negative words.

It should be noted, however, that the significant correlation should be interpreted with caution considering the context of the analyses. A total of 12 correlations (see Appendix B) were computed without multiple comparison correction. Given that the positive correlation has a p-value of .037, this would not survive any correction of this kind.

The reasons for the disparities reported in this experiment are not entirely clear. They are not, however, independent of considerations of the accuracy of the memory measures (i.e. single point measures) that are used to assess links between emotion, memory and memory control. Methods that can overcome this confound between sensitivity and response criterion measures are considered in Experiments 2 and 3, described below.

### 2.2 Experiment 2 – Emotion, Directed Forgetting and Confidence Judgements

In Experiment 1, when controlling for semantic relatedness, sensitivity was higher for both classes of emotional material, and – somewhat surprisingly – there was a more conservative criterion for positive relative to neutral words. Emotionality did not have a significant effect on the ability to remember or forget, although there was evidence for a standard DF effect: participants made more accurate memory judgements to TBR than to TBF words. One possible reason for the inconsistency between the findings in Experiment 1 and the existing literature (Bailey & Chapman, 2012; Brandt et al., 2013; Hauswald et al., 2011; Liu et al., 2017; Myers et al., 1998; Payne & Corrigan, 2007), and indeed between findings in other published studies, is the measures of memory that have been used.

In the most commonly employed Signal Detection Theory (SDT) measure, it is assumed that the strength distributions of old and new items after an encoding manipulation (study phase) are equal and that these two distributions overlap (Snodgrass & Corwin, 1988), whereas empirically this is not the case (see section 1.2 for a detailed description) (Wixted, 2007; Yonelinas, 2001a; Yonelinas et al., 1996; Yonelinas & Parks, 2007; for a review of such findings see Yonelinas, 1994). Under this assumption, *d'* tends to overestimate sensitivity for a conservative criterion and underestimate sensitivity for a liberal criterion when distributions do not overlap (Dougal & Rotello, 2007).

The most common finding regarding emotional material when assessed via recognition memory is a relatively more liberal response criterion relative to neutral material (Bailey & Chapman, 2012; Dougal & Rotello, 2007; Kapucu et al., 2008; Ochsner, 2000). Hence, differences in memory sensitivity may actually reflect changes in response criterion across conditions (Dougal & Rotello, 2007; Phelps & Sharot, 2008). Despite a different set of criterion outcomes, the same argument applies to interpreting the data and analyses for Experiment 1. The same criticism can be directed at previous work (Bailey & Chapman, 2012; Hauswald et al., 2011; Marchewka et al., 2016) and may be one factor contributing to the general variability across studies in the findings already described.

This concern can be ameliorated by acquiring additional measures rather than simply old/new memory decisions from which an assessment of sensitivity and criterion can be derived. One way to do this is via the acquisition of confidence judgements and plotting receiver operating characteristic (ROC) curves.

ROCs offer two important opportunities in the context of the outcomes of Experiment 1. The first is that they provide a way of separately assessing changes in sensitivity and in criterion (see section 1.2.2 for a detailed description). The second is that they provide a means of estimating the contributions of recollection and familiarity to memory judgements, which is not possible using single-point measures. For current purposes, the specific questions of interest are: how does emotion influence response criterion and memory control, how do recollection and familiarity vary according to attempts to control forgetting, and how do they interact with the emotional status of stimuli? In line with previous findings, it is expected to find similar memory sensitivity for emotional and neutral material when controlling for semantic relatedness. While the extant literature would suggest a relatively more liberal criterion for emotional material, the outcomes of Experiment 1 would not support this prediction. However, considering that measures of sensitivity and criterion may have been confounded, the results observed in Experiment 1 may not be an appropriate guideline for the following experiments. Hence, controlling for the sensitivity/criterion confound would lead to the same prediction made previously: a more liberal response criterion for emotional material. The same reasoning applies here for memory sensitivity. In addition, following the observed effects of memory control in Experiment 1, it is expected to find similar effects in the following experiment.

Earlier in this thesis a prediction was made concerning memory control and recognition memory processes. Previous DF studies, to the best of my knowledge, have not used behavioural measures to investigate memory control and recollection and familiarity. However, some investigation of old/new effects in a DF paradigm has been done (Nowicka et al., 2009; Van Hooff et al., 2009), and (partially) based on these findings it is predicted that memory control will not operate over familiarity: it is predicted that there will be a directed forgetting effect only for recollection.

Experiment 2 includes three studies. Each has the same design other than the valence of the words employed, in that each experiment has only two valence categories; negative vs neutral, negative vs positive and positive vs neutral. The negative vs positive gives the opportunity to directly investigate whether negative and positive material have different effects on memory or whether these two classes can be considered as having similar effects. In addition, pairing the valence types across three studies allows for increasing the number of items per category in an experiment of a reasonable length, thereby increasing power. As for the first experiment in this thesis, the experiments described below employ the item-method DF paradigm. ROC data was collected by means of confidence judgements (Glanzer et al., 1999; Yonelinas et al., 1996).

### 2.2.1 Methods

Given the similarities between the designs of the three studies they are described jointly below.

#### 2.2.1.1 Participants

There were 44 undergraduates from the University of Nottingham for each study (study 1: 38 females, *M* = 22.05 years, *SD* = 4.96; study 2: 37 females, *M* = 19.37 years, *SD* = 2.41; study 3: 37 females, *M* = 20.41 years, *SD* = 4.20). After giving informed consent, participants completed the experiment and received either an inconvenience allowance of £3 or credits towards their first-year credit requirement.

#### 2.2.1.2 Design and Materials

Stimuli: Study 1 consisted of negative and neutral words, study 2 negative and positive words and study 3 positive and neutral words.

A 2 (instruction: remember vs forget) x 2 (valence) within-subjects design was employed. Each valence category consisted of 180 words which were selected from the Warriner et al. (2013) database. The valence categories in all studies differed on the basis of valence (negative: *M* = 2.56, *SD* = 0.63; positive: *M* = 7.33, *SD* = 0.57; neutral: *M* = 5.12, *SD* = 0.42). All words in each study were matched for semantic relatedness (distance), word length and word frequency using the methods described in Experiment 1. At the level of the entire word sets independent sample t-tests were conducted to assess the equivalence of the semantic relatedness scores (study 1: range 0.82 – 0.98; study 2: range 0.82 – 0.99; study 3: range 0.82 – 0.99), arousal (study 1: range 2.25 – 7.27; study 2: range 2.25 – 7.27; study 3: range 2.50 – 7.05), word lengths (all studies: range 3-14 letters) and word frequencies (study 1: range 0.30 – 4.83; study 2: range 1.11 – 4.65; study 3: range 0.30 – 4.83) between the valence categories in each study. The only significant differences were in arousal in all studies (but see Appendix D for differences in word frequency). Full details of these outcomes can be seen in Appendix C (Table C1).

The 180 words (each for the three valence categories) were then split into study (90 words) and test (180 words) lists. Three study lists were then formed, each containing 30 words of each valence type (60 words in total). The corresponding test lists each contained 120 words (60 words for each valence category, half of which were also on the study lists). A further set of lists was created in which the words designated as either TBR or TBF were alternated. Again, independent sample t-tests were conducted separately on the study and test lists. There were significant differences for arousal between the emotional and neutral words, and arousal was matched between negative and positive words (see Appendix C, Table C2 and Table C3 and Appendix D). Another set of lists was then created for the counterbalancing of the designated instructions (TBR and TBF) and old/new status.

#### 2.2.1.3 Procedure

Participants were tested individually, and the study lasted no more than 30 minutes. The experiment consisted of three study-test blocks. The procedure for the study phase on each trial was identical to that in Experiment 1 (see section 2.1.1.3).

Test trials commenced with a fixation cross for 500ms, followed by a word. Participants were asked to make an old/new recognition judgment on the word based on confidence ratings, regardless of the TBR or TBF instruction given in the study phase. They responded according to a six-point rating scale (*1 = sure new to 6 = sure old*) by using the computer mouse to select a value on a rating scale. Each word, together with the rating scale, remained on the screen until participants made a response (see Appendix E for an example of the screen). Once a response was made, the fixation cross was presented again for 500ms, followed by the next word.

### 2.2.2 Results

Table 2.6 shows mean probabilities of hits and false alarms across instruction (remember, forget) and valence (negative, positive, neutral) for study 1, 2 and 3 (see Appendix F for proportions of responses for each confidence rating). Summary statistics across instruction and valence for each study are presented in Table 2.7. Recollection and Familiarity were measured using a ROC toolbox in MATLAB based on the DPSD model (Koen et al., 2017).

Table 2.6
*Probabilities of Correct Judgements (Hits) to Old Items and Incorrect Judgements (False Alarms) to New Items across Instruction (Remember/Forget) and Valence (Negative, Positive and Neutral) for Studies 1, 2 and 3. SD = standard deviation*

| | | Study 1 | | Study 2 | | Study 3 | |
|---|---|---|---|---|---|---|---|
| | | Negative M (SD) | Neutral M (SD) | Negative M (SD) | Positive M (SD) | Positive M (SD) | Neutral M (SD) |
| Hits | Remember | 0.74 (0.12) | 0.67 (0.14) | 0.73 (0.11) | 0.70 (0.13) | 0.65 (0.15) | 0.63 (0.19) |
| | Forget | 0.68 (0.15) | 0.59 (0.15) | 0.66 (0.14) | 0.62 (0.15) | 0.59 (0.15) | 0.53 (0.17) |
| False alarms | | 0.25 (0.12) | 0.17 (0.10) | 0.23 (0.08) | 0.22 (0.09) | 0.22 (0.11) | 0.17 (0.09) |

Table 2.7
*Summary Statistics for Each Valence Type across Instructions for Studies 1, 2 and 3. R is the Estimate of Participants' Judgements Driven by Recollection, F is the Estimate of Participants' Judgements Driven by Familiarity, d' is the Estimate of Memory Sensitivity and c is the Estimate of Criterion. SD = standard deviation*

| | | Study 1 | | Study 2 | | Study 3 | |
|---|---|---|---|---|---|---|---|
| | | Negative M (SD) | Neutral M (SD) | Negative M (SD) | Positive M (SD) | Positive M (SD) | Neutral M (SD) |
| R | Remember | 0.27 (0.21) | 0.30 (0.18) | 0.24 (0.16) | 0.22 (0.15) | 0.28 (0.18) | 0.27 (0.16) |
| | Forget | 0.21 (0.20) | 0.26 (0.18) | 0.18 (0.13) | 0.16 (0.14) | 0.20 (0.17) | 0.19 (0.17) |
| F | Remember | 1.08 (0.39) | 1.09 (0.42) | 0.93 (0.30) | 0.84 (0.38) | 0.88 (0.33) | 0.94 (0.40) |
| | Forget | 0.92 (0.40) | 0.91 (0.37) | 0.81 (0.33) | 0.73 (0.31) | 0.76 (0.33) | 0.76 (0.37) |
| *d'* | Remember | 1.47 (0.46) | 1.53 (0.61) | 1.44 (0.36) | 1.35 (0.42) | 1.29 (0.44) | 1.44 (0.48) |
| | Forget | 1.27 (0.54) | 1.32 (0.57) | 1.23 (0.37) | 1.12 (0.37) | 1.12 (0.47) | 1.15 (0.47) |
| c | Remember | 0.04 (0.32) | 0.28 (0.30) | 0.06 (0.30) | 0.11 (0.31) | 0.24 (0.38) | 0.35 (0.35) |
| | Forget | 0.14 (0.32) | 0.41 (0.32) | 0.17 (0.33) | 0.23 (0.32) | 0.32 (0.41) | 0.50 (0.38) |

For each study, 2 x 2 repeated measures ANOVAs were conducted on the four measures shown in Table 2.7: sensitivity (*d'*), response criterion (c) and estimates of recollection (R) and familiarity (F). In each case these analyses initially included the factors of instruction (remember vs forget) and valence (two categories in each case). Only main effects are elaborated on below, as there were no reliable interactions between factors.

#### 2.2.2.1 Effects of Instruction

There were main effects of instruction in all three studies for all measures. In the cases of sensitivity, recollection and familiarity this reflected a directed forgetting effect: *d'*, R and F were all superior for TBR than for TBF words (see Table 2.8). The main effect of instruction for response criterion reflects a more liberal criterion for TBR than for TBF words in each case.

Table 2.8

*Summary of Repeated Measures ANOVA Results for Comparisons between Instruction and Valence in Studies 1, 2 and 3*

| | | Study 1 | | | | Study 2 | | | | Study 3 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Measure (DV) | | *df* | *F* | *p* | $\eta_p^2$ | *df* | *F* | *p* | $\eta_p^2$ | *df* | *F* | *p* | $\eta_p^2$ |
| Instruction | R | 1, 43 | 6.71 | .013* | .14 | 1, 43 | 11.89 | .001** | .22 | 1, 43 | 10.13 | .003** | .19 |
| | F | 1, 43 | 14.31 | < .001*** | .25 | 1, 43 | 14.34 | < .001*** | .25 | 1, 43 | 13.38 | .001** | .24 |
| | *d'* | 1, 43 | 17.46 | < .001*** | .29 | 1, 43 | 23.11 | < .001*** | .35 | 1, 43 | 18.43 | < .001*** | .30 |
| | c | 1, 43 | 17.46 | < .001*** | .29 | 1, 43 | 23.11 | < .001*** | .35 | 1, 43 | 18.43 | < .001*** | .30 |
| Valence | R | 1, 43 | 4.16 | .048* | .09 | 1, 43 | 0.62 | .437 | .01 | 1, 43 | 0.33 | .570 | .01 |
| | F | 1, 43 | 0.00 | .950 | .00 | 1, 43 | 4.71 | .036* | .10 | 1, 43 | 0.35 | .557 | .01 |
| | *d'* | 1, 43 | 1.16 | .287 | .03 | 1, 43 | 6.80 | .012* | .14 | 1, 43 | 2.64 | .112 | .06 |
| | c | 1, 43 | 77.13 | < .001*** | .64 | 1, 43 | 8.40 | .006* | .16 | 1, 43 | 32.62 | < .001*** | .43 |
| Instruction x Valence Interactions | R | 1, 43 | 0.06 | .803 | .00 | 1, 43 | 0.00 | .950 | .00 | 1, 43 | 0.03 | .858 | .00 |
| | F | 1, 43 | 0.05 | .829 | .00 | 1, 43 | 0.02 | .892 | .00 | 1, 43 | 0.92 | .344 | .02 |
| | *d'* | 1, 43 | 0.39 | .540 | .01 | 1, 43 | 0.06 | .803 | .00 | 1, 43 | 3.25 | .078 | .07 |
| | c | 1, 43 | 0.39 | .540 | .01 | 1, 43 | 0.06 | .803 | .00 | 1, 43 | 3.25 | .078 | .07 |

*Notes.* * $p < .05$, ** $p < .01$, *** $p < .001$.
df = degrees of freedom

#### 2.2.2.2 Effects of Valence

There were main effects of valence across all studies for all measures. Sensitivity was superior for negative words relative to positive words in study 2, which is evident in the ROC curves (see Figure 2.2B). The ROC curves in study 2 (B) show that the curve for negative words falls more towards the upper left compared to the curve for positive words. This is the case in both the remember and forget condition. Moreover, the ROCs in study 1 and 3 (see Figure 2.2A and C, respectively) show that the curves for both the valence conditions overlap, indicating no differences in sensitivity between emotional and neutral words (see Table 2.8). Recollection, in study 1, was superior for neutral words relative to negative words. Familiarity was superior for negative relative to positive words in study 2.

For response criterion, there was a more liberal criterion for emotional words compared to neutral words (study 1 and 3) and a more liberal criterion for negative words than for positive words in study 2 (see Table 2.8). The more liberal criterion for emotional words can be seen in the ROC curves (see Figure 2.2A and C). The cumulative hit/false alarm pairings by confidence fall consistently more towards the right for emotional words relative to neutral words. These data points indicate higher hit and false alarm rates at consecutive confidence levels: a more liberal response criterion. Moreover, the more liberal criterion for negative words in study 2 (see Figure 2.2B), is visible in the first few cumulative hit/false pairings that fall more towards the right for negative words relative to positive words. Similar to study 1 and 3, this indicates a more liberal response criterion.

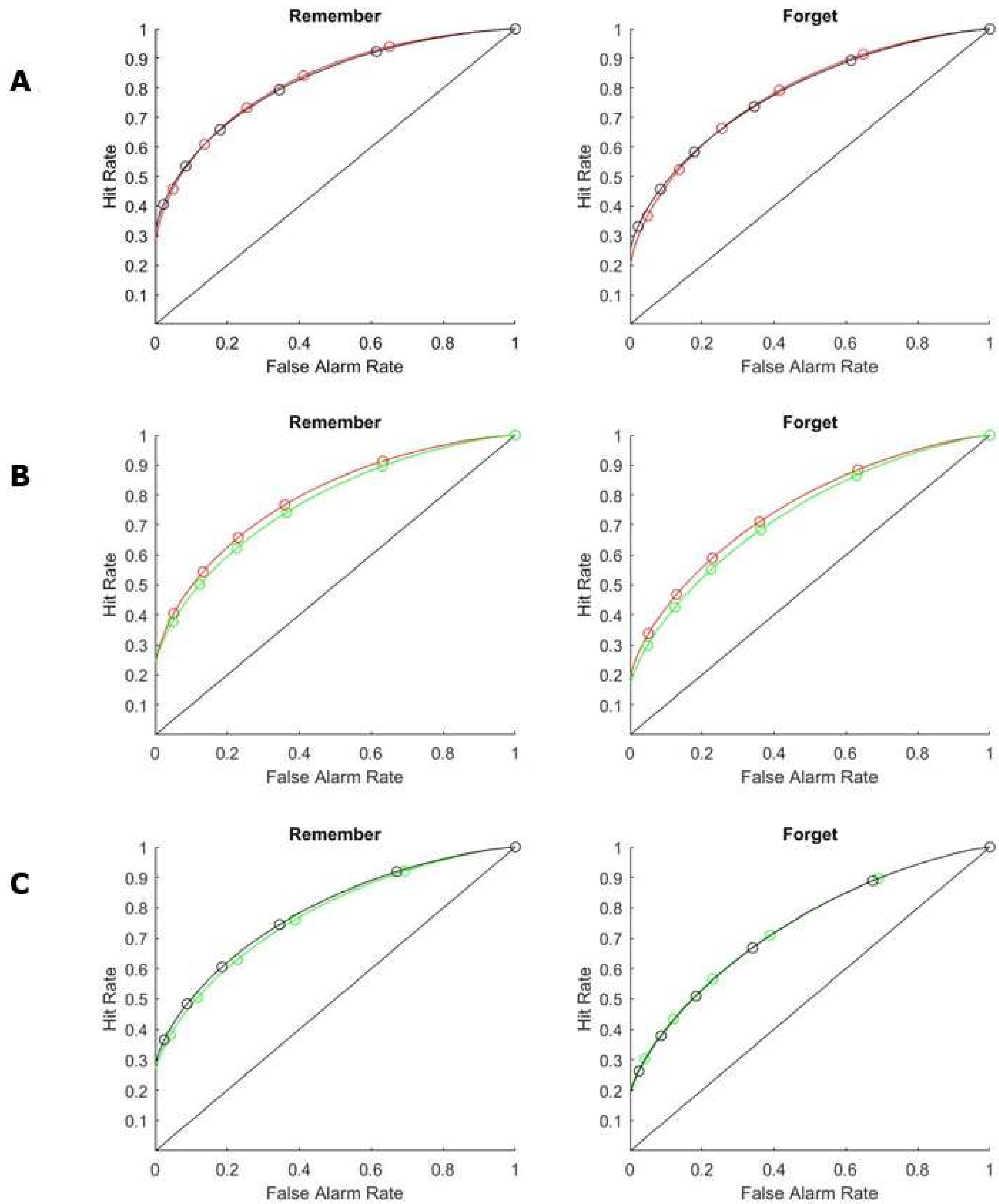



*Figure 2.2.* Receiver operating characteristic (ROC) data for study 1 (A), study 2 (B) and study 3 (C). The red curves denote the negative condition, the green curves denote the positive condition and the black curves denote the neutral condition. The curves indicate memory strength and the circles indicate the level of response confidence (response criterion)

#### 2.2.2.3 ROC Analyses – Area Under the Curve (AUC)

Additional analyses of the ROC curves were conducted using Area Under the Curve (AUC) measures to examine effects of valence and instruction. The AUC was measured using the ROC Toolbox in Matlab (Koen et al., 2017). The results (see Table 2.9 for the ANOVA results) are similar to the results from the *d'* analyses. That is, there are consistent DF effects across all three studies, which were not affected by valence. In study 2, there was also a significant main effect of valence. Memory was superior for negative ($M$ = 0.76, $SD$ = 0.01) than for positive ($M$ = 0.74, $SD$ = 0.01) items.

Table 2.9
*Summary of Repeated Measures ANOVA Results for Comparisons between Instruction and Valence in Studies 1, 2 and 3 on the AUC measures*

| Measure (DV) | Study 1 | | | | Study 2 | | | | Study 3 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | *df* | *F* | *p* | $\eta_p^2$ | *df* | *F* | *p* | $\eta_p^2$ | *df* | *F* | *p* | $\eta_p^2$ |
| Instruction | 1, 43 | 26.61 | <.000*** | .38 | 1, 43 | 33.98 | <.000*** | .44 | 1, 43 | 19.60 | <.000*** | .31 |
| Valence | 1, 43 | 0.02 | .880 | .00 | 1, 43 | 12.32 | .001** | .22 | 1, 43 | 0.12 | .734 | .00 |
| Instruction x Valence Interactions | 1, 43 | 0.35 | .555 | .01 | 1, 43 | 0.02 | .879 | .00 | 1, 43 | 1.09 | .303 | .03 |

*Notes.* ** $p < .01$, *** $p < .001$.
df = degrees of freedom

### 2.2.3 Discussion

One intention in this experiment was to develop an understanding of how control over remembering and forgetting of emotional material is exerted while using ROC measures. A second intention was to characterise how memory differences for emotional materials manifest themselves. Single point measures (e.g. old/new recognition judgements) can yield indicators of sensitivity as well as criterion. They may, however, be conflated as described earlier, which motivated the use of ROCs in this set of studies. They permit a more fine-grained assessment of these two components of performance on memory tasks (Dougal & Rotello, 2007; Kapucu et al., 2008).

#### 2.2.3.1 Memory Sensitivity and Directed Forgetting

There were no interactions between the directed forgetting (Instruction) and emotion (Valence) manipulations, so these elements of the data are considered separately. A directed forgetting effect was found in all studies for the estimates of recollection and familiarity that were derived from the ROC data, and for memory sensitivity. In terms of sensitivity, this is consistent with common findings in the directed forgetting paradigm (MacLeod, 1999). For recollection and familiarity, this outcome is not in line with the prediction that control will operate only over recollection. The outcomes in this experiment suggest that memory control – at least as operationalised as directed forgetting – can influence both processes.

The DF effect did not vary according to emotion, nor were there any differences in sensitivity between emotional and neutral words. The latter outcome is in line with findings from Dougal and Rotello (2007). This is also somewhat comparable with findings from other studies suggesting that emotion itself does not influence sensitivity (Kapucu et al., 2008; Windmann & Kutas, 2001). These data do not align with the outcomes in Experiment 1, however, with two possibilities for this being the use of single-point measures, as already discussed, and/or the use of three valence categories in the same experiment. In this experiment, differences in sensitivity were only observed in study 2, with superior sensitivity for negative relative to positive words. This is somewhat in line with other studies reporting superior sensitivity for negative material relative to positive material (Minnema & Knowlton, 2008; Otani et al., 2012). There are, however, also inconsistent findings compared to previous studies where sensitivity was enhanced for emotional material (both negative and positive) relative to neutral material (Brandt et al., 2013; Hauswald et al., 2011; Minnema & Knowlton, 2008; Otani et al., 2012; Yang et al., 2016), and in which DF effects were diminished for emotional relative to neutral material (Bailey & Chapman, 2012; Hauswald et al., 2011; Otani et al., 2012; Payne & Corrigan, 2007; Yang et al., 2016).

The ROC curves presented in Figure 2.2 demonstrate separately changes in *d'* and criterion. However, the assumptions of the calculation of *d'* – equal variance distributions for old and new items and a slope of 1.0 in zROCs (Snodgrass & Corwin, 1988; Wixted, 2007) – are not commonly met in recognition memory tasks. As mentioned in the Introduction (see section 1.2.2), zROCs in recognition memory tasks commonly have slopes with values less than 1.0 (Yonelinas, 1994), violating this assumption of *d'* (Verde & Rotello, 2003). If this is the case in this experiment then other measures of sensitivity would be more appropriate for measuring possible effects of emotion and directed forgetting.

An evaluation of the zROCs in Experiment 2 confirms previous observations of a slope less than 1.0 in recognition memory (Yonelinas, 1994) (see Appendix H, Figure H1 for the zROCs of Experiment 2 separated for studies 1, 2 and 3). As a consequence, this would suggest that another measure of sensitivity, such as $d_a$, would be more appropriate. The measure of $d_a$ is an alternative to *d'*, that in contrast allows for possible variations in the old and new item distributions (Verde et al., 2006; Verde & Rotello, 2003). The repeated measures ANOVA conducted on *d'* measures as described in section 2.2.2 were also conducted on $d_a$. These analyses revealed the same effects of emotion on directed forgetting as already described (see Appendix G, Table G1 for the ANOVA results based on $d_a$); a directed forgetting effect in all studies, which did not vary according to emotion and enhanced memory sensitivity for negative words compared to positive words in study 2 (see Table 2.8 for the ANOVA results based on *d'*). Thus, using a different measure for sensitivity did not yield different effects.

Despite the similarity between the findings when using *d'* or $d_a$, based on the zROCs presented – which are consistently below 1 – it would be more appropriate to use $d_a$ (Verde &

Rotello, 2003). However, when there are minimal differences in slopes between conditions that are to be compared, using $d_a$ will not lead to very different results as to using *d'* (Simpson & Fitter, 1973). This is the likely reason for the outcomes in Experiment 2. The largest difference between slopes across conditions that were contrasted directly was 0.04 (this was in the contrast between the negative TBF and neutral TBF conditions). In other experiments the largest difference for any give contrast is 0.07 (this is in the contrast between negative TBF and neutral TBF in Experiment 3). The outcome of the analysis in this case for $d_a$ mirrors that for *d'*. A full account of the zROCs for Experiment 3 (see Appendix H, Figure H2) provides reassurance that using $d_a$ or *d'* will not substantially change the findings and conclusions drawn in this thesis. On this basis, *d'* as a measure of sensitivity is reported in the main text in subsequent experiments.

#### 2.2.3.2 Response Criterion

Turning to response criterion, this was more liberal for both emotion categories relative to neutral words and more liberal for negative relative to positive words. While this outcome contrasts with what was found in Experiment 1, it replicates several studies where participants adopted a more liberal response criterion for emotional material, albeit based on assessments of single point measures (Bailey & Chapman, 2012; Dougal & Rotello, 2007; Hauswald et al., 2011; Kapucu et al., 2008; Marchewka et al., 2016; Windmann & Kutas, 2001). A possible explanation for the discrepancies between Experiment 1 and Experiment 2 lies in the number of valence conditions used; three valences combined (Experiment 1) relative to contrasts of two valences in separate studies (Experiment 2). A combination of three or two valence conditions could result in different outcomes as it results in different proportions of emotional and neutral stimuli. In other DF studies, however, a more liberal criterion for emotional material has consistently been observed regardless of the number of valence conditions used (Bailey &

Chapman, 2012; Berger et al., 2018; Gallant et al., 2018; Hauswald et al., 2011; Marchewka et al., 2016; Yang et al., 2012).

Another difference which might explain the discrepancies between Experiment 1 and 2, is the use of ROC measures in Experiment 2. The findings in Experiment 2 are more comparable with the findings from Dougal and Rotello (2007). Using ROC measures, they also found a more liberal criterion for emotional material, whilst there were no differences in memory due to emotion. The separation of sensitivity and criterion measures using ROCs indicates that even though response criterion changes across valence conditions, memory does not. This reinforces the argument that when using single-point measures, as in Experiment 1, it becomes challenging to interpret changes in memory sensitivity when response criterion changes as well. The combination of tight control over stimulus properties and the removal of the sensitivity/criterion confound suggests greater confidence in the findings presented in Experiment 2 than in Experiment 1, and supports the prediction for subsequent studies that the relatively more liberal criterion for emotional material will remain.

#### 2.2.3.3 Recollection and Familiarity

Recollection was superior for neutral compared to negative words in study 1, and familiarity was superior for negative compared to positive words in study 2. Linking this with previous studies that have applied ROC curves to extract measures of recollection and familiarity, in general, emotion has been found to increase recollection (in comparison with neutral material). For familiarity, however, little or no effects of emotion have been found in some studies (McCullough & Yonelinas, 2013; Ritchey et al., 2008; Yonelinas et al., 2011; Yonelinas & Ritchey, 2015). For example, Ritchey et al. (2008) examined subsequent memory effects for negative and neutral images in an fMRI study. Confidence judgements were collected in order to plot ROC curves, which also allowed the researchers to extract estimates of recollection and familiarity. They found recollection to be superior for negative images relative to neutral images, whereas there were no differences in familiarity between negative and neutral images. Similar effects have been reported in studies that investigated the effects of stress on memory for emotional and neutral images (McCullough & Yonelinas, 2013; Yonelinas et al., 2011). Using the R/K paradigm, Ochsner (2000) examined memory for emotional and neutral images. Enhanced memory for emotional images resulted only from increased 'Remember' judgements, which are assumed to reflect recollection (Rugg & Curran, 2007; Voss et al., 2012; Yonelinas, 1994). In a combined ROC and R/K study already described, Dougal and Rotello (2007) observed increased Remember judgements for negative material relative to neutral material. However, they argued that the increased Remember judgements reflected a subjective experience of recollection due to increased 'sense of familiarity' of emotional words. In addition, in their studies, familiarity did not vary according to emotion.

One explanation for the discrepancies in findings between the present experiment and existing literature is because different methods have been used across studies to measure recognition memory and the contributions of recollection and familiarity. In the present experiment, confidence judgements were used. ROC measures were then employed to calculate the contribution of recollection and familiarity. In contrast, Ochsner (2000) used the R/K paradigm, in which participants are explicitly instructed to base their judgements on whether or not they remember contextual information. However, Sharot et al. (2007) used a combination of ROC confidence judgements and the R/K procedure to estimate recollection and familiarity and the influence of emotion. Both methods resulted in similar outcomes: they observed enhanced memory for emotional material driven by recollection for healthy participants (also see Dougal & Rotello, 2007). Importantly, however, confidence judgments were given after a remember/know response was made.

The behavioural findings indicating increased recollection for emotional material have been supported by the outcomes in ERP studies. For example, Zheng et al. (2018) used a standard old/new recognition test, and found a larger left-parietal old/new effect for emotional relative to neutral material. In so far as the parietal old/new effects indexes recollection, this is consistent with the majority of behavioural findings (but see Maratos et al., 2000; Windmann & Kutas, 2001).

It has been observed, however, that memory judgements in recognition memory tasks likely rely heavily on familiarity (Johansson et al., 2004; Tulving, 1985b) by virtue of the task demands. This may not be the case, however, in the same way, when confidence judgements or R/K judgements are required. Under these circumstances, the task instructions might encourage a greater reliance on recollection (assuming it is a consciously controlled process) than when only old/new judgements are required. This consideration is important for the ERP experiments described in Chapter 3, but before that the outcomes in another set of behavioural experiments are reported.

#### 2.2.3.4 Employing the R/K Procedure

In order to investigate the effect of emotion on memory control and processes of recollection and familiarity further, the Remember/Know procedure is used in Experiment 3. Based on a particular set of assumptions, as already described, ROCs can be used to estimate recollection and familiarity. Under the assumptions employed here (Parks & Yonelinas, 2007; Yonelinas, 1994; Yonelinas et al., 1996; Yonelinas & Jacoby, 1995; Yonelinas & Parks, 2007) recollection is assumed to be a pure reflection of high confidence judgements, whereas the remaining (lower confidence judgements) are assumed to be based on familiarity. This is an important assumption, because it precludes the highest confidence judgements being based on familiarity, as well as encompassing the strong view that recollection is not graded in the way it is reflected in confidence judgements. In his extensive review of research into recollection and familiarity, Yonelinas (2002) argued that stronger conclusions about the likely roles played by recollection and familiarity in supporting memory judgements could be inferred when similar outcomes are obtained via different approaches designed to assess the processes of interest (Yonelinas & Jacoby, 1995). In keeping with this logic, one way to investigate the generality of the findings in Experiment 2 is to maintain the same stimulus sets and employ a different measure for estimating recollection and familiarity. This was done by using the R/K procedure combined with confidence judgements in Experiment 3.

## 2.3 Experiment 3 – Emotion, Directed Forgetting and the R/K Procedure

The main objective of Experiment 2 was to estimate the contributions of sensitivity and criterion separately in tasks where valence and directed forgetting were manipulated. The experiments also provided an opportunity to assess the contributions of recollection and familiarity and how they interact with valence and DF instructions, because estimates of these processes can be derived from ROC data (Yonelinas, 1994). In Experiment 2 there were no differences in memory sensitivity due to emotion, nor did emotion influence the directed forgetting (DF) effect. There was also a relatively more liberal criterion for emotional words. This finding is important because it was obtained in a task where there was not the dependence on single-point measures. This had only been done, in a somewhat different design, in one previous study (Dougal & Rotello, 2007). Using negative, positive and neutral words, they also found a more liberal criterion while memory was equivalent for emotional and neutral words.

Another means of extracting measures of these two memory processes is the Remember/Know (R/K) procedure (Tulving, 1985b). This procedure was employed here as an attempt to assess the generality of the findings from Experiment 2: to determine whether the ways in which recollection and familiarity vary with emotion and instructions is stable across different means of estimating recollection and familiarity.

In the R/K procedure it is assumed that Remember judgements reflect recollection and Know judgements reflect familiarity (Tulving, 1985b; Yonelinas & Parks, 2007). Common assumptions concerning the processes of recollection and familiarity are that they are independent and Know judgements reflect familiarity in absence of recollection (Yonelinas & Jacoby, 1995). This assumption means that some items are both familiar and recollected, and hence in order to measure recollection and familiarity in the R/K procedure a correction needs to be applied, otherwise the measure K (Know) alone will commonly underestimate the contribution of familiarity. To make estimates of familiarity under the independence assumption, the following equation is used:

$$F = \frac{K}{1 - R}$$

In this equation F represents familiarity, K represents the probability of Know responses and R represents recollection (Yonelinas, 2002; Yonelinas & Jacoby, 1995).

In summary, the key objective in this experiment is to establish whether the results obtained using ROCs generalise when another measure of estimating recollection and familiarity (involving slightly different assumptions) is employed. Stimulus sets and all other aspects of the experimental design were held constant. Furthermore, as described earlier, single-point measures may not be sufficient to make accurate interpretations on the influence of emotion on memory sensitivity. Therefore, as in Experiment 2, the experiments described below include a combination of the R/K procedure and confidence judgements. Based on findings of Experiment 2, it is expected to find similar memory sensitivity for emotional and neutral material and no effect of emotion on the directed forgetting effect. It is also expected to find a relatively more liberal criterion for emotional material. Furthermore, based on previous findings (Dougal & Rotello, 2007; McCullough & Yonelinas, 2013; Ritchey et al., 2008;

Yonelinas et al., 2011) it would be expected only to find differences in recollection between emotional and neutral material. This, however, contrasts with the findings in Experiment 2. According to Yonelinas (2002) the ROC measures and the R/K procedure should lead to similar results, although they are based on somewhat different assumptions. If the two methods of estimating recollection and familiarity correspond (which depends primarily on the assumption that a R (Remember) response equates with a high confidence 'old' response), it is expected to observe superior recollection for neutral compared to negative material and superior familiarity for negative compared to positive material. Moreover, in line with the findings in Experiment 2, a directed forgetting effect is expected to be observed for recollection and for familiarity.

### 2.3.1 Methods

Given the similarities between the designs of the three studies they are described jointly below.

#### 2.3.1.1 Participants

There were 44 undergraduates from the University of Nottingham for each study (study 1: 33 females, *M* = 21.33 years, *SD* = 3.23; study 2: 37 females, *M* = 19.93 years, *SD* = 2.45; study 3: 40 females, *M* = 19.07 years, *SD* = 1.80). After giving informed consent, participants completed the experiment and received either an inconvenience allowance of £3 or 0.5 credits towards their first-year credit requirement.

#### 2.3.1.2 Design and Materials

The design and materials used in Experiment 3 were the same as in Experiment 2 (see section 2.2.1.2).

#### 2.3.1.3 Procedure

Study and test procedures were very similar to Experiment 2. The only changes in Experiment 3 were the response possibilities that were set according to the R/K procedure combined with confidence judgements. Participants were asked to respond according to a 5-point rating scale (Yonelinas et al., 2005) by using the keys 1 to 5 on a keyboard. They were instructed to judge whether each word was old or new, and to make one of the 5 responses. They were instructed to respond Remember (5) when they were able to remember specific details about the word at study, such as remembering the position of the word in the study list or what they thought at the time of presentation (see Appendix J for the specific instructions given to participants, based on the instructions used in Rajaram (1993)). For words they believed to be old, but for which they could not recover specific details, they were instructed to respond Know, and to indicate their confidence in their Know judgement (*4 = sure old, 3 = maybe old*). They were also asked to separate their new judgements by confidence (*1 = sure new, 2 = maybe new*).

### 2.3.2 Results

Table 2.10 displays mean probabilities of hits and false alarms separated by instruction (remember, forget) and valence (negative, positive, neutral) for studies 1, 2 and 3 in Experiment 3 (see Appendix I for proportions of responses for each confidence rating). Hits are the sum of Remember and Know judgements for old words. False alarms are the same sum of incorrect responses for new words. Measures of sensitivity and criterion for studies 1, 2 and 3 are presented in Table 2.11. *d'* (sensitivity) and c (criterion), were calculated as was done in Experiments 1 and 2. Familiarity was measured based on the independence assumption proposed by Yonelinas and Jacoby (1995). The estimates of recollection (R) were taken to be the probabilities of Remember judgements. The estimates of familiarity (F) were taken to be the probabilities of Know judgements collapsed across confidence using the equation described in the introduction to Experiment 3:

$$F = \frac{K}{1 - R}$$

Table 2.10

*Probabilities of Correct Judgements (Hits) to Old Items and Incorrect Judgements (False Alarms) to New Items across Instruction (Remember/Forget) and Valence (Negative, Positive and Neutral) for Studies 1, 2 and 3. SD = standard deviation*

| | | Study 1 | | Study 2 | | Study 3 | |
|---|---|---|---|---|---|---|---|
| | | Negative M (SD) | Neutral M (SD) | Negative M (SD) | Positive M (SD) | Positive M (SD) | Neutral M (SD) |
| Hits | Remember | 0.73 (0.12) | 0.67 (0.15) | 0.74 (0.12) | 0.70 (0.11) | 0.69 (0.15) | 0.65 (0.16) |
| | Forget | 0.63 (0.15) | 0.55 (0.19) | 0.67 (0.14) | 0.64 (0.16) | 0.54 (0.16) | 0.51 (0.16) |
| False alarms | | 0.28 (0.14) | 0.19 (0.12) | 0.23 (0.12) | 0.22 (0.11) | 0.25 (0.13) | 0.20 (0.13) |

Table 2.11

*Summary Statistics for Each Valence across Instructions for Studies 1, 2 and 3. R is the Estimate of Recollection Driven by Remember Judgements, F is the Estimate of Familiarity Driven by Know Judgements (according to the Independence Assumption), d' is the Estimate of Memory Sensitivity and c is the Estimate of Criterion. SD = standard deviation*

| | | Study 1 | | Study 2 | | Study 3 | |
|---|---|---|---|---|---|---|---|
| | | Negative M (SD) | Neutral M (SD) | Negative M (SD) | Positive M (SD) | Positive M (SD) | Neutral M (SD) |
| R | Remember | 0.30 (0.20) | 0.27 (0.18) | 0.30 (0.17) | 0.27 (0.16) | 0.31 (0.18) | 0.28 (0.15) |
| | Forget | 0.18 (0.17) | 0.17 (0.15) | 0.20 (0.16) | 0.20 (0.17) | 0.15 (0.10) | 0.15 (0.11) |
| F | Remember | 0.59 (0.18) | 0.54 (0.16) | 0.63 (0.14) | 0.59 (0.13) | 0.55 (0.19) | 0.51 (0.20) |
| | Forget | 0.54 (0.16) | 0.46 (0.18) | 0.60 (0.15) | 0.56 (0.15) | 0.47 (0.18) | 0.42 (0.17) |
| *d'* | Remember | 1.24 (0.47) | 1.34 (0.51) | 1.52 (0.55) | 1.43 (0.50) | 1.30 (0.50) | 1.42 (0.57) |
| | Forget | 0.94 (0.41) | 1.00 (0.54) | 1.30 (0.45) | 1.24 (0.53) | 0.87 (0.36) | 0.99 (0.37) |
| c | Remember | -0.02 (0.33) | 0.22 (0.35) | 0.06 (0.32) | 0.14 (0.32) | 0.09 (0.39) | 0.27 (0.41) |
| | Forget | 0.12 (0.38) | 0.39 (0.41) | 0.17 (0.38) | 0.23 (0.36) | 0.31 (0.39) | 0.48 (0.46) |

For each study, 2 x 2 repeated measures ANOVAs were conducted on the four measures shown in Table 2.11: sensitivity (*d'*), response criterion (c) and estimates of recollection (R) and familiarity (F). In each case these analyses initially included the factors of instruction (remember vs forget) and valence (two categories in each case). ROC curves for each study are attached in Appendix K.

### 2.3.2.1 Effects of Instruction

There were main effects of instruction in all three studies for all measures. In the cases of sensitivity, recollection and familiarity this reflects a directed forgetting effect: *d'*, R and F were all superior for TBR than for TBF words (although see the interaction term for R in study 2 reported below; see Table 2.12). The main effect of instruction for response criterion reflects a relatively more liberal criterion for TBR than for TBF words.

Table 2.12

*Summary of Repeated Measures ANOVA Results for Comparisons between Instruction and Valence in Studies 1, 2 and 3*

| | | Study 1 | | | | Study 2 | | | | Study 3 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Measure (DV) | | *df* | *F* | *p* | $\eta_p^2$ | *df* | *F* | *p* | $\eta_p^2$ | *df* | *F* | *p* | $\eta_p^2$ |
| Instruction | R | 1, 43 | 28.26 | < .001*** | .40 | 1, 43 | 19.62 | < .001*** | .31 | 1, 43 | 37.32 | < .001*** | .47 |
| | F | 1, 43 | 13.22 | .001** | .24 | 1, 43 | 6.72 | .013* | .14 | 1, 43 | 20.05 | < .001*** | .32 |
| | *d'* | 1, 43 | 35.40 | < .001*** | .45 | 1, 43 | 18.27 | < .001*** | .30 | 1, 43 | 37.50 | < .001*** | .47 |
| | c | 1, 43 | 35.40 | < .001*** | .45 | 1, 43 | 18.27 | < .001*** | .30 | 1, 43 | 37.50 | < .001*** | .47 |
| Valence | R | 1, 43 | 3.95 | .053 | .09 | 1, 43 | 4.45 | .041* | .09 | 1, 43 | 3.50 | .068 | .08 |
| | F | 1, 43 | 15.94 | < .001*** | .27 | 1, 43 | 6.18 | .017* | .13 | 1, 43 | 10.75 | .002** | .20 |
| | *d'* | 1, 43 | 2.49 | .122 | .06 | 1, 43 | 4.13 | .048* | .09 | 1, 43 | 7.38 | .009** | .15 |
| | c | 1, 43 | 94.85 | < .001*** | .69 | 1, 43 | 4.45 | .041* | .09 | 1, 43 | 56.46 | < .001*** | .57 |
| Instruction x Valence Interactions | R | 1, 43 | 1.07 | .307 | .02 | 1, 43 | 4.09 | .049* | .09 | 1, 43 | 2.39 | .129 | .05 |
| | F | 1, 43 | 0.33 | .567 | .01 | 1, 43 | 0.00 | .948 | .00 | 1, 43 | 0.03 | .867 | .00 |
| | *d'* | 1, 43 | 0.24 | .625 | .01 | 1, 43 | 0.46 | .502 | .01 | 1, 43 | 0.09 | .768 | .00 |
| | c | 1, 43 | 0.24 | .625 | .01 | 1, 43 | 0.46 | .502 | .01 | 1, 43 | 0.09 | .768 | .00 |

*Notes.* * $p < .05$, ** $p < .01$, *** $p < .001$.
df = degrees of freedom

### 2.3.2.2 Effects of Valence

There were main effects of valence across all studies for all four measures and an interaction between valence and instruction for recollection in study 2. Sensitivity was superior for negative words relative to positive words (study 2) and for neutral words relative to positive words (study 3; see Table 2.12). There was no effect of valence in study 1 – the negative/neutral contrast, although the absolute magnitude of the differences (neutral > negative) was similar to the differences in the other two studies.

Turning to recollection, in study 2 recollection was superior for negative relative to positive words and there was also an interaction term. Recollection was superior for negative words in the remember condition and there was no effect of emotion in the forget condition. Furthermore, there was an effect of emotion for familiarity across all studies. Familiarity was superior for emotional relative to neutral words and for negative relative to positive words (see Table 2.12).

For response criterion, there was a more liberal response criterion for emotional words relative to neutral words and a more liberal criterion for negative words relative to positive words (see Table 2.12).

#### 2.3.2.3 ROC Analyses – Area Under the Curve (AUC)

Additional analyses of the ROC curves were conducted for the Area Under the Curve (AUC) measures to examine possible effects of valence and instruction. The AUC was measured using the ROC Toolbox in Matlab (Koen et al., 2017). The results (see Table 2.13 for the ANOVA results) are similar to the results from the *d'* analyses. That is, there are consistent DF effects across all three studies, which are not affected by valence. In study 2, there was also a significant main effect of valence. Memory was higher for negative ($M$ = 0.79, $SD$ = 0.01) than for positive ($M$ = 0.78, $SD$ = 0.01) items. Although the *d'* analyses resulted in a main effect of valence in study 3 as well, this was not the case in the AUC analyses.

Table 2.13
*Summary of Repeated Measures ANOVA Results for Comparisons between Instruction and Valence in Studies 1, 2 and 3 on the AUC measures*

| Measure (DV) | Study 1 | | | | Study 2 | | | | Study 3 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | *df* | *F* | *p* | $\eta_p^2$ | *df* | *F* | *p* | $\eta_p^2$ | *df* | *F* | *p* | $\eta_p^2$ |
| Instruction | 1, 43 | 41.11 | <.000*** | .49 | 1, 43 | 24.21 | <.000*** | .36 | 1, 43 | 43.23 | <.000*** | .50 |
| Valence | 1, 43 | 0.10 | .750 | .00 | 1, 43 | 7.98 | .007** | .16 | 1, 43 | 3.05 | .088 | .07 |
| Instruction x Valence Interactions | 1, 43 | 0.20 | .660 | .01 | 1, 43 | 1.08 | .306 | .02 | 1, 43 | 1.53 | .224 | .03 |

*Notes.* ** $p < .01$, *** $p < .001$.
df = degrees of freedom

### 2.3.3 Discussion

One intention in this experiment was to investigate the generality of the findings in Experiment 2, where ROC measures were used to extract estimates of recollection and familiarity. This was done in Experiment 3 by introducing the Remember/Know procedure in combination with confidence judgements. The confidence judgements were applied in order to have an experimental design close to Experiment 2 that allows a direct comparison of both methods. The experiment also allowed for assessment of criterion and the important separation between sensitivity and criterion.

#### 2.3.3.1 Memory Sensitivity and Directed Forgetting

Consistent with the findings in Experiment 2, there was a reliable DF effect for all measures, the DF effect did not vary according to emotion, nor did emotion enhance memory sensitivity. Moreover, memory sensitivity was superior for negative relative to positive words. Recollection as well as familiarity were superior for TBR words relative to TBF words in all three studies. This is partially in line with previous findings in an ERP study investigating the neural correlates of recollection and familiarity in directed forgetting. Nowicka et al. (2009)

reported a left-parietal old/new effect for TBR items, suggesting that remembering TBR items was based on recollection processes. There was no left-parietal old/new effect for correctly recognized TBF words, and the outcomes of this ERP study did not speak to putative contributions via familiarity.

For memory sensitivity, the DF effect observed here is consistent with common findings in the directed forgetting paradigm (MacLeod, 1999). The absence of changes in the DF effect with emotion is consistent with some previous studies (Berger et al., 2018; Gallant et al., 2018; Taylor et al., 2018). This consistency, however, does not extend to other studies that have found a smaller DF effect for emotional material (Bailey & Chapman, 2012; Brandt et al., 2013; Hauswald et al., 2011; Minnema & Knowlton, 2008; Otani et al., 2012; Yang et al., 2016), or a smaller DF effect for neutral material (Brandt et al., 2013). Moreover, the observation that emotion did not enhance memory sensitivity is in contrast with the assumption that emotion enhances memory (Kensinger & Corkin, 2003; Levine & Edelstein, 2009; Talmi et al., 2008), as well as findings in some DF studies (Brandt et al., 2013; Gallant & Yang, 2014; Payne & Corrigan, 2007; Taylor et al., 2018). In contrast with Experiment 2, memory sensitivity was superior for neutral words relative to positive words, which is consistent with some previous findings of studies investigating directed forgetting for emotional and neutral material (Bailey & Chapman, 2012; Berger et al., 2018), and other studies (Maratos et al., 2000). This effect of emotion between neutral and positive words was not, however, evident in the AUC analyses. In summary, and perhaps inevitably given the heterogeneity in the existing literature, the data in Experiment 3 are aligned with some prior data but not others.

The superior sensitivity for negative relative to positive material in study 2 suggests that negative and positive material have different effects on memory. Based on this finding, the effect of emotion on memory here cannot be linked to properties of arousal, rather this may be linked with properties of valence (Dolcos et al., 2004; Kensinger & Corkin, 2003, 2004;

Kensinger & Kark, 2018). The consistency in this case with the findings in Experiment 2 reinforces this argument. Another consistency between the present experiment and Experiment 2, both using multiple points measures, is that emotion did not enhance memory overall, emphasizing that response criterion can create illusory indications of changes in memory when single-point measures are used (Dougal & Rotello, 2007; Kapucu et al., 2008).

In summary, the findings observed here are in line with some of the predictions made in this experiment but not others. Briefly, emotion did not enhance sensitivity, the DF effect was not influenced by emotion and a DF effect was observed for recollection and familiarity. In contrast, however, sensitivity was superior for neutral words relative to positive words. The consistent DF effect across the three behavioural experiments described above is further evidence of a robust pattern. Moreover, the finding that the DF effect did not differ according to valence is also a consistent outcome. So far, these outcomes suggest that factors not controlled for in prior DF studies have resulted in the published inconsistencies.

#### 2.3.3.2 Response Criterion

Response criterion was more liberal for emotional words relative to neutral and for negative words relative to positive words. This is in line with findings from Experiment 2. These findings are inconsistent, however, with the findings in Experiment 1 where there was a more conservative response criterion for emotional words. As discussed in Experiment 2, the number of valence conditions used and the use of single-point vs multiple-point measures may provide an explanation for these discrepancies. The consistency of the findings across Experiments 2 and 3, in designs arguably better suited to assess the measures of interest than those used in Experiment 1, increases the degree of confidence that criterion change with emotion is a robust outcome.

The findings discussed here are consistent with previous DF and other studies, where participants adopted a more liberal criterion for emotional material (Bailey & Chapman, 2012; Dougal & Rotello, 2007; Hauswald et al., 2011; Kapucu et al., 2008; Marchewka et al., 2016; Windmann & Kutas, 2001). In these studies, however, there were no direct paired contrasts between negative and positive material. The criterion changes here and in Experiment 2 for this paired contrast are thus a relatively novel outcome, and these outcomes cannot be attributed to differences in arousal, either, given that this was matched for these two valence categories.

#### 2.3.3.3 Recollection and Familiarity

In study 2, recollection was superior for negative relative to positive TBR words, whereas emotion did not influence recollection for TBF words. This interaction illustrates that forgetting negative and positive words was equally challenging whereas remembering negative words was easier than remembering positive words. Previous studies of emotional directed forgetting reporting an interaction between directed forgetting and emotion involved only negative and neutral material, or three valence categories, so in no previous cases was the negative/positive contrast isolated (Bailey & Chapman, 2012; Hauswald et al., 2011; Nowicka et al., 2011; Payne & Corrigan, 2007). These studies suggest that emotional information may be particularly resistant to forgetting compared to neutral items, whereas the findings in the present experiment reveal important differences within the emotional category (negative vs positive). Zheng et al. (2018) found a larger left-parietal old/new effect for negative relative to positive and neutral images, which may indicate that memory judgements were to a greater extent based on recollection for negative images relative to positive images, but this was not mirrored in the behavioural outcomes in their study.

In this experiment there were no differences in recollection estimates between emotional and neutral words, which contrasts with the findings from Experiment 2 and some previous findings (Dougal & Rotello, 2007; Ochsner, 2000; Ritchey et al., 2008). In Experiment 2 recollection was superior for neutral relative to negative words. Linking the present findings with the existing literature, behavioural R/K studies have reported enhanced recollection for emotional material based on increases in Remember judgements (Ochsner, 2000; Sharot et al., 2007).

Familiarity estimates in Experiment 3 were increased for emotional relative to neutral words. This is also inconsistent with the findings from Experiment 2, but consistent with Ochsner (2000) who found, at least for negative images, increased Know judgements relative to neutral images. They also found increased Know judgements for negative relative to positive images, which is consistent with Experiment 2 as well as the present findings. Perhaps the strongest conclusion that this pattern of outcomes permits is that it adds further weight to the view that the effects of emotion on memory cannot be reduced to the influence of arousal.

The DF effect again observed for recollection as well as familiarity confirms the revised prediction that was made based on the findings in Experiment 2 (prior to that experiment, DF was expected to influence recollection only). One explanation for this observation is that directed forgetting is a consequence of enhanced processing of TBR items that influences both recollection and familiarity. This account is in line with findings that there are some experimental manipulations that increase both recollection and familiarity (Yonelinas & Jacoby, 1995).

With respect to the effects of emotion on recollection and familiarity, there are some inconsistencies as described above. Also, the findings here are in contrast to what has been predicted based on Experiment 2. To repeat, the key differences are: (i) superior recollection for neutral compared to negative words in Experiment 2 and superior recollection only for negative compared to positive words in Experiment3 and (ii) familiarity was superior for emotional words compared to neutral words in Experiment 3, whereas this effect was not evident in Experiment 2. One explanation for these inconsistencies is that different measures were used in both experiments. This did not, however, result in different observations for response criterion and directed forgetting. In addition, both experiments used a different set of participants, so individual differences cannot be ruled out. While recognising that a definitive explanation is unlikely, a closer inspection of the recollection and familiarity outcomes using two different methods (i.e. the ROC measure and R/K method of measuring recollection and familiarity) was undertaken.

#### 2.3.3.4 Comparing ROC and R/K Analyses

As mentioned earlier, one intention in Experiment 3 was to investigate the generality of the findings in Experiment 2 (ROCs) using a slightly different method, the R/K procedure. This was also done by directly comparing the outcomes using the R/K method in Experiment 2 and the ROC measure in Experiment 3. If high confidence old judgements are based on recollection only, and this is also true for Remember judgements, then ROCs in the two cases should overlap. If, however, the highest level of confidence judgements can also be made on the basis of familiarity, then the intercept of the ROC will be lower in the R/K procedure. In this experiment, a combination of the R/K procedure and confidence judgements was applied, and the question to be considered is how ROCs plotted in this design should relate to ROCs plotted on the basis of confidence judgements, as in Experiment 2. Previously, studies have investigated precisely this question and have found that estimates of recollection and familiarity derived from the R/K procedure correspond with the ones derived from ROC measures (Yonelinas et al., 1996; for a review of findings see Yonelinas, 2001a). In addition, R/K measures of recollection and familiarity generally fit with the predicted ROCs according to the DPSD model (Parks & Yonelinas, 2007; Yonelinas et al., 1996). A comparison of the ROC figures between Experiment 2 and 3 suggests that the ROC curves result in similar patterns (see Appendix K for the ROC curves in Experiment 3). In the negative/neutral contrast there is an overlap of the curves indicating no differences in sensitivity, whereas in the negative/positive contrast the two curves do not overlap indicating differences in sensitivity. These observations are identical to the observations in Experiment 2 (see Figure 2.2). Similarly, the ROC curves in Experiment 3 indicate a more liberal criterion for emotional words compared to neutral words and for negative words compared to positive words. This is very similar indeed to the observations in Experiment 2.

For a direct comparison of the ROC and R/K method within each experiment, the R/K method in Experiment 2 was applied by taking the highest confidence old judgements to old items as the proportion of Remember responses, as the estimate of recollection. The other two confidence judgements for old items (probably old and maybe old) were considered as Know responses and used to calculate familiarity using the independence R/K method. In Experiment 3, the 5-point confidence scale was used to derive 5-point ROC curves (instead of 6-point ROC curves in Experiment 2) and estimates of recollection and familiarity were based on the DPSD model. Remember responses were used as the highest confidence old judgement that reflected recollection (i.e. the left most point on the ROC curve). The remainder of the confidence scale (4 points) were used for the remainder of the ROC curve. The estimates of recollection and familiarity measured based on the ROC measures and the R/K procedure in Experiment 2 and Experiment 3 for all three studies can be found in Table 2.14 (Experiment 2) and Table 2.15 (Experiment 3).

Table 2.14
*Measures of Recollection (R) and Familiarity (F) Calculated using the R/K Approach and the ROC Analyses across Instructions (Remember/Forget) and Valence (Negative, Positive and Neutral) in Studies 1, 2 and 3 in Experiment 2. SD = standard deviation*

| | | | Study 1 | | Study 2 | | Study 3 | |
|---|---|---|---|---|---|---|---|---|
| | | | Negative M (SD) | Neutral M (SD) | Negative M (SD) | Positive M (SD) | Positive M (SD) | Neutral M (SD) |
| ROC | R | Remember | 0.26 (0.22) | 0.29 (0.19) | 0.23 (0.16) | 0.21 (0.15) | 0.27 (0.18) | 0.26 (0.16) |
| | | Forget | 0.20 (0.21) | 0.25 (0.19) | 0.17 (0.13) | 0.15 (0.14) | 0.19 (0.17) | 0.18 (0.17) |
| | F | Remember | 1.08 (0.39) | 1.09 (0.52) | 0.93 (0.30) | 0.84 (0.38) | 0.86 (0.34) | 0.92 (0.41) |
| | | Forget | 0.92 (0.40) | 0.91 (0.37) | 0.81 (0.33) | 0.73 (0.31) | 0.74 (0.34) | 0.74 (0.38) |
| R/K | R | Remember | 0.46 (0.18) | 0.41 (0.18) | 0.46 (0.16) | 0.43 (0.16) | 0.38 (0.20) | 0.37 (0.17) |
| | | Forget | 0.37 (0.21) | 0.33 (0.19) | 0.38 (0.17) | 0.34 (0.18) | 0.30 (0.20) | 0.26 (0.18) |
| | F | Remember | 0.51 (0.20) | 0.45 (0.16) | 0.57 (0.17) | 0.53 (0.19) | 0.43 (0.14) | 0.43 (0.16) |
| | | Forget | 0.48 (0.17) | 0.39 (0.15) | 0.51 (0.17) | 0.48 (0.16) | 0.42 (0.17) | 0.37 (0.16) |

Table 2.15
*Measures of Recollection (R) and Familiarity (F) Calculated using the R/K Approach and the ROC Analyses across Instructions (Remember/Forget) and Valence (Negative, Positive and Neutral) in Studies 1, 2 and 3 in Experiment 3. SD = standard deviation*

| | | | Study 1 | | Study 2 | | Study 3 | |
|---|---|---|---|---|---|---|---|---|
| | | | Negative M (SD) | Neutral M (SD) | Negative M (SD) | Positive M (SD) | Positive M (SD) | Neutral M (SD) |
| ROC | R | Remember | 0.17 (0.21) | 0.22 (0.18) | 0.20 (0.15) | 0.18 (0.14) | 0.18 (0.18) | 0.21 (0.17) |
| | | Forget | 0.08 (0.13) | 0.12 (0.14) | 0.12 (0.13) | 0.12 (0.15) | 0.09 (0.09) | 0.09 (0.09) |
| | F | Remember | 0.87 (0.90) | 1.05 (0.43) | 1.27 (0.52) | 1.18 (0.52) | 1.02 (0.51) | 0.92 (1.11) |
| | | Forget | 0.79 (0.42) | 0.79 (0.46) | 1.11 (0.42) | 1.06 (0.44) | 0.46 (1.46) | 0.78 (0.36) |
| R/K | R | Remember | 0.29 (0.20) | 0.26 (0.19) | 0.30 (0.17) | 0.27 (0.16) | 0.30 (0.19) | 0.27 (0.16) |
| | | Forget | 0.17 (0.18) | 0.16 (0.15) | 0.20 (0.16) | 0.20 (0.17) | 0.14 (0.10) | 0.14 (0.11) |
| | F | Remember | 0.58 (0.19) | 0.53 (0.17) | 0.63 (0.14) | 0.59 (0.13) | 0.55 (0.19) | 0.51 (0.20) |
| | | Forget | 0.53 (0.17) | 0.45 (0.18) | 0.60 (0.15) | 0.56 (0.15) | 0.47 (0.18) | 0.42 (0.17) |

Inspection of these data points suggests that the measures correspond and indicate similar effects of emotion on the processes of recollection only in study 2 (negative vs positive) in both experiments. For example, in Experiment 3 recollection and familiarity are increased for negative relative to positive words using both methods. Moreover, broadly equivalent recollection in Experiment 3 for negative and positive words in the forget condition is also visible in both methods (see Table 2.15). This is, however, not the case in study 1 and 3, where both measures estimate different effects of emotion on recollection and familiarity. In study 1 of Experiment 2, for example, recollection was superior for neutral TBR words relative to negative TBR words under the ROC method, whereas the opposite is the case under the R/K method. These outcomes are at least suggestive of the conclusion that one or more of the assumptions underlying these methods does not hold.

The results of these comparisons bear some similarities with those in other studies reporting discrepancies when using both methods (e.g. Rotello, Macmillan, Reeder, & Wong, 2005). For example, Rotello et al. (2005) employed a combination of confidence judgements and R/K judgements and found participants' Remember judgements were accompanied with various levels of confidence. Thus, Remember judgements were not a pure measure of recollection (also see Rotello et al., 2006). Wixted (2007) has argued that the processes of recollection and familiarity are not independent and that Remember responses or high confidence judgements may not reflect only recollection. By this view, while the two processes are considered to exist, each response on a memory test is a reflection of a combined contribution from the two processes, albeit to differing extents (for a detailed review, see Wixted, 2007). The data in Experiments 2 and 3 cannot adjudicate between different models of the relationships between recollection and familiarity. They do, none the less, align themselves with findings suggesting that there is a less than 1:1 correspondence between their accuracy in estimating process contributions.

# CHAPTER 3

# 3. EXAMINING EMOTION IN DIRECTED FORGETTING USING ELECTROPHYSIOLOGICAL MEASURES (ERPs)

## 3.1 Experiment 4 – Emotion and Directed Forgetting using ERPs

Introducing confidence judgements in Experiment 2 allowed for a separation between estimates of memory sensitivity and response criterion that was not possible in Experiment 1, where only old/new judgements were required at test. One of the reasons for conducting Experiments 2 and 3 was to understand the data that is obtained when single-point measures have been employed. As described in section 2.2, an issue with using single-point measures is that memory sensitivity and response criterion are not necessarily independent. Because of this, conclusions about whether emotion affects memory sensitivity, response criterion or both can be made only tentatively when differences in criterion occur across critical conditions.

This potential confound motivated the use of confidence judgements in Experiment 2 and 3, because they are one way of providing separate estimates of the contributions of sensitivity and criterion. The outcomes of these experiments shed light on the ways in which sensitivity and criterion interact with emotion (see section 2.2). Additionally, the use of confidence judgements in Experiment 2 permitted an investigation of how the processes of recollection and familiarity contribute to memory for emotional and neutral material, and how they are influenced by instructions to remember and forget. This was also accomplished via the use of the Remember/Know procedure (combined with confidence judgments) in Experiment 3. Across these three experiments the main findings were: (i) a consistent liberal criterion for emotional compared to neutral material in Experiment 2 and 3, and (ii) no differences between the directed forgetting effect for emotional and neutral material. In Experiment 1 there was a more liberal criterion for neutral material, which is inconsistent with others finding the opposite using a variety of methods (Bailey & Chapman, 2012; Dougal & Rotello, 2007; Hauswald et al., 2011; Marchewka et al., 2016), as was found in Experiment 2 and 3.

If, however, changes to the memory response requirements at test change markedly how participants approach tasks – for example, by relying to a greater extent on recollection when confidence or R/K judgements are made – then the findings in Experiment 2 and 3 might not be reasonably compared to those in Experiment 1, where only old/new recognition memory judgements were required. For this reason, in this chapter the experiment that is reported requires old/new judgements only at test, and the contributions of recollection and familiarity to memory for emotionally valenced material as well as the influence of directions to remember and forget are assessed via a neural measure (event-related potentials: ERPs). The rationale for using ERPs is to provide a means of assessing the extent to which recollection and familiarity are engaged in a directed forgetting task requiring only old/new decisions.

Previous studies using event-related potentials (ERPs) have identified dissociable neural correlates that are linked with recollection and familiarity (Rugg & Curran, 2007; Rugg et al., 1998; Vilberg & Rugg, 2008; for a review see Friedman & Johnson, 2000). The left-parietal old/new effect, which onsets around 400-500ms post-stimulus has a left-sided posterior maximum and lasts for 400-500ms, has been linked with recollection (Rugg & Curran, 2007; Sanquist et al., 1980). This effect is larger for Remember than for Know judgements in the R/K procedure (Duarte, Ranganath, Winward, Hayward, & Knight, 2004; Friedman & Johnson, 2000; Smith, 1993), and when source judgements are correct rather than when they are incorrect (Leynes & Phillips, 2008; Senkfor & Van Petten, 1998; Wilding, 2000; Wilding et al., 1995; Wilding & Rugg, 1996). The process of familiarity has been linked with the mid-frontal old/new effect (which is also referred to as the FN400; Curran, 1999; Rugg et al., 1998). This effect, is observed in a time window of 300-500ms with a fronto-central maximum (Rugg et al., 1998; and for reviews and discussions see Curran, 2000; Paller et al., 2007).

ERPs have also been used to examine how recollection and familiarity are influenced by emotion. Maratos et al. (2000) investigated memory for negative and neutral words. Participants were instructed to study a set of words while rating the words on emotionality. During a subsequent recognition memory test, Maratos et al. (2000) observed reduced memory sensitivity for negative words. Electrophysiologically, there were no differences in the magnitude of the mid-frontal old/new effects according to emotion. The left-parietal ERP old/new effect, however, was smaller for negative words. Based on these findings, they suggested that emotion influences memory based on recollection and not familiarity (Maratos et al., 2000). The opposite ERP outcome was, however, obtained in a study by Zheng et al. (2018). Behaviourally, using negative, positive and neutral images, they found no differences in memory sensitivity. They did, however, observe a larger left-parietal old/new effect for negative images relative to positive and neutral images. They concluded that, for negative images, increased recollection-based memory judgements were made, although it remains problematic that this was not reflected in the behavioural data. In another study that investigated memory for negative and neutral words, there were no differences in memory sensitivity (Windmann & Kutas, 2001). Moreover, in contrast to the ERP findings described above, Windmann and Kutas (2001) found no effect of emotion on the magnitude of the old/new effects in mid-frontal and left-parietal regions. Therefore, they concluded that emotion did not influence recollection or familiarity, but again the absence of a behavioural effect makes interpretation of the ERP old/new effects somewhat difficult.

Researchers have also investigated how old/new effects vary for TBR and TBF items. Nowicka et al. (2009) found a reliable old/new effect for correctly recognized TBR items over central and parietal sites, whereas this old/new effect was absent for correctly recognized TBF items. According to Nowicka et al. (2009), these findings suggest that recollective processes support TBR items more than TBF items. Van Hooff, Whitaker and Ford (2009), also examined

ERPs for TBR and TBF test words and found a similar pattern. They also found a larger left-parietal old/new effect for TBR items relative to TBF items.

To re-iterate, the rationale for using ERP old/new effects is to provide a means of investigating memory and emotion that complements the ROC and R/K measures. In this experiment the focus is on conducting targeted analyses at mid-frontal and left-parietal sites over the 300-500ms and 500-800ms time windows, in order to isolate activity which is assumed to reflect the processes of recollection and familiarity (Rugg & Curran, 2007; Rugg et al., 1998; Sanquist et al., 1980; Wilding & Rugg, 1996). Two studies were conducted in which memory effects were investigated between negative and neutral words (study 1) and between positive and neutral words (study 2). The reason for separating negative and positive material relative to neutral material in different studies was to increase power and maintain sufficient trials in key conditions for the ERP data analyses. If the ways in which recollection and familiarity are influenced by emotion and directed forgetting are not dependent upon response demands at test, then the ERP modulations will behave in a way that is comparable to how the estimates for recollection and familiarity varied with emotion and directed forgetting instructions in Experiments 2 and 3. That is, both the left-parietal and mid-frontal old/new effects will be larger for TBR than for TBF items.

### 3.1.1 Methods

#### 3.1.1.1 Participants

Participants in both studies were undergraduates from the University of Nottingham. Participants with sensitivity (*d'*) scores $< 0.24$ for recognition memory and/or $<16$ trials in the critical response categories (see below for definitions) in the ERP data were excluded. In addition, participants for whom more than 25% of trials were rejected due to noise were also excluded. In study 1 there were 20 participants of which 4 participants were excluded for having fewer than 16 trials in at least one critical response category. In study 2 there were 21 participants of which 5 participants were excluded; one participant did not meet the inclusion criterion for sensitivity, one was excluded because of extensive noise in the EEG data and three were excluded for having fewer than 16 trials in one or more of the critical response categories. In both studies data from 16 remaining participants were used for behavioural and EEG analyses (study 1: 13 females, $M$ = 23 years, $SD$ = 3.73; study 2: 11 females, $M$ = 19.81 years, $SD$ = 1.87). After giving informed consent, participants completed the experiment and received either an inconvenience allowance of £10 or 2 credits towards their first-year credit requirement.

#### 3.1.1.2 Design and Materials

The two studies had very similar designs and procedures and differed only in the categories of valence used. Study 1 consisted of negative and neutral words, study 2 positive and neutral words.

A 2 (instruction: remember vs forget) x 2 (valence) within-subjects design was employed. Each valence category consisted of 240 words which were selected from the Warriner et al. (2013) database. The valence categories in both studies differed on the basis of valence (negative: $M$ = 2.70, $SD$ = 0.67; positive: $M$ = 7.28, $SD$ = 0.48; neutral: $M$ = 5.12, $SD$ = 0.38). All words were matched for semantic relatedness, word length and word frequency. At the level of the entire word sets independent sample t-tests were conducted to assess the equivalence of the semantic relatedness scores (study 1: range 0.79 – 0.95; study 2: range 0.79 – 0.98), arousal (study 1: range 2.25 – 7.27; study 2: range 2.50 – 7.05), word length (both studies: range 3 – 14) and word frequencies (both studies: range 0.30 – 5.40). The only significant differences were in arousal. Full details of these outcomes can be seen in Appendix L (Table L1).

The 240 words were then split into study (120 words) and test (240 words) lists. Six study lists were formed, each containing 20 words for each valence type (40 words in total). The corresponding test lists each contained 80 words in total (40 words for each valence category, half of which were also on the study lists). A further set of lists was created in which the words designated as either TBR or TBF were alternated. DF instructions (TBR and TBF) and the old/new status of words were counterbalanced. Again, independent sample t-tests were conducted separately on the study and test lists for the list of properties described above. The outcomes are consistent with those already described (see Appendix L, Table L2 (study lists) and Table L3 (test lists)).

#### 3.1.1.3 EEG Acquisition

EEG was recorded at a sample rate of 1024 Hz using a Biosemi ActiveTwo system (Amsterdam, Netherlands) with 64 active electrodes placed according to the 10-20 system (Jasper, 1958, see Figure 1.4B). Two additional electrodes located at mid-parietal sites, the Common Mode Sense (CMS) and the Driven Right Leg (DRL), were used for online referencing (see Figure 1.4B). Additional electrodes were placed on the left and right mastoids for offline re-referencing. Electrooculogram (EOG) was recorded using two electrodes above and below the right eye (VEOG) and from the outer canthi (HEOG) of both eyes for monitoring eye blinks and lateral eye movements.

Offline pre-processing of the EEG data was done via the Brain Products Analyzer software. The data were filtered using a high-pass filter of 0.1 Hz and a low-pass filter of 30 Hz. Trials containing large EOG artefacts (eye blinks and movements) were rejected using the Ocular Correction ICA function in Analyzer, followed by a visual inspection. The data were then re-referenced using the average of the signal at the two mastoids and segmented into epochs of 1200ms time locked to stimulus presentation (including a 200ms pre-stimulus baseline). A semi-automated procedure was used for the rejection of large artefacts exceeding ±100 μV, which were identified initially via an algorithm and followed by visual inspection. The critical response categories were: correct old judgements (hits) to remember and forget words for negative, positive and neutral words and correct new responses (correct rejections to new negative, positive and neutral words). See Table 3.1 for mean numbers of trials contributing to each of the response categories in studies 1 and 2. An average of 87.6% of trials in study 1 and 87.7% of trials in study 2 per participant was retained for analyses after pre-processing.

Table 3.1
*Mean Numbers of Trials used for Averaging separated by Response Category (Correct Rejections and Hits separated by Instruction (Remember/Forget)) and Valence (Negative, Positive and Neutral). SD = standard deviation*

| | Study 1 | | Study 2 | |
|---|---|---|---|---|
| | Negative<br>M (SD) | Neutral<br>M (SD) | Positive<br>M (SD) | Neutral<br>M (SD) |
| Hit (Remember) | 40 (6.92) | 37 (7.75) | 36 (9.04) | 32 (8.29) |
| Hit (Forget) | 32 (9.53) | 27 (9.42) | 31 (8.99) | 28 (9.57) |
| Correct Rejection | 85 (19.95) | 89 (22.22) | 80 (11.47) | 84 (15.78) |

#### 3.1.1.4 Procedure

Participants were tested individually, and the study lasted no more than 50 minutes. The entire time spent with participants was up to 2 hours. The remaining 70 minutes was occupied by the application of the cap and electrodes, and removal and debrief after the end of the study.

The experiment consisted of six study-test blocks. In the study phase, a fixation cross appeared on the screen for 1000ms followed by a 200ms blank screen. Words were then presented individually for 1500ms in the centre of the screen. After each word a blank screen appeared for 500ms followed by a remember cue (VVVVV in the colour green) or a forget cue (XXXXX in the colour red). Participants were instructed to attempt to remember the preceding word following a remember cue, and to forget the word when a forget cue followed. The instructions remained on the screen for 500ms and the order of remember and forget words was determined randomly for each participant. Following the offset of the cue, a blank screen appeared for a varying duration of between 700 and 1300ms, before the next trial started.

Test trials commenced with a blank screen for 1000ms, followed by a fixation cross for another 1000ms. After the fixation cross another blank screen appeared for 200ms before a word was presented for 300ms. Participants were asked to make an old/new recognition judgment on the word, regardless of the TBR or TBF instruction given in the study phase, by pressing designated keys with their left and right index fingers on a keyboard. The hands used for old and new responses were counterbalanced across participants. Once a response was made, the next trial commenced.

Participants were instructed to minimise unnecessary movement during the study and test blocks and limit eye blinking to the periods when the fixation cross appeared on the screen. They were allowed to take short breaks between study and test blocks. Within these breaks they were allowed to move carefully and blink as they wished.

### 3.1.2 Results

#### 3.1.2.1 Behavioural Data

Table 3.2 shows mean probabilities of hits and false alarms across instruction (remember, forget) and valence (negative vs neutral and positive vs neutral). Estimates for sensitivity and criterion separated for instruction and valence are presented in Table 3.3.

Table 3.2
*Probabilities of Correct Judgements (Hits) to Old Items and Incorrect Judgements (False Alarms) to New Items across Instruction (Remember/Forget) and Valence (Negative, Positive and Neutral) for Studies 1 and 2. SD = standard deviation*

| | | Study 1 | | Study 2 | |
|---|---|---|---|---|---|
| | | Negative M (SD) | Neutral M (SD) | Positive M (SD) | Neutral M (SD) |
| Hits | Remember | 0.74 (0.11) | 0.70 (0.13) | 0.69 (0.12) | 0.61 (0.14) |
| | Forget | 0.59 (0.21) | 0.50 (0.19) | 0.58 (0.16) | 0.52 (0.15) |
| False alarms | | 0.18 (0.13) | 0.15 (0.13) | 0.22 (0.13) | 0.18 (0.13) |

Table 3.3
*Summary Statistics for Each Valence Type across Instructions for Studies 1 and 2. d' is the Estimate of Memory Sensitivity and c is the Estimate of Criterion. SD = standard deviation*

| | | Study 1 | | Study 2 | |
|---|---|---|---|---|---|
| | | Negative M (SD) | Neutral M (SD) | Positive M (SD) | Neutral M (SD) |
| *d'* | Remember | 1.72 (0.53) | 1.77 (0.74) | 1.35 (0.56) | 1.31 (0.52) |
| | Forget | 1.32 (0.50) | 1.25 (0.48) | 1.05 (0.44) | 1.08 (0.32) |
| c | Remember | 0.12 (0.40) | 0.29 (0.42) | 0.15 (0.30) | 0.38 (0.35) |
| | Forget | 0.32 (0.46) | 0.55 (0.53) | 0.30 (0.38) | 0.49 (0.42) |

For each study, 2 x 2 repeated measures ANOVAs were conducted on the two measures shown in Table 3.3: sensitivity (*d'*) and response criterion (c). In each case these analyses initially included the factors of instruction (remember vs forget) and valence (two categories in each case). Only main effects are elaborated on below, as there were no reliable interactions between factors.

### *3.1.2.1.1 Effects of Instruction*

There were main effects of instruction in both studies for sensitivity and response criterion. In the case of sensitivity this reflects a directed forgetting effect: *d'* was superior for TBR compared to TBF words (see Table 3.4). The main effect of instruction for response criterion reflects a more liberal criterion for TBR than for TBF words.

#### *3.1.2.1.2 Effects of Valence*

There were main effects of valence only for response criterion (see Table 3.4). A more liberal response criterion was observed for emotional words relative to neutral words in both studies.

Table 3.4
*Summary of Repeated Measures ANOVA Results for Comparisons between Instruction and Valence in Studies 1 and 2*

| | | Study 1 | | | | Study 2 | | | |
|---|---|---|---|---|---|---|---|---|---|
| Measure (DV) | | *df* | *F* | *p* | $\eta_p^2$ | *df* | *F* | *p* | $\eta_p^2$ |
| Instruction | *d'* | 1, 15 | 15.34 | .001** | .51 | 1, 15 | 8.69 | .010** | .37 |
| | c | 1, 15 | 15.34 | .001** | .51 | 1, 15 | 8.69 | .010** | .37 |
| Valence | *d'* | 1, 15 | 0.02 | .891 | .00 | 1, 15 | 0.01 | .920 | .00 |
| | c | 1, 15 | 20.52 | < .001*** | .58 | 1, 15 | 20.37 | < .001*** | .58 |
| Instruction x Valence Interactions | *d'* | 1, 15 | 1.96 | .182 | .12 | 1, 15 | 0.52 | .481 | .03 |
| | c | 1, 15 | 1.96 | .182 | .12 | 1, 15 | 0.52 | .481 | .03 |

*Notes.* ** $p < .01$, *** $p < .001$.
df = degrees of freedom

### 3.1.2.2 EEG Data

Analyses of the neural data were targeted at mid-frontal sites from 300-500ms to assess changes in the putative index of familiarity according to valence and response categories (i.e. remember hit, forget hit and correct rejections). Furthermore, analyses were targeted at left-parietal sites from 500-800ms to assess changes in the index of recollection according to the same factors.

Figure 3.1 (negative vs neutral) and Figure 3.2 (positive vs neutral) show the grand average ERPs for the three critical response categories over mid-frontal (F3, Fz, F4) and left-parietal (P3, P5, P7) regions. The translucent grey areas indicate the critical time windows for the analyses of the mid-frontal (300-500ms) and left-parietal (500-800ms) old/new effects. The scalp maps at the foot of each figure are computed on the basis of difference scores obtained by subtracting the mean amplitude measures for correct rejections from those for the TBR and TBF hit categories. The method used is a spherical spline interpolation (Perrin et al., 1989).

For study 1 (see Figure 3.1) the old/new effects are small with no marked changes according to valence. Moreover, in contrast to what is often observed, some of the old/new effects are negative-going, most notably for TBF items. For study 2 (see Figure 3.2), there are only small differences between old/new effects at frontal sites between 300 and 500ms. Over parietal scalp between 500 and 800ms the old/new effects for positive words are larger than those for neutral words. There are no marked differences according to instruction (TBR/TBF). The key outcomes are shown in Figure 3.3, which depicts the mean amplitudes of the key old/new effects and comprise a summary of the data that were submitted to the analyses described below.

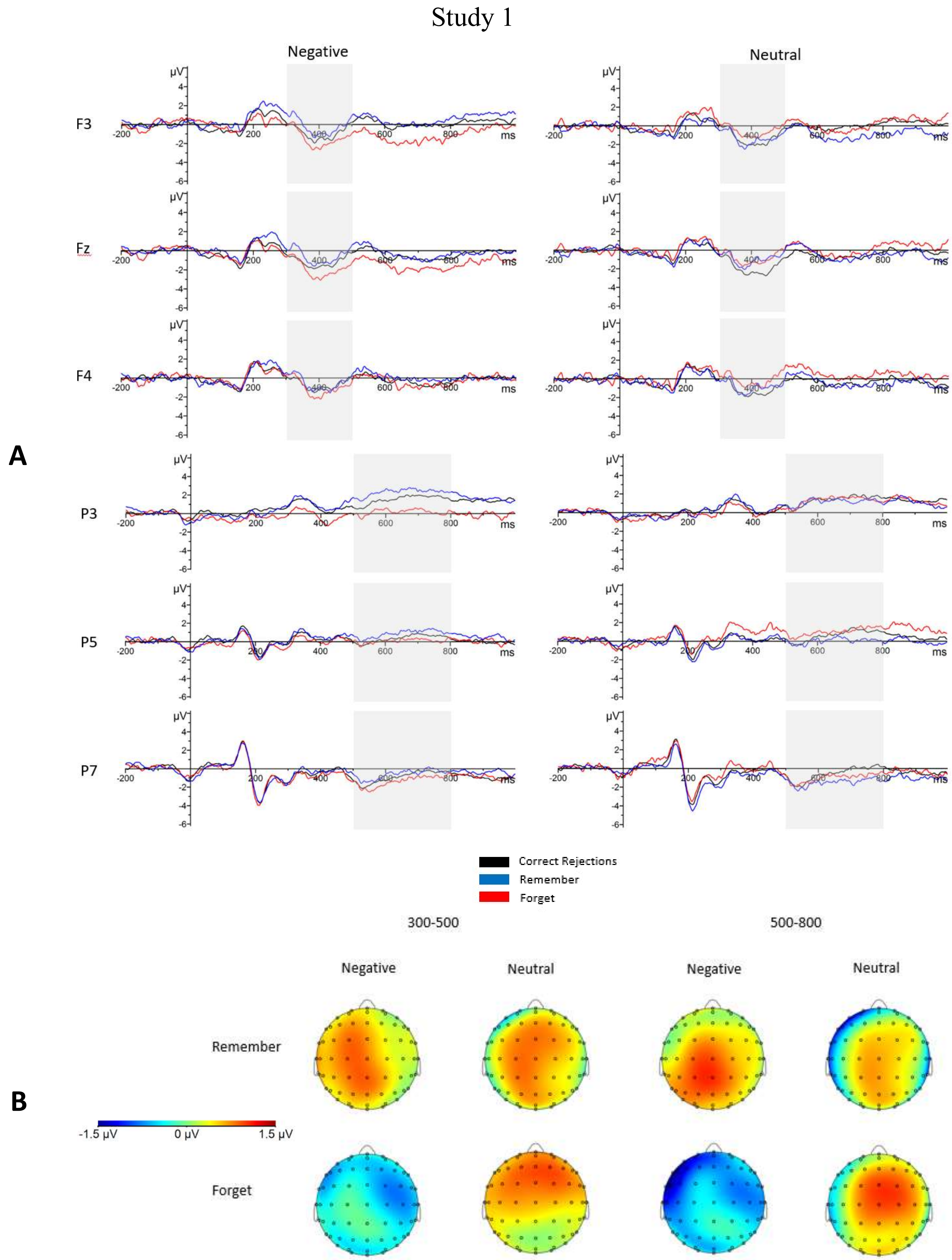
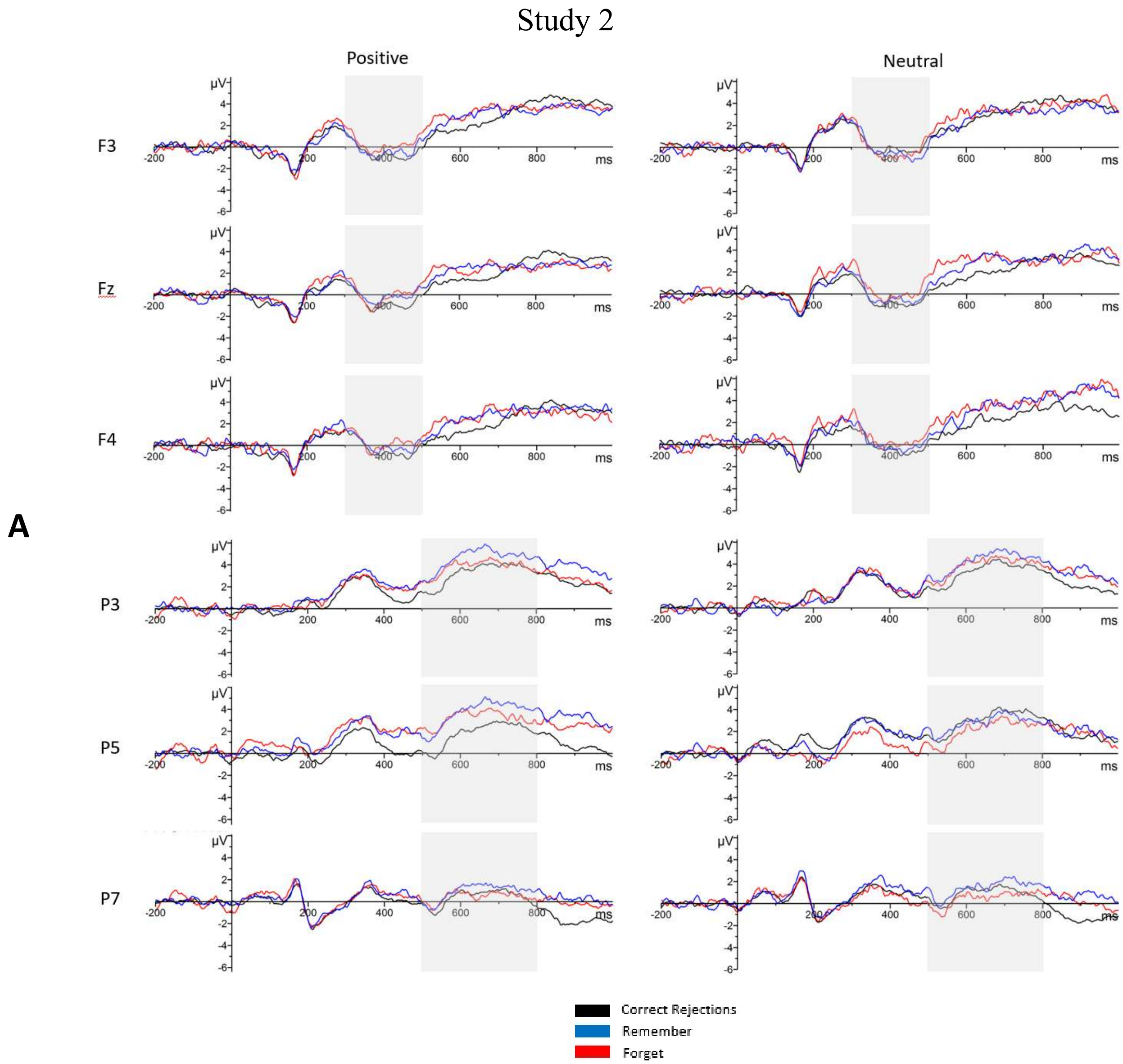



*Figure 3.1.* (A) Grand average ERPs for frontal sites (channel F3, Fz, F4) in a time window of 300-500ms and for parietal sites (channels P3, P5, P7) in a time window of 500-800ms for negative (left column) and neutral (right column) material, separated for remember hits (RH), forget hits (FH) and correct rejections (CR). (B) Scalp maps showing the scalp distributions of the differences between the neural activities elicited by correct old/new judgements to old and new items for the time windows 300-500ms and 500-800ms for negative and neutral material.



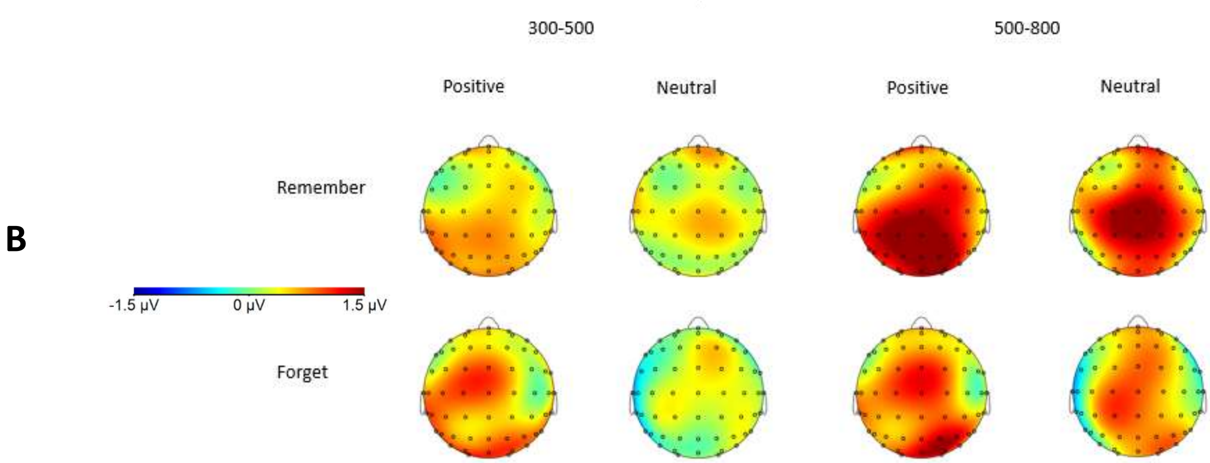




*Figure 3.2*. (A) Grand average ERPs for frontal sites (channel F3, Fz, F4) in a time window of 300-500ms and for parietal sites (channels P3, P5, P7) in a time window of 500-800ms for positive (left column) and neutral (right column) material, separated for remember hits (RH), forget hits (FH) and correct rejections (CR). (B) Scalp maps showing the scalp distributions of the differences between the neural activities elicited by correct old/new judgements to old and new items for the time windows 300-500ms and 500-800ms for positive and neutral material.

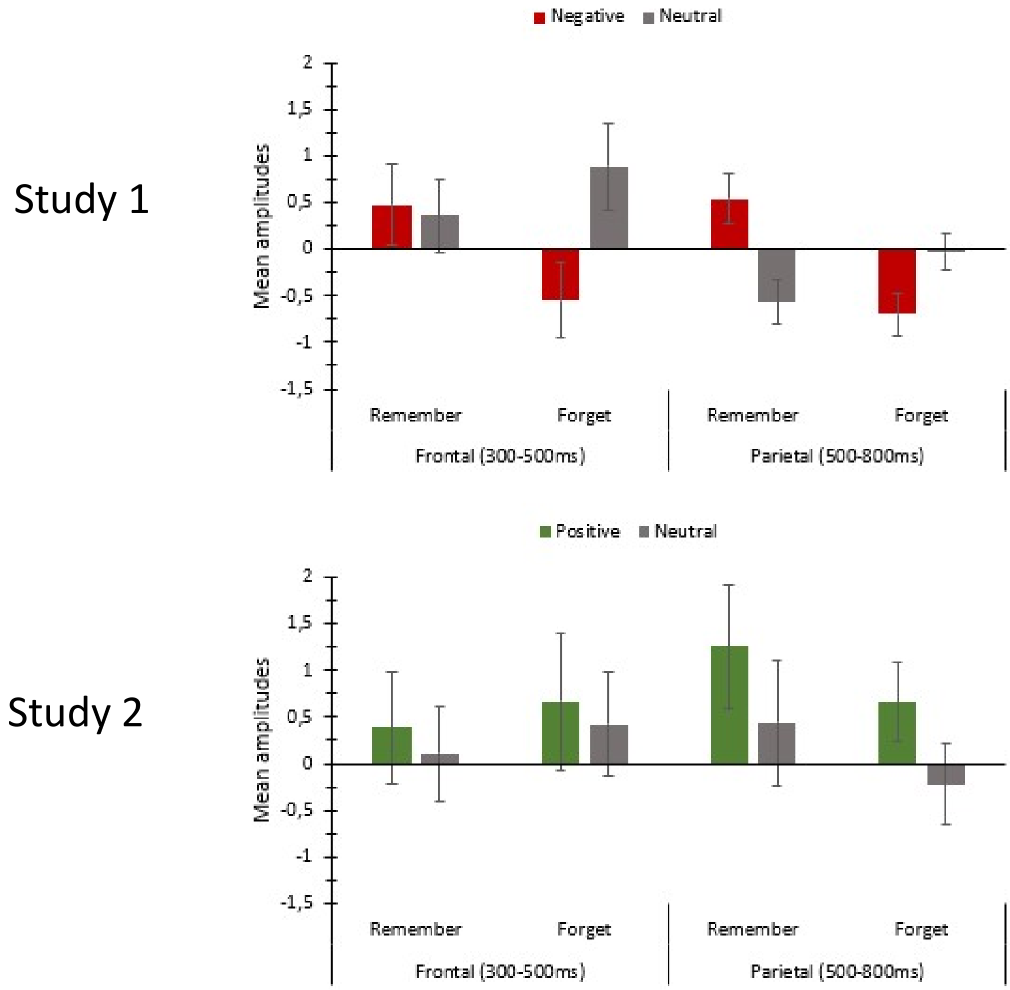



*Figure 3.3.* Mean amplitudes of the TBR and TBF old/new effects in studies 1 and 2 separated by valence and instruction. The old/new effects were calculated by subtracting the mean amplitudes for correct rejections from those for TBR and TBF hits. The mean amplitudes are averages taken across channels F3, Fz, F4) for the 300-500ms window and channels P3, P5, P7 for the 500-800ms window. Error bars represent standard error.

For each study, 3 x 2 x 3 repeated measures ANOVAs were conducted on mean amplitudes for the 300-500ms (mid-frontal old/new effect; F3, Fz, F4) and the 500-800ms (left-parietal old/new effect; P3, P5, P7) data shown in Figure 3.3. In each case these analyses initially included the factors of response category (remember hits, forget hits and correct rejections), valence (two categories in each case) and site as described above.

#### *3.1.2.2.1 Mid-Frontal Old/New Effect*

There were no main effects of valence or response category and no interactions for the analyses of the mid-frontal old/new effects in either of the studies (see Table 3.5).

Table 3.5

*Summary of Repeated Measures ANOVA Results for Comparisons between Response Category and Valence for the Mid-Frontal Old/New Effects in Studies 1 and 2*

| | Study 1 | | | | Study 2 | | | |
|---|---|---|---|---|---|---|---|---|
| Measure (DV) | *df* | *F* | *p* | $\eta_p^2$ | *df* | *F* | *p* | $\eta_p^2$ |
| Response Category | 2, 30 | 1.11 | .344 | .07 | 2, 30 | 0.67 | .517 | .04 |
| Valence | 1, 15 | 0.03 | .861 | .00 | 1, 15 | 0.11 | .745 | .01 |
| Response Category*Valence | 2, 30 | 2.84 | .074 | .16 | 2, 30 | 0.10 | .904 | .01 |

*Note. df* = degrees of freedom

#### *3.1.2.2.2 Left-Parietal Old/New Effect*

There was a significant interaction for valence and response category in study 1 (see Table 3.6). However, simple main effects were not significant and there were no reliable differences in mean amplitudes in old/new judgements (forget, remember and new words) for negative words ($F(2,14) = 2.29$, $p = .138$) and neutral words ($F(2,14) = 0.78$, $p = .477$). The interaction term likely reflects the presence of a small positive-going old/new effect for negative TBR items, and either no or small negative-going old/new effects in the other three instances.

Table 3.6

*Summary of Repeated Measures ANOVA Results for Comparisons between Response Category and Valence for the Left-Parietal Old/New Effects in Studies 1 and 2*

| | Study 1 | | | | Study 2 | | | |
|---|---|---|---|---|---|---|---|---|
| Measure (DV) | *df* | *F* | *p* | $\eta_p^2$ | *df* | *F* | *p* | $\eta_p^2$ |
| Response Category | 2, 30 | 0.75 | .482 | .05 | 2, 30 | 1.80 | .183 | .11 |
| Valence | 1, 15 | 0.00 | .988 | .00 | 1, 15 | 0.00 | .988 | .00 |
| Response Category*Valence | 2, 30 | 3.66 | .038* | .20 | 2, 30 | 0.85 | .439 | .05 |

*Note.* $^*p < .05$.
df = degrees of freedom

### 3.1.3 Discussion

The overall objective in this experiment was to contribute to our understanding of how control over remembering and forgetting of emotional material and the processes of recollection and familiarity are linked. A key element of this experiment was the use of ERPs to index the engagement of recollection and familiarity. The motivation for using a neural measure was to provide a measure of recollection and familiarity when using only old/new judgements in a recognition test, recognising that, as described earlier, when confidence judgements or different response options other than old/new judgements are required at test participants may rely differently on recollection and familiarity.

Behaviourally, and consistent with previous experiments in this thesis and the wider literature, there was a directed forgetting effect in both studies: better memory for TBR words compared to TBF words (MacLeod, 1999). Furthermore, consistent with the findings in Experiments 2 and 3, emotion did not influence the ability to remember and forget nor did emotion enhance memory sensitivity. Also, in line with earlier findings, there was a more liberal criterion for emotional words (negative and positive) compared to neutral words.

Electrophysiologically, there were no reliable old/new effects in the mid-frontal and left-parietal sites, with the exception of a small and difficult to interpret interaction at parietal sites in study 1. Consequently, the data have little say about how directed forgetting and emotion influence memory processes. Although not reliable, in study 2 there was a typical old/new effect pattern in which ERPs for hits (remember and forget) were more positive-going than for correct rejections for both positive and neutral words. The absence of reliable effects is surprising, and further comment is deferred pending the report of a replication attempt for study 1, given the unexpected pattern of outcomes.

### 3.2 A Full Replication of Study 1 (Negative and Neutral Words)

The absence of reliable positive-going mid-frontal and left-parietal old/new effects in the first ERP experiment in this thesis is inconsistent with the vast majority of published findings. For both effects there is an extensive literature documenting their scalp distributions, polarity and sensitivity to experiment manipulations (Friedman & Johnson, 2000; Rugg & Curran, 2007; Sanquist et al., 1980). These effects have also been revealed in DF studies (Nowicka et al., 2009), suggesting that this specific manipulation of instructions during encoding is unlikely to be a reason for the observed data patterns in study 1. Indeed, the published literature contains numerous reports of the consistency in the polarity of old/new effects across many encoding manipulations. The outcomes in study 2 are qualitatively similar to what is commonly observed.

Inspection of the data for individual participants in study 1 did not reveal any outliers, and the proportion of trials rejected due to noise is not excessive, and is comparable to that for study 2 (~12% in both cases). The levels of response accuracy in this study, while not very high, are similar to those in studies where positive-going ERP old/new effects have been revealed (see, for example Tsivilis et al., 2015). These considerations do not provide a clear explanation for the findings in study 1, but they do argue for another study designed to assess how robust the outcomes are. If it is the case that old/new effects comparable to those in study 1 are replicated, then it would suggest that there are specific elements of the design employed here that are responsible for effects that differ strikingly from those that are typically reported. The most salient of these would be the stimuli employed. The outcomes and a discussion of the findings are provided below.

### 3.2.1 Methods

#### 3.2.1.1 Participants

These were undergraduates and postgraduates from the University of Nottingham. The same inclusion criteria were applied as in study 1 and 2 (see section 3.1.1.1). There were a total of 22 participants of which 6 participants were excluded. 3 participants were excluded because they did not meet the inclusion criteria for sensitivity, 1 was excluded because of extensive noise in the EEG data resulting in rejection of more than 25% of trials and 2 participants were excluded because of extensive alpha activity that was specific to some critical response categories. Data from the remaining 16 participants were used for behavioural and EEG analyses (8 females, $M$ = 23.25, $SD$ = 3.32). After giving informed consent, participants completed the experiment and received an inconvenience allowance of £10.

#### 3.2.1.2 Design, Materials and Procedure

The design, materials and procedure of the task, as well as the EEG recording protocols were kept the same as in the EEG studies described above (see sections 3.1.1.2 to 3.1.1.4). See Table 3.7 for mean numbers of trials contributing to each of the response categories in the EEG data compared to study 1. An average of 93.8% of trials per participant was retained for analyses after pre-processing the EEG data (in study 1 there was an average of 87.6% of trials per participant retained).

Table 3.7
*Mean Numbers of Trials used for Averaging separated by Response Category (Correct Rejections and Hits separated by Instruction (Remember/Forget) and Valence (Negative, Positive and Neutral) for Study 1 and the Replication of Study 1. SD = standard deviation*

| | Study 1 | | Replication Study 1 | |
|---|---|---|---|---|
| | Negative M (SD) | Neutral M (SD) | Negative M (SD) | Neutral M (SD) |
| Hit (Remember) | 40 (6.92) | 37 (7.75) | 40 (7.84) | 36 (9.99) |
| Hit (Forget) | 32 (9.53) | 27 (9.42) | 32 (6.74) | 27 (4.97) |
| Correct Rejection | 85 (19.95) | 89 (22.22) | 93 (15.16) | 98 (13.34) |

### 3.2.2 Results

#### 3.2.2.1 Behavioural Data

Table 3.8 shows mean probabilities of hits and false alarms across instruction (remember, forget) and valence (negative vs neutral) for study 1 and the replication of study 1. Estimates for sensitivity and criterion separated for instruction and valence are presented in Table 3.9.

Table 3.8
*Probabilities of Correct Judgements (Hits) to Old Items and Incorrect Judgements (False Alarms) to New Items across Instruction (Remember/Forget) and Valence for Study 1 and the Replication of Study 1. SD = standard deviation*

| | | Study 1 | | Replication Study 1 | |
|---|---|---|---|---|---|
| | | Negative M (SD) | Neutral M (SD) | Negative M (SD) | Neutral M (SD) |
| Hit | Remember | 0.74 (0.11) | 0.70 (0.13) | 0.72 (0.16) | 0.64 (0.17) |
| | Forget | 0.59 (0.21) | 0.50 (0.19) | 0.56 (0.11) | 0.47 (0.10) |
| False alarm | | 0.18 (0.13) | 0.15 (0.13) | 0.17 (0.12) | 0.13 (0.09) |

Table 3.9
*Summary Statistics for Each Valence Type across Instructions for Study 1 and the Replication of Study 1. d' is the Estimate of Memory Sensitivity and c is the Estimate of Criterion. SD = standard deviation*

| | | Study 1 | | Replication Study 1 | |
|---|---|---|---|---|---|
| | | Negative M (SD) | Neutral M (SD) | Negative M (SD) | Neutral M (SD) |
| $d'$ | Remember | 1.72 (0.53) | 1.77 (0.74) | 1.69 (0.61) | 1.63 (0.63) |
| | Forget | 1.32 (0.50) | 1.25 (0.48) | 1.23 (0.42) | 1.17 (0.43) |
| c | Remember | 0.12 (0.40) | 0.29 (0.42) | 0.22 (0.41) | 0.43 (0.36) |
| | Forget | 0.32 (0.46) | 0.55 (0.53) | 0.45 (0.38) | 0.66 (0.31) |

While the outcomes of study 1 have already been reported, they are repeated here for ease of the direct comparison with the replication outcomes. For the replication of study 1, 2 x 2 repeated measures ANOVAs were conducted on the two measures shown in Table 3.9: sensitivity ($d'$) and response criterion (c). In each case these analyses initially included the factors of instruction (remember vs forget) and valence (negative vs neutral).

#### *3.2.2.1.1 Effects of Instruction*

Similar to the results in study 1, there were main effects of instruction for sensitivity and response criterion. In the case of sensitivity this reflects a directed forgetting effect: *d'* was superior for TBR in comparison to TBF words (see Table 3.10). In what is a fairly typical pattern in the work reported in this thesis, the main effect of instruction for response criterion reflects a more liberal criterion for TBR than for TBF words.

Table 3.10
*Summary of Repeated Measures ANOVA Results for Comparisons between Instruction and Valence in Study 1 and the Replication of Study 1*

| | | Study 1 | | | | Replication Study 1 | | | |
|---|---|---|---|---|---|---|---|---|---|
| Measure (DV) | | *df* | *F* | *p* | $\eta_p^2$ | *df* | *F* | *p* | $\eta_p^2$ |
| Instruction | *d'* | 1, 15 | 15.34 | .001** | .51 | 1, 15 | 26,49 | < .001*** | .64 |
| | c | 1, 15 | 15.34 | .001** | .51 | 1, 15 | 26,49 | < .001*** | .64 |
| Valence | *d'* | 1, 15 | 0.02 | .891 | .00 | 1, 15 | 1.28 | .276 | .08 |
| | c | 1, 15 | 20.52 | < .001*** | .58 | 1, 15 | 25.94 | < .001*** | .63 |
| Instruction x Valence Interactions | *d'* | 1, 15 | 1.96 | .182 | .12 | 1, 15 | 0.00 | .951 | .00 |
| | c | 1, 15 | 1.96 | .182 | .12 | 1, 15 | 0.00 | .951 | .00 |

*Notes.* ** $p < .01$, *** $p < .001$.
df = degrees of freedom

#### *3.2.2.1.2 Effects of Valence*

There were main effects of valance only for response criterion (see Table 3.10). A more liberal response criterion was observed for negative words relative to neutral words in both cases.

### 3.2.2.2 EEG Data

As for the two initial studies, analyses of the neural data were targeted at mid-frontal sites from 300-500ms and at left-parietal sites from 500-800ms.

Figure 3.4 shows the grand average ERPs for the three critical response categories (remember hits, forget hits and correct rejections) at frontal (F3, Fz, F4) and left-parietal (P3, P5, P7) sites. The translucent grey areas indicate the critical time windows for the analyses of the mid-frontal (300-500ms) and left-parietal (500-800ms) old/new effects. The scalp maps at the foot of the figure are computed in the same way as for studies 1 and 2.

There appear to be, for the most part, small positive-going old/new effects at frontal and parietal sites. The exception to this general pattern is what appears to be an early negative-going mid-frontal ERP old/new effect for TBR words in the negative condition. The key outcomes are shown in Figure 3.5, which depicts the mean amplitudes of the key old/new effects and comprise a summary of the data that were submitted to the analyses described below compared to the mean amplitudes observed in study 1.

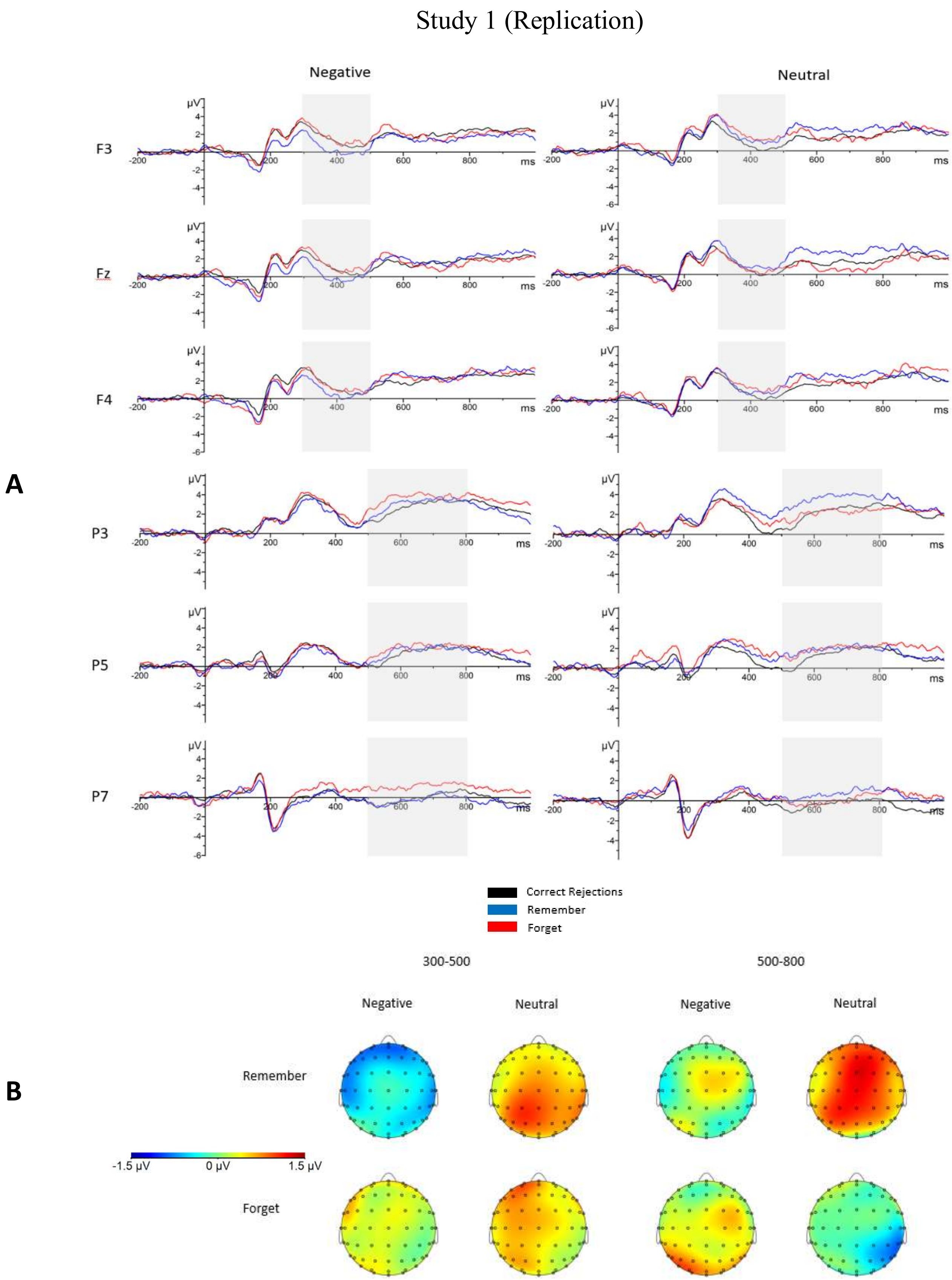



*Figure 3.4.* (A) Grand average ERPs for frontal sites (channel F3, Fz, F4) in a time window of 300-500ms and for parietal sites (channels P3, P5, P7) in a time window of 500-800ms for negative (left column) and neutral (right column) material, separated for remember hits (RH), forget hits (FH) and correct rejections (CR). (B) Scalp maps showing the scalp distributions of the differences between the neural activities elicited by correct old/new judgements to old and new items for the time windows 300-500ms and 500-800ms for negative and neutral material.

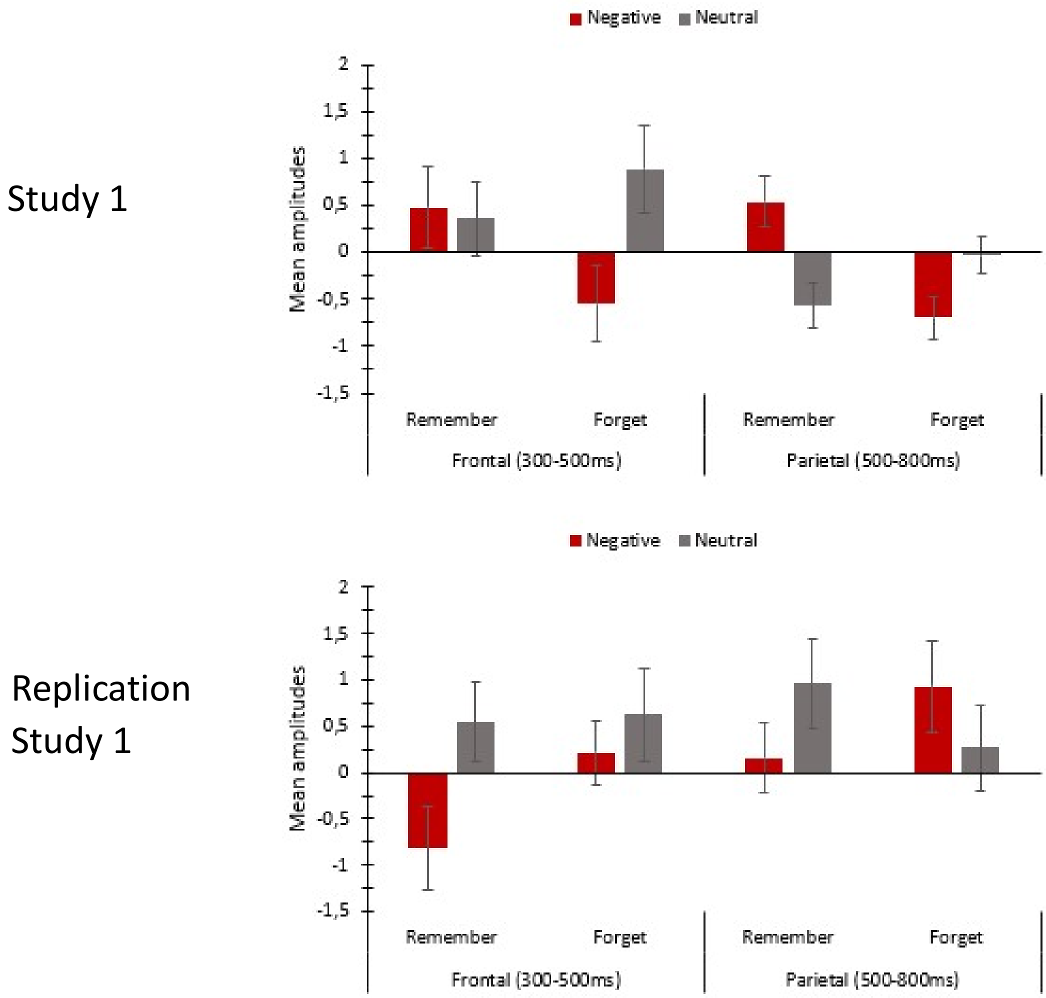



*Figure 3.5*. Mean amplitudes of the TBR and TBF old/new effects in studies 1 and the replication of study 1 separated by valence and instruction (remember and forget). The old/new effect magnitudes were calculated by subtracting the mean amplitudes for correct rejections from those for TBR and TBF hits. The mean amplitudes are averages taken across channels F3, Fz, F4) for the 300-500ms window and channels P3, P5, P7 for the 500-800ms window. Error bars represent standard error.

For the replication of study 1, 3 x 2 x 3 repeated measures ANOVAs were conducted on mean amplitudes for the 300-500ms (mid-frontal old/new effect; F3, Fz, F4) and the 500-800ms (left-parietal old/new effect; P3, P5, P7) data shown in Figure 3.5. In each case these analyses initially included the factors of response category (remember hits, forget hits and correct rejections), valence (negative and neutral) and site as described above.

### *3.1.2.2.1 Mid-Frontal Old/New Effect*

There was a significant interaction for valence and response category (see Table 3.11) in the replication study, which was followed up via simple main effects tests. Simple main effects showed a significant effect for TBR words ($F(1,15) = 3.60$, $p = .037$, $\eta_p^2 = .19$), whereas no reliable differences for TBF words were apparent. The outcomes likely reflect the small positive-going old/new effects for neutral words, and the absence of an effect (TBF) or a negative-going effect (TBR) for negative words.

Table 3.11
*Summary of Repeated Measures ANOVA Results for Comparisons between Response Category and Valence for the Mid-Frontal Old/New Effects in Study 1 and the Replication of Study 1*

| | Study 1 | | | | Replication Study 1 | | | |
|---|---|---|---|---|---|---|---|---|
| Measure (DV) | *df* | *F* | *p* | $\eta_p^2$ | *df* | *F* | *p* | $\eta_p^2$ |
| Response Category | 2, 30 | 1.11 | .344 | .07 | 2, 30 | 1.32 | .283 | .08 |
| Valence | 1, 15 | 0.03 | .861 | .00 | 1, 15 | 0.69 | .420 | .04 |
| Response Category*Valence | 2, 30 | 2.84 | .074 | .16 | 2, 30 | 3.60 | .040* | .19 |

*Note.* * $p < .05$
df = degrees of freedom

### *3.1.2.2.2 Left-Parietal Old/New Effect*

There were no main effects of valence or response category and no interactions between these factors (see Table 3.12).

Table 3.12
*Summary of Repeated Measures ANOVA Results for Comparisons between Response Category and Valence for the Left-Parietal Old/New Effects in Study 1 and the Replication of Study 1*

| | Study 1 | | | | Replication Study 1 | | | |
|---|---|---|---|---|---|---|---|---|
| Measure (DV) | *df* | *F* | *p* | $\eta_p^2$ | *df* | *F* | *p* | $\eta_p^2$ |
| Valence | 1,15 | 0.00 | .988 | .00 | 1, 15 | 0.53 | .477 | .03 |
| Response Category | 2, 30 | 0.75 | .482 | .05 | 2, 30 | 1.69 | .202 | .10 |
| Valence*Response Category | 2, 30 | 3.66 | .038* | .20 | 2, 30 | 2.87 | .072 | .16 |

*Note.* * $p < .05$.
df = degrees of freedom

### 3.2.3 Discussion

For the replication study, sensitivity was higher for TBR items than TBF items, indicating a directed forgetting (DF) effect. There was no effect of emotion on memory sensitivity nor did the DF effect vary according to emotion. These findings are consistent with those in study 1. Also consistent with study 1 is the finding of a more liberal response criterion for negative items than for neutral items. These behavioural findings are also consistent with the findings from study 2, as well as Experiments 2 and 3.

The EEG data analyses in the replication study indicated that the left-parietal old/new effect did not differ according to DF instructions or emotion. There was, however, some indication of divergences at mid-frontal electrode locations. While there were small positive-going old/new effects of similar magnitude for neutral words, the old/new effect for negative words was negative-going for remember instructions and negligible for the forget instruction. These findings are contrary to what was found in study 1, where there were no effects of DF instructions and emotion on both the mid-frontal and left-parietal old/new effect.

Overall, the outcomes in the replication – primarily positive-going old/new effects- are more consistent with what has often been observed (Friedman & Johnson, 2000; Rugg & Curran, 2007; Sanquist et al., 1980; Vilberg & Rugg, 2008). In study 1, the majority of old/new effects were negative-going, and while there have been a few reports of negative-going effects (e.g. Woodruff et al., 2006), they are few and difficult to interpret.

How should the different data patterns in the initial study and the replication be interpreted? The behavioural data in the two studies are broadly similar, making it unlikely that performance changes contribute markedly to the different EEG outcomes. One difference across the studies is the percentage of trials rejected (~6% difference), although the mean trials numbers submitted for averaging and then analysis are reasonable in both cases (see Table 3.7). Moreover, inspection of the means for old/new effects on a by participant basis did not reveal any marked outliers contributing to the overall averaged pattern. These considerations do not offer an immediate explanation for the disparate findings.

One reason for employing ERPs in this experiment was to provide a means of assessing how recollection and familiarity contribute to directed forgetting and vary with emotion when only recognition memory judgements are made. It may be, however, that the old/new instructions make it unlikely to assess changes in recollection, if memory judgements in an old/new recognition memory test rely heavily on familiarity based judgements (Johansson et al., 2004; Tulving, 1985b). However, this explanation does not tend to extend easily to the absence of reliable mid-frontal old/new effects. Nor is it consistent with previous ERP studies in which an old/new recognition memory task was applied (Maratos et al., 2000; Zheng et al., 2018). In short, for recollection, some have found a smaller left-parietal old/new effect for negative material (e.g. Maratos et al., 2000), whereas the opposite has also been observed: a larger left-parietal old/new effect for negative relative to positive and neutral material (e.g. Zheng et al., 2018). Perhaps one way to explain these null results and discrepancies with previous studies showing positive results is to consider experiment design and response accuracy. Using words, Maratos et al. (2000) matched the negative and neutral words only for word length and the degree of semantic relatedness and not on word frequency. Previously, it has been shown that word frequency does influence the magnitude of the old/new effects (Rugg & Doyle, 1992). Frequency was controlled for in the studies reported in this thesis. Perhaps more important are the levels of memory sensitivity, as quite understandably old/new effects increase in magnitude as response accuracy increases. Maratos et al. (2000) and Zheng et al. (2018) had higher levels of response accuracy compared to what was observed in the present experiments. In present studies, memory sensitivity was on average between 1.20 and 1.52 (*d'* collapsed across remember and forget instructions). Memory performance in Maratos et al.'s (2000) study was between 1.64 (negative) and 2.03 (neutral) and in Zheng et al.'s (2018) study between ±2.4 (neutral) and ±2.6 (negative + positive). Hence, the lower levels of sensitivity in the studies reported here are a candidate explanation for the non-signifcant old/new effects.

Whatever the real explanation is, the somewhat disapointing functional consequence is that the ERP data recorded at test have little to say about the links between emotion, control, recollection and familiarity. There are, however, other opportunities for functional insights that the ERP data offer. A particularly interesting question to ask is whether and/or how emotion and DF instructions influence processes that are active during the encoding of emotional and neutral words. Analyses designed to contribute to an understanding of how these elements are linked are described below.

## 3.3 Encoding, Emotion and Directed Forgetting

### 3.3.1 Neural Measures of Encoding and Responses to Directed Forgetting Instructions

Neural activity recorded at the time items are studied can be employed to investigate encoding processes. The most common way that this is done is by separating and contrasting neural activity for items that received either correct or incorrect memory judgements on a subsequent test. Differences between activities of this kind have been termed subsequent memory or Dm (Difference according to memory) effects (Paller et al., 1987), and may well reflect processes important for successful memory encoding.

The subsequent memory approach has been employed extensively and successfully in fMRI studies (for a review of findings see Henson, 2005; Paller & Wagner, 2002), identifying the brain systems that can support successful encoding (e.g. Henson, Rugg, Shallice, Josephs, & Dolan, 1999; Wagner et al., 1998). The same approach with ERPs has a longer history (Paller et al., 1987, 1988), and the results from several studies have provided insights into the time-course and nature of memory encoding processes (e.g. Otten, Henson, & Rugg, 2001; Otten & Rugg, 2001; Otten, Quayle, & Puvaneswaran, 2010).

Often ERP subsequent memory effects when the accompanying test is recognition memory yield small or non-significant differences between neural activities for items attracting correct or incorrect recognition memory judgements. Larger effects are often observed when the test requirements might be considered to rely on recollection (Otten et al., 2001; Otten & Rugg, 2001; Otten et al., 2010; for a review of findings see Friedman & Johnson, 2000). In light of this, Otten and Rugg (2001) argued that when using a recognition memory task to examine subsequent memory effect, it is beneficial to employ confidence judgements. This is an effective means of removing low confidence responses, and/or correct guesses, from the subsequent memoy contrast.

That said, Yick et al. (2015) investigated processes supporting encoding of emotional materials using negative and neutral images and an old/new recognition memory task combined with confidence judgements. To ensure larger effects, ‘guess’ responses were excluded from the ERP analyses. They observed a broadly distributed positive-going effect for negative relative to neutral images that onset around 400ms and lasted for at least 600ms. These outcomes are consistent with the view that different processes operate at encoding according to valence, although the use of only these two valence categories means that ascribing effects to valence or arousal is not straightforward.

The findings of Yick et al. (2015) give some indication that there is merit in analyzing ERP subsequent memory effects in the ERP experiments in this thesis, despite the fact that their results are surprising in the context of the broader research literature, where subsequent memory effects in old/new recognition memory tasks are often small. To anticipate, however, and in keeping with the observations made in the preceding paragraph, and probably also related to the relatively modest levels of response accuracy in these experiments, there were no significant subsequent memory effects in the first ERP study and the replication. There was a reliable subsequent memory effect in the second study, but it did not vary according to valence. For this reason, these subsequent memory effects are summarized, along with the reports of their analysis, in Appendix M.

The designs of the ERP experiments in this thesis, however, also permit an analysis of the neural activities elicited by the cues to remember/forget. These ERPs offer access to a related set of questions regarding how memory encoding can be investigated. A first step is to establish whether ERPs vary according to remember/forget instructions. If they do, then a second question is whether subsequent memory effects for remember/forget items are equivalent. If both remembering and forgetting are active (and different) processes, then the subsequent memory effects would be different in the two cases. While the presence of subsequent memory effects for remember cues only would not rule out the same interpretation (ERPs might just not be sensitive to active forgetting processes), this outcome would also be consistent with the view that active processing of items is engaged only following a remember cue.

The designs of the ERP experiments in this thesis also permit assessment of how instructions to remember/forget vary with emotion. In each experiment it is possible to separate remember and forget cues according to the valence of the preceding item. If the differences are comparable in the two cases the most straightforward interpretation is that the effects simply reflect the valence of the preceding items. Of greater interest, however, would be different effects according to valence in the two cases. To consider one outcome: differences according to valence only for remember cues would be consistent with the interpretation offered in the previous paragraph whereby active processing in pursuit of encoding is engaged when a remember but not when a forget cue is encountered.

For DF instructions, differences in brain activity have been reported in a handful of studies. The most common finding is a temporally extended greater relative positivity for TBR items relative to TBF items over posterior sites (Bailey & Chapman, 2012; Gallant et al., 2018; Hauswald et al., 2011). This difference typically onsets around 300ms post-stimulus and continues for at least 300ms. While this finding has been interpreted as an indication of more enhanced processing of items given a TBR instruction, in none of these studies has it been possible to provide a strong basis for this claim: because it was simply a paired contrast in each case the polarity of the difference and claims about greater processing receive only weak support. Hauswald et al. (2011) linked frontal activity (more positive-going for TBF than TBR) to suppression for TBF items (also see Gallant et al., 2018), but again the empirical basis for this claim is weak.

Bailey and Chapman (2012) investigated the effect of valence on the activities elicited by TBR and TBF cues. Besides differential processing of directed forgetting instructions, they found an enhanced late posterior positivity (~450 to 600ms post-stimulus) for emotional material relative to neutral material (also see Gallant & Dyson, 2016), which they have interpreted as an indication of enhanced attention allocation for emotional material (Kok, 1997). They did not, however, observe that emotion influenced the activity associated with TBR and TBF cues differently. Likewise, other researchers have failed to observe any evidence of differential processing following either a remember or forget cue according to emotion (Gallant et al., 2018; Hauswald et al., 2011). These findings are consistent across the different materials used, such as words (Bailey & Chapman, 2012; Gallant et al., 2018) or images (Hauswald et al., 2011), as well as the number of the valence categories that were contrasted.

There are, however, some contrasting data points. Brandt et al. (2013) found a more positive-going posterior modulation following a TBR cue for negative words relative to neutral words. Neural activity did not vary according to emotion following a TBF cue. Likewise, Gallant and Dyson (2016) reported more positive-going posterior activity for a TBR cue following negative words relative to neutral words and an absence of an emotional effect on the TBF cue. The effect was not found when comparing positive with neutral words. These findings have been interpreted as enhanced processing for emotional material that participants were instructed to remember and are consistent with the view that: (a) ERPs elicited by remember/forget cues are not uniformly sensitive to the valence of the preceding memoranda, and (b) forget cues are not subject to active forgetting processes that vary by valence.

In another study, Yang et al. (2012), however, observed enhanced positivity over posterior sites following a TBR cue for neutral images relative to negative images. They linked this effect to the observation that neutral material is more difficult to remember relative to emotional material (Hamann, 2001). In addition, they did observe differences according to valence for forget cues. This comprised more positive-going frontal activity for TBF cues for negative images compared to neutral images. They interpreted these findings as an indication of inhibition processes acting over TBF cues for negative images. At a minimum, their findings do suggest differential and active processing of TBF items in comparison to TBR items, but the inference linking to inhibition derives support only from data in other studies. The frontal activity observed by Yang et al. (2012) and not by Brandt et al. (2013) and Gallant and Dyson (2016), may not necessarily be linked to the different materials used, as images were used by Hauswald et al. (2011) who also did not find any effect of valence on TBF cues.

Using a combination of words and images, Liu et al. (2017) manipulated the encoding context by using negative or neutral images and neutral words that were embedded in these images. Consistent with Brandt et al. (2013) and Gallant and Dyson (2016), they observed a larger positive-going posterior effect for TBR cues that were preceded by a negative image relative to a neutral image. In addition, relatively more positive-going frontal activity was observed for TBF cues relative to TBR cues, which did not vary according to valence.

Taken together, the existing literature does not provide a consistent basis for predicting whether ERPs elicited by TBF cues will vary with emotion. For TBR cues, however, the prior literature suggests a greater relative positivity for emotional material. In order to test this prediction, and to capture effects with possible maxima and posterior and/or anterior locations, a broad analysis of the distribution of any ERP divergences by cue type and emotion was conducted, as described in greater detail below.

### 3.3.2 Analyses

As already stated, there were no reliable subsequent memory effects that varied according to valence (see Appendix M). For this reason, the remainder of the section below is focused on the neural activity that is elicited by remember and forget instructions and assessments of whether valence interacts with these. An overview of these analyses (they are the same for each study) and then the results are reported below.

Based on previous literature investigating the effects of emotion on remembering and forgetting (Hauswald et al., 2011; Liu et al., 2017), analyses were targeted at two distinct time windows: 300-500ms and 500-700ms. Mean amplitudes were extracted for the following electrodes: at posterior regions CP1, CPz, CP2, P1, Pz, P2 and at anterior regions FP1, FPz, FP2, F1, Fz, F2. For each time window 6 x 2 x 2 repeated-measure ANOVAs were computed for both electrode groups including the following factors: Channel, Valence (negative vs neutral or positive vs neutral) and DF instructions (remember vs forget).

### 3.3.3 Results

Figure 3.6 (negative vs neutral), Figure 3.7 (negative vs neutral; replication study) and Figure 3.8 (positive vs neutral) show the grand average ERPs elicited by remember and forget cues separated by valence. They are shown for representative posterior (CPz, Pz) and anterior

(FPz, Fz) electrode locations. The scalp maps at the foot of each figure illustrate the amplitude distributions of the relevant paired contrasts in each of the targeted conditions.

In study 1 (Figure 3.6) the only marked differences are over mid-posterior sites for TBR cues. The figure shows a greater relative positivity for cues preceded by negative rather than neutral words which onsets around 300ms post-stimulus and continues for the duration of the recording epoch. In the replication of study 1 (Figure 3.7), there are similar selective patterns, although the greater relative positivity extends a little more anteriorly. For TBR words in study 2 (Figure 3.8), a similar pattern can be seen, in this case comprising in a greater relative positivity for positive relative to neutral. Moreover, in this study there is some evidence for a relatively greater negativity for positive relative to neutral in response to TBF cues, which is largest over centro-parietal regions.

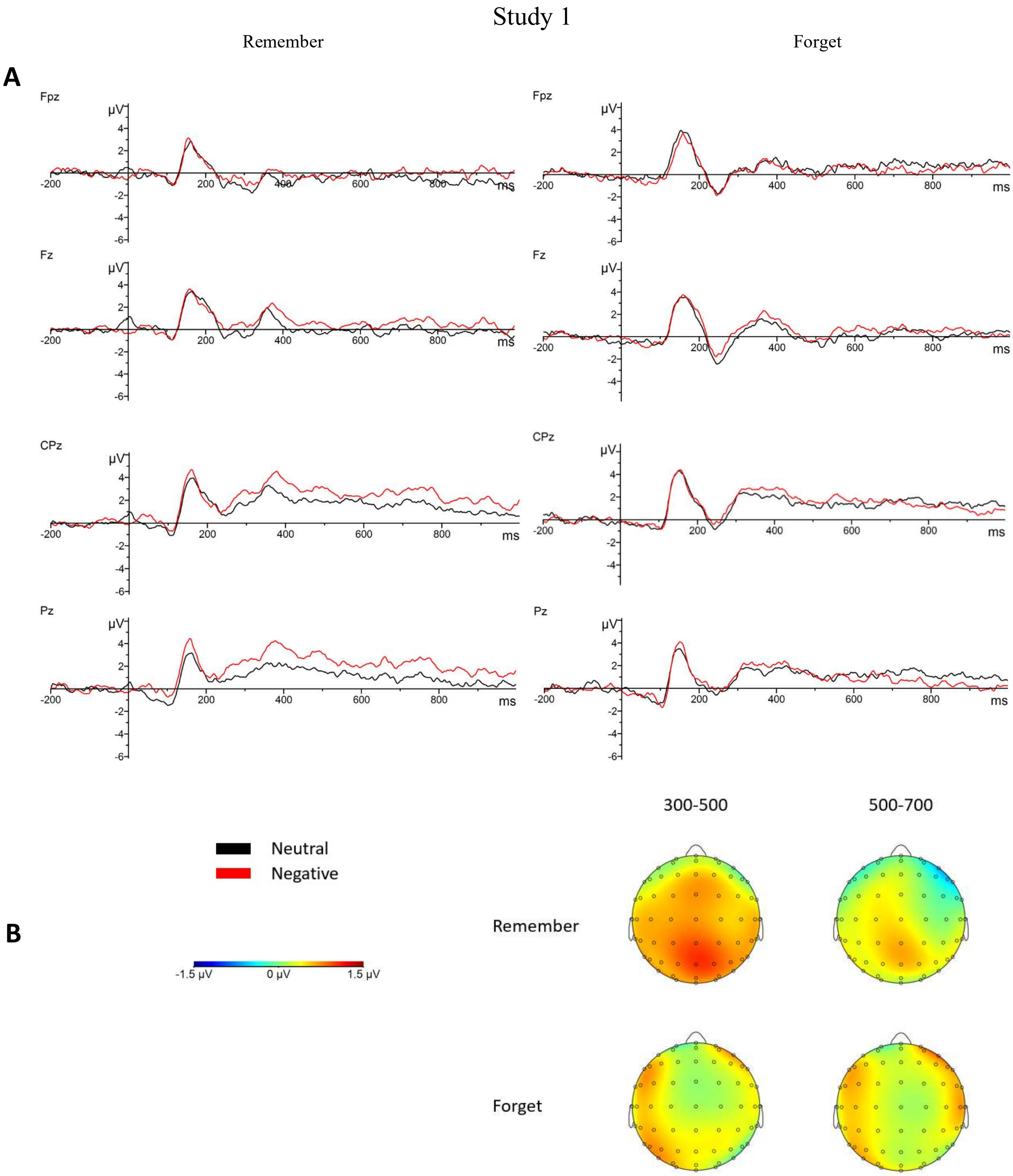



*Figure 3.6.* (A) Grand average ERPs for posterior regions (channels CPz, Pz) and for anterior regions (channels FPz, Fz) for remember (left column) and forget (right column) items, separated for negative and neutral words. (B) Scalp maps showing the differences between neural activities elicited by negative and neutral words in each targeted condition: TBR and TBF for the time windows 300-500ms and 500-700ms.

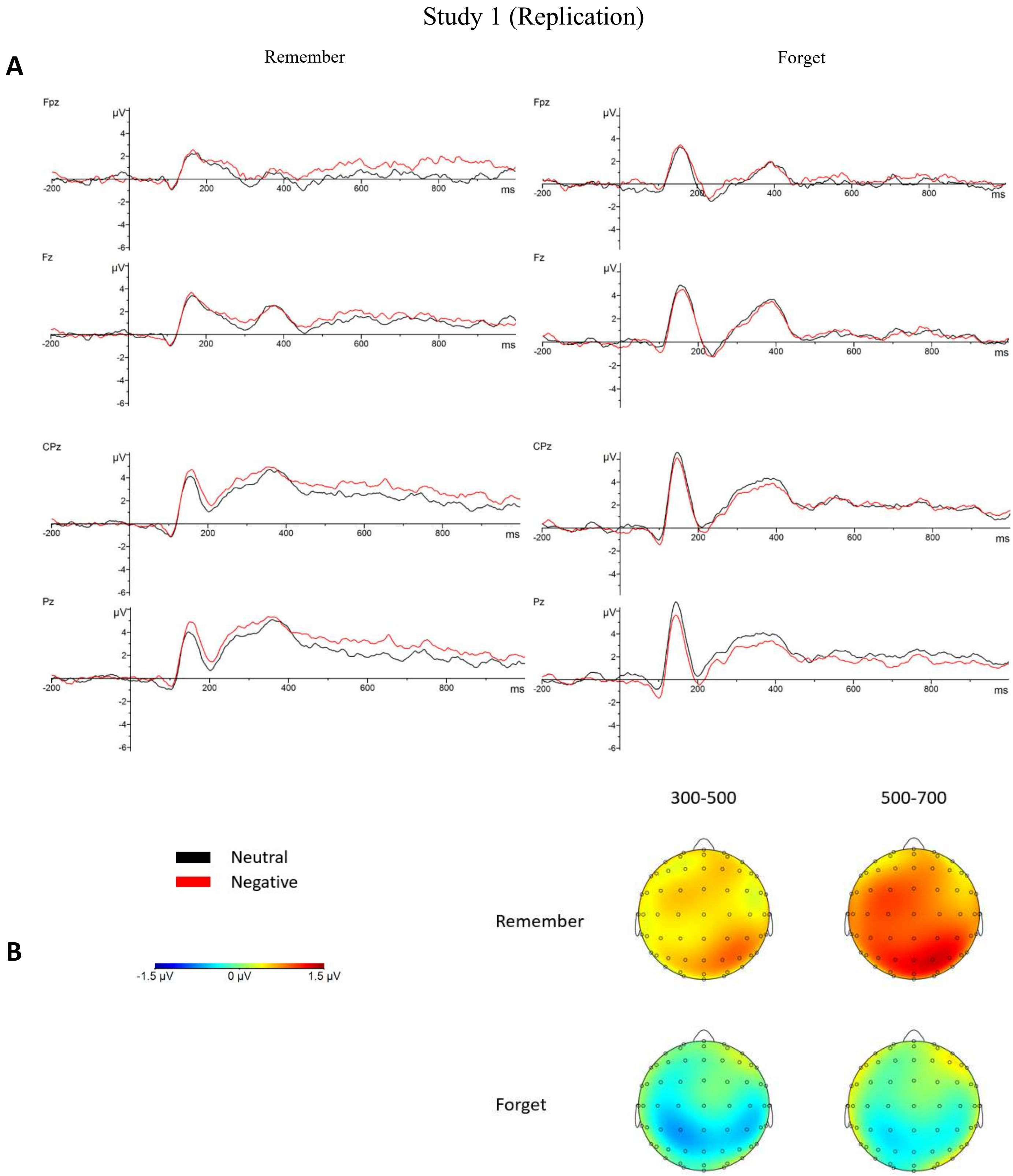



*Figure 3.7.* (A) Grand average ERPs for posterior regions (channels CPz, Pz) and for anterior regions (channels FPz, Fz) for remember (left column) and forget (right column) items, separated for negative and neutral words. (B) Scalp maps showing the differences between neural activities elicited by negative and neutral words in each targeted condition: TBR and TBF for the time windows 300-500ms and 500-700ms.

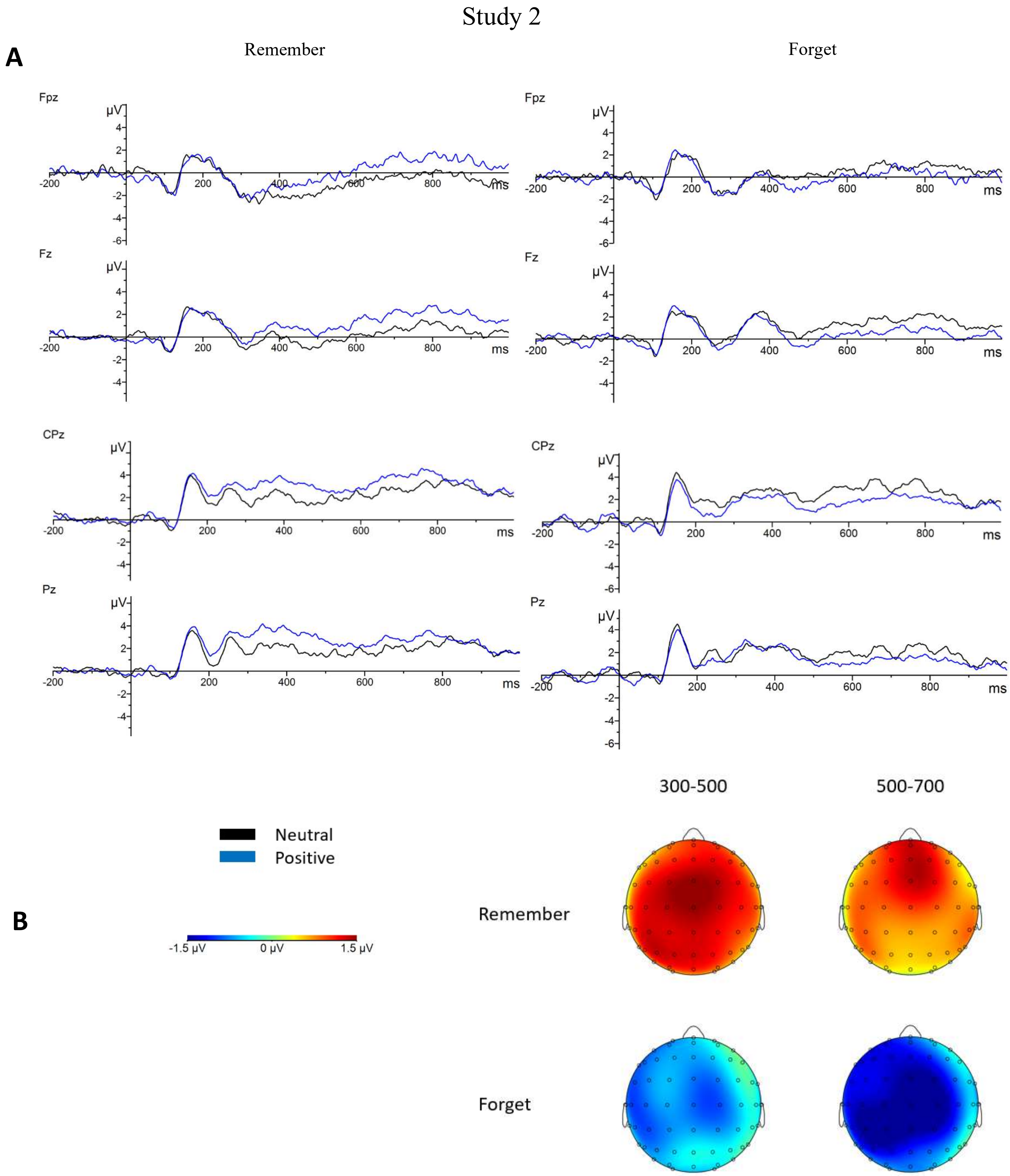



*Figure 3.8.* (A) Grand average ERPs for posterior regions (channels CPz, Pz) and for anterior regions (channels FPz, Fz) for remember (left column) and forget (right column) items, separated for positive and neutral words. (B) Scalp maps showing the differences between neural activities elicited by negative and neutral words in each targeted condition: TBR and TBF for the time windows 300-500ms and 500-700ms.

For each study, 6 x 2 x 2 repeated measures ANOVAs were conducted on mean amplitudes for the 300-500ms time window (posterior regions; CP1, CPz, CP2, P1, Pz, P2) and the 500-700ms time window (anterior regions; FP1, FPz, FP2, F1, Fz, F2). In each case these analyses initially included the factors of channel, instructions (remember and forget) and valence (two categories in each case).

#### 3.3.3.1 300-500ms Time Window

##### *3.3.3.1.1 Anterior Sites*

There were no significant effects in study 1. There was, however, a reliable interaction between valence and instruction in the replication of study 1 and in study 2 (see Table 3.13). However, analysis of simple main effects in the replication of study 1 revealed no significant effects according to valence in the remember condition ($F(1, 15) = 2.96$, $p = .106$) or in the forget condition ($F(1, 15) = 0.44$, $p = .516$). The most likely reason for the significant interaction term is the relatively larger differences according to valence for TBR than for TBF cues.

In a comparable outcome, the interaction term in study 2 arises for similar reasons: simple main effect analyses in study 2 revealed a reliable greater negativity for remember cues following neutral ($M = -1.24$, $SD = 0.70$) words relative to those following positive ($M = -0.25$, $SD = 0.79$) words ($F(1, 15) = 10.98$, $p = .005$, $\eta_p^2 = .42$) . There were no reliable differences in the forget condition.

Table 3.13

*Summary of Repeated Measures ANOVA Results for Comparisons between Instruction and Valence in the 300-500ms time window over Anterior Regions for Study 1, Replication Study 1 and Study 2*

| | Study 1 | | | | Study 1 (Replication) | | | | Study 2 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Measure (DV) | *df* | *F* | *p* | $\eta_p^2$ | *df* | *F* | *p* | $\eta_p^2$ | *df* | *F* | *p* |
| Instruction | 1, 15 | 0.81 | .382 | .05 | 1, 15 | 1.36 | .261 | .08 | 1, 15 | 3.63 | .076 |
| Valence | 1, 15 | 0.56 | .465 | .04 | 1, 15 | 0.46 | .509 | .03 | 1, 15 | 1.17 | .296 |
| Instruction*Valence | 1, 15 | 0.67 | .425 | .04 | 1, 15 | 5.49 | .033* | .27 | 1, 15 | 8.59 | .010* |

*Note.* * $p < .05$.
df = degrees of freedom

#### *3.3.3.1.2 Posterior Sites*

There were no significant effects in study 1, despite the greater relative positivity that is evident in Figure 3.6. There were reliable effects in the replication of study 1 and in study 2. In the replication study main effects of instructions reflect greater relative positivity for TBR ($M$ = 4.06, $SD$ = 0.88) cues relative to TBF ($M$ = 2.91, $SD$ = 0.57) cues over posterior scalp (see Table 3.14). Of greater importance, there were significant interactions between valence and instruction in both the replication of study 1 and in study 2. Simple main effects in the replication of study 1 revealed a greater relative positivity over posterior regions for remember cues following negative ($M$ = 4.37, $SD$ = 0.86) words compared to those following neutral ($M$ = 3.76, $SD$ = 0.92) words ($F(1, 15) = 15.00$, $p = .044$, $\eta_p^2 = .24$). There were no reliable effects in the forget condition. Simple main effects in study 2 reflect an increased positivity over posterior regions for remember cues following positive ($M$ = 3.51, $SD$ = 1.02) words compared to those following neutral ($M$ = 2.25, $SD$ = 0.88) words ($F(1, 15) = 11.75$, $p = .004$, $\eta_p^2 = .44$) only, despite the apparent reversal in the forget condition (see Figure 3.8).

Table 3.14
*Summary of Repeated Measures ANOVA Results for Comparisons between Instruction and Valence in the 300-500ms time window over Posterior Regions for Study 1, Replication Study 1 and Study 2*

| | Study 1 | | | | Study 1 (Replication) | | | | Study 2 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Measure (DV) | *df* | *F* | *p* | $\eta_p^2$ | *df* | *F* | *p* | $\eta_p^2$ | *df* | *F* | *p* |
| Instruction | 1, 15 | 1.73 | .208 | .10 | 1, 15 | 7.20 | .017* | .32 | 1, 15 | 0.63 | .441 |
| Valence | 1, 15 | 3.49 | .081 | .19 | 1, 15 | 0.01 | .920 | .00 | 1, 15 | 5.75 | .118 |
| Instruction*Valence | 1, 15 | 0.60 | .450 | .04 | 1, 15 | 5.71 | .031* | .28 | 1, 15 | 9.97 | .006** |

*Note.* * $p < .05$, ** $p < .01$.
df = degrees of freedom

### 3.3.3.2 500-700ms Time Window

#### *3.3.3.2.1 Anterior Sites*

Reliable effects were observed only in study 2, where there was an interaction between valence and instruction (see Table 3.15). Simple main effects revealed a greater relative positivity over anterior regions for remember cues following positive (*M* = 0.70, *SD* = 0.69) words compared to those following neutral (*M* = -0.43, *SD* = 0.54) words ($F(1, 15) = 7.72$, $p$ = .014, $\eta_p^2 = .34$). In the forget condition, by contrast, there was an increased positivity for neutral (*M* = 0.04, *SD* = 0.45) words relative to positive (*M* = 1.04, *SD* = 0.42) words ($F(1, 15) = 5.46$, $p = .034$, $\eta_p^2 = .27$).

Table 3.15
*Summary of Repeated Measures ANOVA Results for Comparisons between Instruction and Valence in the 500-700ms time window over Anterior Regions for Study 1, Replication Study 1 and Study 2*

| | Study 1 | | | | Study 1 (Replication) | | | | Study 2 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Measure (DV) | *df* | *F* | *p* | $\eta_p^2$ | *df* | *F* | *p* | $\eta_p^2$ | *df* | *F* | *p* | $\eta_p^2$ |
| Instruction | 1, 15 | 0.17 | .689 | .01 | 1, 15 | 1.98 | .180 | .12 | 1, 15 | 0.57 | .463 | .04 |
| Valence | 1, 15 | 0.08 | .778 | .01 | 1, 15 | 2.25 | .154 | .13 | 1, 15 | 0.04 | .836 | .00 |
| Instruction*Valence | 1, 15 | 0.01 | .908 | .00 | 1, 15 | 0.95 | .345 | .06 | 1, 15 | 15.40 | .001** | .51 |

*Note.* * $p < .05$, *** $p < .001$.
df = degrees of freedom

### *3.3.3.2.2 Posterior Sites*

There were no reliable effects in study 1. There was a main effect of instruction in the replication of study 1 (see Table 3.16), which reflected a greater relative positivity for TBR (*M* = 2.76, *SD* = 0.68) words relative to TBF (*M* = 1.79, *SD* = 0.48) words over posterior scalp. In study 2 there was an interaction between valence and instruction. Simple main effects reflect a greater relative positivity for forget cues following neutral (*M* = 2.42, *SD* = 0.71) words relative to those following positive (*M* = 1.15, *SD* = 0.75) words ($F(1, 15) = 9.51$, $p = .008$, $\eta_p^2 = .39$) only, despite the apparent reversal in the remember condition (see Figure 3.8).

Table 3.16

*Summary of Repeated Measures ANOVA Results for Comparisons between Instruction and Valence in the 500-700ms time window over Posterior Regions for Study 1, Replication Study 1 and Study 2*

| | Study 1 | | | | Study 1 (Replication) | | | | Study 2 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Measure (DV) | *df* | *F* | *p* | $\eta_p^2$ | *df* | *F* | *p* | $\eta_p^2$ | *df* | *F* | *p* | $\eta_p^2$ |
| Instruction | 1, 15 | 1.44 | .249 | .09 | 1, 15 | 6.17 | .025* | .29 | 1, 15 | 0.57 | .463 | .04 |
| Valence | 1, 15 | 1.12 | .306 | .07 | 1, 15 | 1.61 | .225 | .10 | 1, 15 | 1.21 | .288 | .08 |
| Instruction*Valence | 1, 15 | 0.17 | .690 | .01 | 1, 15 | 3.26 | .091 | .18 | 1, 15 | 8.07 | .012* | .35 |

*Note.* * $p < .05$.

df = degrees of freedom

### 3.3.4 Discussion

The objective of these analyses was to use the ERP data to understand memory encoding in the context of emotion and directed forgetting. The absence of subsequent memory effects is arguably unsurprising given the previous literature (although see Yick et al., 2015). The outcomes for the analyses of the activities elicited by TBR and TBF cues are, however, of particular interest.

For the replication of study 1 and for study 2, there was a common greater relative positivity for TBR cues that were preceded by emotional rather than neutral words over posterior scalp. This is consistent with previous findings (Brandt et al., 2013; Gallant & Dyson, 2016; Liu et al., 2017). One proposal is that these differences are linked to enhanced attention allocation for emotional material during encoding (Hauswald et al., 2011; Kok, 1997). The presence of this modulation only for TBR is also important because this outcome is consistent with the view that following TBF cues items are subject to relatively little effortful processing, at least in study 1 (replication). Over anterior scalp, in study 2, the same effect was found for positive words followed by a TBR cue relative to neutral words. The greater relative positivity, however, for TBF cues preceded by neutral words over anterior scalp in study 2 suggests a different conclusion. That is, that TBR and TBF cues engage different processes that are differentially influenced by valence. In addition, this anterior effect for neutral TBF cues is also visible over posterior scalp in a later time window (500-700ms). These outcomes for TBR cues are in line with the predictions. No confident prediction was made for TBF cues, and these data points are broadly consistent with those of Yang et al. (2012), but not with those of Brandt et al. (2013) and Gallant and Dyson (2016).

These outcomes – notwithstanding the inconsistencies with some other reports – suggest that people do not apply the same forgetting operations for positive and negative material. A stronger demonstration of this would stem from a within-participant assessment, but the data across assessments are at least suggestive of a common set of processes engaged in response to TBR cues that are agnostic about the positive/negative dimension, and a second set that are specific to positive (relative to neutral) material. Support for accounts that posit active suppression/inhibition as a mechanism for forgetting (Anderson & Hanslmayr, 2014; Basden et al., 1993; Hauswald et al., 2011; Liu et al., 2017; Yang et al., 2012) is therefore evident in study 2. The null result for TBF in study 1 (replication) needs to be treated cautiously, none the less it is worth noting the qualitative similarity with the outcomes in study 2. Moreover, the findings for TBF in study 2 are consistent with the view that active processing of TBF items differed by valence, but whether this neural activity is an index of inhibition or suppression is more challenging on the basis of the current data alone. The N2 component has often been linked with processes of suppression or inhibition (Eimer, 1993; Kok, 1986), which is most commonly examined in a go/no-go task in which participants have to respond to certain stimuli and refrain from responding to other stimuli (Eimer, 1993; Levy & Anderson, 2002). The N2 is observed over frontal scalp in a time window commonly spanning 200-300ms (Eimer, 1993). Linking this to observations in DF studies, enhanced frontal activities for TBF relative to TBR cues have therefore been interpreted as the engagement of inhibitory processes. Yang et al. (2012) observed a similar frontal activity comparable to that seen in go/no-go task for TBF cues. The majority of DF studies, however, have observed frontally distributed activity in a later time window, commonly from 500-600ms onwards (Hauswald et al., 2011; Hsieh et al., 2009; Liu et al., 2017; Paz-Caballero et al., 2004).

# CHAPTER 4

# 4. THE INFLUENCE OF ENCODING CONTEXT IN DIRECTED FORGETTING

## 4.1 Experiment 5 – Encoding Context, Emotion and Directed Forgetting

Across all experiments in this thesis a reliable directed forgetting effect has been observed. When only two classes of valence were compared (Experiments 2–4), emotion did not have an effect on the amount of directed forgetting, nor did it affect memory sensitivity. Furthermore, with the exception of Experiment 1, there was a relatively more liberal criterion for emotional words compared to neutral words. The consistency of these behavioural results, which were observed in the behavioural data for the ERP experiments as well, attests to the robust nature of these findings. The variable findings for recollection and familiarity across some experiments provide somewhat lower confidence in overarching claims that can be made, and as already acknowledged the use of different measures at retrieval across experiments is also a potential factor.

The focus in this experiment is on the ways in which effects due to emotion – most notably criterion change – might be due to processes that operate at encoding and/or retrieval. This builds on prior work where it has been noted that attributing outcomes to either stage is difficult when emotion is manipulated during encoding and retrieval (Maratos & Rugg, 2001; Smith et al., 2004).

As discussed in the Introduction (see sections 1.4.3-1.4.5), how emotion is manipulated and how it affects memory needs to be considered carefully. Emotional materials have increased arousal levels relative to what can be described as neutral materials, and of course differ in valence (Kensinger & Corkin, 2003). For words, other properties including semantic relatedness, word frequency and word length may be important. Another factor to consider when using emotional material, and of primary importance in this chapter, is that when using emotional material during study as well as test, there is an additional potential confound (Liu et al., 2017; Maratos & Rugg, 2001; Smith, Dolan, & Rugg, 2004). This occurs because the emotional attributes of a retrieval cue are confounded with the emotional attributes of the information that needs to be retrieved. Thus, whether emotion influences encoding and/or retrieval processes is hard to distinguish. A generally accepted principle of memory is that attributes of a retrieval cue can influence what information is retrieved, and how the memory trace is experienced (Keele, 1972; Moscovitch, 1992; Schacter et al., 1998). It is therefore possible that differences between the emotional properties of items at test and/or at encoding are responsible for different outcomes on memory assessments (Liu et al., 2017; Maratos & Rugg, 2001; Smith et al., 2004).

One way to avoid this potential confound, if the intention is to understand the influence of emotion at the time of encoding, is to use neutral materials during study and test and manipulate the emotional encoding context in which the material is studied. During a subsequent memory test, the neutral items are then presented. As a result, the emotional attributes of a retrieval cue are (arguably) eliminated during test and there is no potential confound between the retrieval cues and the information to be retrieved. In keeping with this rationale, Liu et al. (2017) used emotional images and neutral words in an item-method directed forgetting task. During the study phase, neutral words were shown against a background of emotionally negative or neutral images. To ensure that participants paid attention to the words and the background images, they were instructed to click on the part of the image that they believed was related to the embedded word. Participants were instructed to either remember or forget the words. Recognition memory was superior for words studied in the neutral context. In addition, the directed forgetting effect was larger for words shown against a neutral background, and this was due to the fact that memory for TBR words experienced in a negative context was decreased compared to TBR words in a neutral context.

The authors suggested that if the negative images attracted relatively more attention than neutral images (Talmi et al., 2008), thereby limiting attention directed towards the neutrally embedded words, this outcome would come about. If attention towards the neutral words is limited, then presumably so is the extent to which these words are encoded, making it less challenging to forget than to remember these words. This argument can be seen as an example of 'attentional narrowing'; memory is worse for peripheral details when there is a central arousing/emotional item in a scene (Bradley et al., 1992; Mather & Knight, 2008), although in the manipulation used by Liu et al., (2017) the arousing/emotional item was peripheral and the central item was neutral. In another study, albeit without a directed forgetting manipulation, Maratos and Rugg (2001) used emotional and neutral sentences as the encoding context for neutral words. Memory was assessed via a recognition memory task. The encoding context did not influence the memorability of the neutral words, and they argued that this is perhaps because a recognition test is not sufficient to detect an effect of contextual manipulations, because recognition judgements can be based on familiarity as well as on recollection (Yonelinas & Jacoby, 1995). However, this does not explain the differences in memory according to the emotional encoding context observed in Liu et al.'s (2017), who used a recognition memory test. In a second experiment, Maratos and Rugg (2001) used an additional source memory test by instructing participants to make judgements on whether a word was presented in a negative or neutral context. Contrary to the outcomes in their first experiment, they observed enhanced memory for words from a negative context relative to a neutral context. In addition, they observed increased correct judgements for neutral contexts relative to negative contexts. They proposed that these discrepancies arise because participants prepared differently: they were expecting to make source judgements as well in the second experiment.

More broadly, the effects of emotion on source memory have been studied extensively and the findings are inconsistent (for a review of findings see: Mather, 2007; Yonelinas &

Ritchey, 2015). Negative valence has been shown to impair associative memory (Madan et al., 2012; Madan, Fujiwara, et al., 2017; Maratos & Rugg, 2001), whereas positive valence has been found to enhance it (Madan et al., 2019). Using a cued-recall task, Madan et al., (2012) examined the effect of negative arousal on associative memories of moderately arousing negative and neutral word-pairs. Participants were instructed to remember the word pairs and memory was tested in a cued-recall task. They observed an impairment of associative memory between word pairs due to negative arousal. On the contrary, in a free recall test, memory for negative arousing words was enhanced compared to neutral words. These findings were later replicated using images (Madan et al., 2017). Based on these findings they suggested that negative arousal impairs the encoding of associations, whereas it enhances the memory of an item itself (Madan et al., 2012; Madan et al., 2017).

In contrast to these findings, there are also reports of enhanced associative memory for negative material. Smith et al. (2004) used positive, negative and neutral images as background context and neutral images of objects that were superimposed on the background images. Participants were instructed to imagine an association between the objects and the images and were tested (via yes/no recognition) on their memory for the object images. In addition, after the recognition judgement participants were required to make a source judgement by indicating whether the background image was negative, positive or neutral or whether they didn't remember the valence of the image. They found superior memory for objects encoded in a positive context relative to a negative and neutral context. Moreover, the number of correct source judgements was higher for objects that were encoded in both the negative and positive context compared to the neutral context. In contrast to what was observed by Madan et al. (2012, 2017), therefore, under these circumstances associative memory is not impaired for negative material. These findings also diverge somewhat from those of Liu et al. (2017) who found that memory for neutral material associated with a negative context was impaired relative to material associated with a neutral context. One explanation for this disparity is that different valence contrasts were used (negative vs neutral in one case, and positive vs negative vs neutral in the other) and therefore the effect of encoding context differs. In a valence contrast in which negative, positive and neutral material is applied, the ratio between emotionally valenced and neutral material is unequal. Whereas, when, for example, negative and neutral material is contrasted the ratio is equal. In addition, the valence dimension is increased when using both negative and positive material intermixed with neutral material. These two differences in combination might explain why differences in the effects of emotion occur when using a different combination of valence categories.

Turning to considerations for the current experiment, manipulating the encoding context to potentially avoid the confound between the influence of emotional attributes across study and test also merits some additional considerations. One could argue that because the neutral words are embedded in different emotional encoding contexts, the emotional attributes of the image become associated with the words thereby influencing their perceived valence. If this was the case, then this experiment would be prone to the same confound it is designed to avoid. Mather and Knight (2008) suggested that items previously associated with emotional material attract more attention than items not previously associated with emotional material, which they referred to as the emotional harbinger effect. They conducted a series of experiments to investigate this and in each they created an emotional harbinger effect by presenting a neutral item simultaneously with an emotional item (either auditory or visual material in separate experiments) 16 times in total and found that memory for contextual details was impaired for neutral material that was previously associated with an emotional item. They proposed that the impaired memory was a result of changes in emotionality of the neutral items due to being associated with emotional material previously. Thus, according to this account, even though the item is neutral, when it is associated with emotional material, this changes the perceived emotionality. It is important to note, however, that there might be a differential effect when a neutral item is associated with emotional material only once (Liu et al., 2017; Smith et al., 2004) or 16 times (Mather & Knight, 2008). None-the-less, given this consideration, the experiment described below provides a means of assessing the emotional harbinger effect.

The objective for this experiment was to remain as close as possible to the experimental design in Liu et al.'s (2017) study and therefore a similar manipulation was implemented. In short, the item-method directed forgetting procedure was implemented using neutral words embedded in emotional and neutral images presented during encoding. Neutral words were presented during retrieval in an old/new recognition memory test. It might seem surprising that no measure of recollection and familiarity was included given the broad focus on these measures in this thesis. However, this was done to create broadly similar circumstances and task demands for the participants in this experiment and in Liu et al.'s (2017) study. Additionally, in response to the issues discussed above, an association memory test and a valence rating task were included. The intention behind the association memory test is based on arguments from Maratos and Rugg (2001) that recognition memory alone may not always be sufficient to detect an effect of contextual manipulations, in this case emotion. Moreover, in order to evaluate whether there is a case of the emotional harbinger effect – a change in perceived valence of words – a valence rating task is included. In line with previous findings, it was expected to find: (i) a directed forgetting effect, (ii) enhanced memory for words encoded in a neutral context due to attentional narrowing for words encoded in an emotional context, (iii) a diminished directed forgetting effect for words encoded in an emotional context, and (iv) enhanced associative memory for positive and neutral encoding contexts. Moreover, and perhaps most pertinently, if the consistent criterion change reported in this series is a consequence of effects of emotion operating over the valence of words at retrieval, then they will be absent in this experiment.

### 4.1.1 Methods

#### 4.1.1.1 Participants

These were 44 undergraduates and postgraduates (28 females, *M* = 24.30 years, *SD* = 3.68) from the University of Nottingham. After giving informed consent, participants completed the experiment and received an inconvenience allowance of £8.

#### 4.1.1.2 Design and Materials

A 2 (instruction: remember vs forget) x 3 (context valence: negative vs positive vs neutral) within-subjects design was employed. There were 360 neutral words and 180 images. The words were selected from the Warriner et al. (2013) database and were presented along with one of the three categories of valenced images.

Images were selected from the EmoMadrid affective picture database (http://www.uam.es/CEACO/EmoMadrid.htm; Carretié, Tapia, López-Martín, & Albert, 2019). Valence and arousal were measured via a rating scale from -2 (very negative/very calming) to 2 (very positive/very arousing). The images differed on the basis of valence (negative: *M* = -1.27, *SD* = 0.36; positive: *M* = 1.22, *SD* = 0.26; neutral: *M* = 0.18, *SD* = 0.20), and on the basis of arousal (negative: *M* = 1.08, *SD* = 0.31; positive: *M* = 1.03, *SD* = 0.40; neutral: *M* = -0.03, *SD* = 0.36) between emotional (negative and positive) and neutral images (see Appendix N). The images were matched for visual complexity, using measures of edge density and feature congestion (Madan et al., 2018) across all valence categories. At the level of the entire image sets independent sample t-tests were conducted to assess the equivalence of edge density (range 0.00 – 0.20) and feature congestion (range 1.57 – 10.38). Full details of these matching criteria and the outcomes of the analyses can be found in Appendix N.

The 360 words were split into 180 old words shown during the study phase and 180 new words shown at test phase together with the old words. Half of the words presented during the study phase were followed by an instruction to remember and half of them by an instruction to forget. This created six critical (background – instruction) conditions of interest; Negative-Remember, Negative-Forget, Positive-Remember, Positive-Forget, Neutral-Remember and Neutral-Forget. The pairing of words with the images and the assignment of remember and forget instructions were randomized.

#### 4.1.1.3 Procedure

##### *4.1.1.3.1 Study Phase*

Participants were tested individually, and the study lasted no more than 1 hour. The experiment consisted of one study-test block. In the study phase, 180 images were presented individually for 4500ms in the centre of a computer screen. After 1000ms of the 4500ms a fixation cross appeared in the middle of the image for 500ms, followed by a word that was presented for 500ms. After the word disappeared, the image remained on the screen for 2500ms during which time participants were instructed to imagine an association between the word and the image. The image was then replaced by a fixation cross for 500ms followed by a remember cue (VVVVV in the colour green) or a forget cue (XXXXX in the colour red). Participants were instructed to attempt to remember the preceding word following a remember cue, and to forget the word following a forget cue. The instructions remained on the screen for 1500ms and the order of remember and forget words was determined randomly for each participant. Following the offset of the cue, there was a blank screen for 500ms after which the next image appeared on the screen. See Figure 4.1 for an example of one trial.

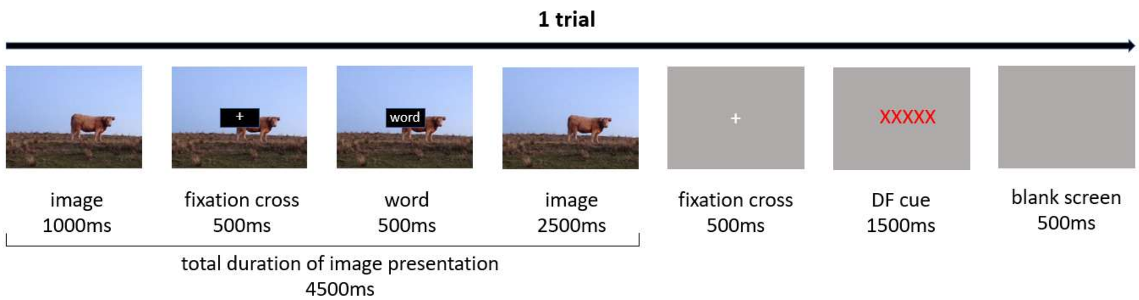



*Figure 4.1.* Experimental procedure for the study phase with and example trial of a word embedded in a neutral image and followed by a forget cue.

#### *4.1.1.3.2 Old/New Recognition Memory*

Test trials commenced with a fixation cross for 500ms, followed by a word for 500ms. Participants were asked to make an old/new recognition judgment to each word, regardless of the TBR or TBF instruction given in the study phase, by pressing designated keys with their left and right index fingers on a keyboard. Once a response was made, a blank screen appeared for 500ms, after which the next trial commenced with a fixation cross that was visible for 500ms.

#### *4.1.1.3.3 Association Memory*

After the old/new recognition test, a surprise associative memory test commenced with a fixation cross for 500ms at the centre of the screen. The fixation cross was then replaced with a word at the centre of the screen and 6 images – 3 images above the word and 3 below the word (see Figure 4.2 for an example). One of the images had been paired with the word at study. The other 5 (‘lure’ images) had not, but had been presented at study. Each image was presented 6 times during the associative memory test; once as the image paired with the word and the other 5 times as a lure image. Participants were instructed to select the image that was paired with the word on the screen in the previous study phase by using a computer mouse and clicking on the image. The images and the word remained on the screen until a response was made, which was then followed by a blank screen for 100ms. After the blank screen, a new trial commenced with a fixation cross. There were two test conditions. In the valence consistent condition, the valence of the lure images was consistent with the valence of the target image. In the valence mixed condition, there were 2 images of each valence category (negative, positive and neutral). An equal number of trials (90) were in the valence consistent and inconsistent condition.


*Figure 4.2.* One trial of the association memory test presenting a word in the middle with 6 images (3 above and 3 under the word).

#### *4.1.1.3.4 Valence Rating Task*

After the association memory test, participants were asked to rate how happy or unhappy the 360 words presented during the old/new recognition test made them feel. Each word was presented on the screen individually and participants were asked to respond via keyboard according to a 9-point Likert scale (*1 = extremely unhappy to 9 = extremely happy*). They were instructed to respond 1 when a word made them feel extremely unhappy, 9 when a word made them feel extremely happy and to use numbers in between to describe intermediate feelings.

When a word made them feel completely neutral, they were instructed to respond 5. This rating scale is similar to the rating scale used to collect valence ratings in the database that was used to select the words that were employed in this experiment (Warriner et al., 2013). Each word remained on the screen until a response was made.

### 4.1.2 Results

Table 4.1 shows mean probabilities of hits and false alarms as well as sensitivity and criterion across instruction (remember, forget) and background image valence (negative, positive, neutral). The proportions of correct source judgements in the association memory test are shown in Table 4.2, separated according to the consistent/inconsistent dimension.

Table 4.1
*Probabilities of Correct Judgements (Hits) to Old Items, Incorrect Judgements (False Alarms) to New Items and Estimates of Memory Sensitivity (d') and Criterion (c) for each Background Image Valence, across Instruction (Remember/Forget). SD = standard deviation*

| | Hits | | FA | *d'* | | c | |
|---|---|---|---|---|---|---|---|
| | Remember | Forget | | Remember | Forget | Remember | Forget |
| | M (SD) | M (SD) | M (SD) | M (SD) | M (SD) | M (SD) | M (SD) |
| Negative | 0.61 (0.15) | 0.55 (0.19) | | 1.35 (0.60) | 1.17 (0.60) | 0.35 (0.34) | 0.44 (0.41) |
| Positive | 0.59 (0.17) | 0.54 (0.16) | | 1.27 (0.53) | 1.14 (0.56) | 0.39 (0.38) | 0.46 (0.35) |
| Neutral | 0.57 (0.17) | 0.53 (0.17) | 0.17 (0.10) | 1.23 (0.58) | 1.11 (0.54) | 0.41 (0.36) | 0.47 (0.38) |

Table 4.2
*Probabilities of Correct Source Judgements across Instruction (Remember/Forget) and Source Condition (Consistent and Mixed) for Each Valence (Negative, Positive and Neutral) Category of the Background Images*

| | Remember | | Forget | |
|---|---|---|---|---|
| | Consistent | Mixed | Consistent | Mixed |
| | M (SD) | M (SD) | M (SD) | M (SD) |
| Negative | 0.60 (0.22) | 0.60 (0.24) | 0.54 (0.24) | 0.55 (0.21) |
| Positive | 0.58 (0.20) | 0.54 (0.24) | 0.50 (0.23) | 0.50 (0.23) |
| Neutral | 0.57 (0.25) | 0.62 (0.23) | 0.56 (0.21) | 0.58 (0.21) |

#### 4.1.2.1 Old/New Recognition Test

2 (instruction: remember vs forget) x 3 (valence: negative vs positive vs neutral) repeated measures ANOVAs were conducted separately on the measures of sensitivity (*d'*) and response criterion (c). Only effects of DF instruction were reliable (see Table 4.3). *d'* was superior for TBR than for TBF items, and there was a more liberal criterion for TBR items.

Table 4.3
*Summary of Repeated Measures ANOVA Results for Comparisons between Instruction and Valence for Memory Sensitivity and Criterion in the Old/New Recognition Memory Test and Comparisons between Instruction, Valence and Association (Consistent and Mixed Valence) for Memory Accuracy in the Association Memory Test*

| | Old/new recognition memory | | | | | | | | Association memory | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Memory sensitivity *(d')* | | | | Criterion (c) | | | | Accuracy | | | |
| Measure (DV) | *df* | *F* | *p* | $\eta_p^2$ | *df* | *F* | *p* | $\eta_p^2$ | *df* | *F* | *p* | $\eta_p^2$ |
| Instruction | 1, 43 | 10.88 | .002** | .20 | 1, 43 | 10.88 | .002** | .20 | 1, 43 | 11.09 | .002** | .21 |
| Valence | 2, 86 | 2.31 | .106 | .05 | 2, 86 | 2.31 | .106 | .05 | 2, 86 | 6.46 | .002** | .13 |
| Association | | | | | | | | | 1, 43 | 0.52 | .475 | .01 |
| Instruction*Valence | 2, 86 | 0.27 | .761 | .01 | 2, 86 | 0.27 | .761 | .01 | 2, 86 | 0.54 | .583 | .01 |
| Instruction*Association | | | | | | | | | 1, 43 | 0.09 | .771 | .00 |
| Valence*Association | | | | | | | | | 2, 86 | 2.09 | .130 | .05 |
| Instruction*Valence* Association | | | | | | | | | 2, 86 | 1.43 | .244 | .03 |

*Notes.* ** $p < .01$.
*df* = degrees of freedom

#### 4.1.2.2 Association Memory Test

A 2 (instruction: forget vs. remember) x 3 (valence: negative vs. positive vs. neutral) x 2 (association condition: consistent vs. mixed) repeated measures ANOVA revealed only main effects (see Table 4.3). Association memory was superior for images that were paired at study with words that participants were instructed to remember ($M$ = 0.59, $SD$ = 0.03) relative to instructions to forget ($M$ = 0.54, $SD$ = 0.03). The main effect for valence was followed up via t-tests. Memory was superior for negative ($M$ = 0.57, $SD$ = 0.19) and neutral ($M$ = 0.58, $SD$ =

0.20) images relative to positive ($M$ = 0.53, $SD$ = 0.19) images (negative vs positive: $t(43)$ = 2.74, $p$ = .009, $d$ = .41 and neutral vs positive: $t(43)$ = -3.37, $p$ = .002, $d$ = .50). There were no differences in association memory between negative and neutral images.

#### 4.1.2.3 Valence Rating Task

A one-way ANOVA with four levels (negative, positive, neutral context and new words) was conducted. A significant effect ($F$(2.68, 960,18) = 13.67, $p$ < .001, $\eta_p{}^2$ = .04) was followed up for all possible paired comparisons via t-tests using a Bonferroni adjusted alpha level of .008 (.05/6). When words were presented in a positive ($M$ = 5.36, $SD$ = 0.65) context they were rated more positively relative to when presented in a negative context ($M$ = 5.15, $SD$ = 0.64; t(359) = -4.68, $p$ < .001, $d$ = .25) or as new words ($M$ = 5.16, $SD$ = 0.40; t(359) = 5.63, $p$ < .001, $d$ = .31). Words presented in a neutral context ($M$ = 5.31, $SD$ = 0.67) were also rated more positively relative to when presented in a negative context (t(359) = -3.63, $p$ < .001, $d$ = .20) or as new words (t(359) = 4.30, $p$ < .001, $d$ = .23). There were no other reliable differences. There was, however, a positive correlation between the ratings for the new words ($M$ = 5.16, $SD$ = 0.40) and the ratings from the database ($M$ = 5.12, $SD$ = 0.31), $r$ (358) = .33, $p$ <.001.

In order to evaluate whether the valence of the images accompanying words at study influenced perceived valence of the words subsequently, t-tests were conducted on the means of the valence ratings for each participant using a Bonferroni adjusted alpha level of .008 (.05/6). While not differing from each other, words in a positive ($M$ = 5.34, $SD$ = 0.46) and neutral context ($M$ = 5.29, $SD$ = 0.44) were rated as being more positive than words in a negative context ($M$ = 5.17, $SD$ = 0.44; positive vs negative: t(43) = -2.87, $p$ = .006, $d$ = .42; neutral vs negative t(43) = -2.32, $p$ = .025, $d$ = .34). In comparison to new words, those studied in a positive or neutral context were rated more highly (new words: $M$ = 5.17, $SD$ = 0.41; positive vs new words: t(43) = 4.27, $p$ < .001, $d$ = .62; neutral vs new words: t(43) = 3.33, $p$ = .002, $d$ = .50).

### 4.1.3 Discussion

Memory was enhanced for TBR compared to TBF words. This directed forgetting (DF) effect was not influenced by the valence of the encoding context nor was the memorability of the words. This is in contrast with the findings reported by Liu et al. (2017): memory for words in a neutral context was enhanced compared to words in a negative context and the directed forgetting effect was diminished for words in a negative context, due to better memory for TBR words in a neutral context than in a negative one.

Methodological factors might explain the discrepancies between these findings. One difference between the experiments is the method used for creating an association between the words and the images. Liu et al. (2017) instructed their participants to click somewhere on the image that was related to the meaning of the word. This may have been an easier way for participants to create an association than the method used in the present experiment, where participants were merely instructed to imagine an association between the word and image and no action was needed. The degree to which the encoding context influences the encoding of a word may be linked with the strength of an association between the word and the encoding context. When a strong association has been made between an emotional image and neutral word, this may enhance memory for the neutral words. However, as argued by Liu et al. (2017), an emotional encoding context may attract enhanced attention leading to poorer encoding, which then results in lower memory for words encoded in an emotional encoding context. Thus, perhaps the degree to which an image is distracting may also be different between Liu et al.'s (2017) study and those in this experiment.

Perhaps more important is the degree of visual complexity of the images used in the two experiments. Images that are high in arousal (negative and positive images) are commonly also higher in visual complexity (Madan et al., 2018). Matching for complexity was achieved in this experiment, but this was not included in Liu et al.'s (2017) design. In their study, it might be that their outcomes arise because of increased visual complexity of the negative images they employed. Because visual complexity was matched across negative, positive and neutral images in this experiment any outcomes obtained are not subject to this possible confound. This is a good argument for weighting heavily the findings in this experiment relative to those reported by Liu et al. (2017).

As described earlier, Maratos and Rugg (2001), as well as many others (Yonelinas, 2002), suggested that recognition memory is not the most sensitive means of detecting effects of encoding manipulations that primarily influence recollection. Using negative and neutral sentences, they examined the effect of the encoding context on neutral words using a recognition memory task (i.e. Experiment 1) and a combination of a recognition/source memory task (i.e. Experiment 2). In their second experiment they observed enhanced memory for words encoded in a negative context relative to words encoded in a neutral context. This effect was absent in their first experiment and they ascribed these discrepancies to differences in task demands: because participants anticipated making source judgements in the second experiment, this may have resulted in the encoding context having a larger effect on the encoding processes of the neutral words embedded in these sentences. Linking these observations to the findings in the present experiment, this would explain why there was no effect of encoding context on memory because the association memory test was incidental. If participants had been aware of a subsequent source memory test, the encoding context might have affected the memorability of the neutral words. However, Smith et al. (2004) observed superior memory for objects encoded in a positive context relative to a negative and neutral context regardless of whether the source memory task was intentional or incidental. Perhaps other factors such as methodological elements (e.g. control of stimulus properties) or the manipulation of the encoding context (e.g. sentences or images) are at the root of these discrepancies. Furthermore, the findings reported here don't seem to be in line with the attentional narrowing effect (Mather & Knight, 2008), since the encoding context did not affect the memorability of the words.

Even though emotion did not affect the memorability of words, it did influence the word-context associations. Identification of the images associated with words was superior for negative and neutral images relative to positive images. The direction of the effect of emotion is the opposite to what has been reported in some previous studies. In some other studies negative emotion has tended to impair associative memory (Madan et al., 2012; Madan et al., 2017; Maratos & Rugg, 2001; but see Smith et al., 2004), while positive emotion has enhanced it (Madan et al., 2019; Pierce & Kensinger, 2011; Smith et al., 2004). For negative emotion it has been proposed that it is more challenging to integrate a neutral item with negative items, because people tend to focus their attention more towards the negative item (Otani et al., 2012; Pierce & Kensinger, 2011). Positive emotion has been proposed to broaden attention and cognitive processing, making it easier to integrate positive and neutral material, also referred to as attention broadening (Madan et al., 2019; Pierce & Kensinger, 2011). Taken together, these accounts predict an impairment of word-context associations for negative images and an enhancement for positive images. Linking these to the present findings, there seems to be a case of attentional narrowing (Mather & Knight, 2008) for positive images – the positive images attract enhanced attention leading to a decreased attention to the neutral word. This may have caused an impairment in creating word-context associations, leading to decreased memory for positive images. However, this is not consistent with the attention broadening argument and with previous findings (Madan et al., 2019; Smith et al., 2004). The findings reported here are also inconsistent with other studies using a similar encoding context manipulation (Maratos & Rugg, 2001; Smith et al., 2004). For example, in Smith et al.'s (2004) study, memory for contextual details was enhanced for negative and positive contexts relative to a neutral context. There are, however, differences in task demands between these studies and the present experiment. Smith et al. (2004) required participants to make a judgement on the valence of the encoding context, whereas participants in this experiment were required to indicate the specific image associated with a word. Thus, besides making a judgement on the valence of the context, participants were required to make an attempt to retrieve the specific image.

Another explanation for poorer memory for word-context associations for positive images stems from considerations regarding limitations in creating associations between stimuli. According to Mather (2007) emotion could impair the creation of associations between an emotional and neutral stimulus due to an emotional stimulus attracting more attention. As a result, this impairment decreases the likelihood of associating an emotional stimuli with a neutral one. Mather (2007) further suggested that an enhanced effect of emotion on the creation of associations is only seen when it is targeted at features of an emotional item itself. This account does not explain, however, the differences between negative and positive contexts reported in this experiment. These differences (the effects of negative and positive emotion on memory for word-context associations) suggest that arousal is not the factor responsible for these effects. Rather, valence is more likely to be implicated. A more direct comparison between negative and positive word-context associations, as was done in Experiment 2 and 3, would provide a more comprehensive answer to the question of whether arousal and/or valence influence word-context associations.

There was also a directed forgetting effect for the images in this experiment. Comparable DF effects have been reported previously (Bancroft et al., 2013; Hockley et al., 2016; Wang et al., 2016), although in these studies the DF instructions were direct. For example, in the study by Bancroft et al. (2013), in an item-method DF design, participants were directly instructed to apply the DF instructions (i.e. remember or forget) to word pairs that were presented. Participants were to create associations between word pairs when instructed to remember while not doing so when instructed to forget. Bancroft et al. (2013) observed a DF effect for the word pairs. In the present study the DF effect, by contrast, occurred despite the fact that the instructions to remember/forget were directed only at the words and participants were not instructed to either remember or forget the word-image associations. In addition, the DF instructions were not an indicator of whether to create an association or not. In another study (Burgess, Hockley & Hourihan, 2017) an encoding context structure similar to the present experiment (albeit without manipulating emotion) was used. Words were presented against neutral images and participants were instructed to only remember or forget the words. Although they found a DF effect for the words, they did not observe a DF effect for the images. In contrast to the design of present experiment, they did not instruct participants to create an association between the words and the images, which could explain the different outcomes.

The absence of an effect of emotion on the memorability of the words and the directed forgetting effect for words is consistent with findings in the previous experiments in this thesis. The motivation for this experiment was to disentangle what happens during encoding and retrieval and to investigate in what way emotion might influence memory control. This was done by removing – or at least attempting to minimise – the emotional attributes affiliated with retrieval cues. In light of this manipulation, the fact that response criterion was not different for words that were encoded in an emotional or neutral context is important: an effect of emotion on criterion has been observed consistently in previous experiments in this thesis. This raises the possibility that it is the emotional attributes of retrieval cues that results in participants adopting a more liberal criterion for emotional material. One explanation for a more liberal response criterion for emotional contents is that it results from an increased sense of familiarity for emotional relative to neutral material that is generated simply by a test item (Kensinger & Corkin, 2003; Kensinger & Kark, 2018). This would then result in an increased likelihood of responding old to emotional material (for a similar explanation in other contexts, see discussions of the revelation effect (Azimian-Faridani & Wilding, 2004; Lecompte, 1995;

Watkins & Peynircioglu, 1990)). The fact that the changes in criterion in earlier studies were based on an increase in the likelihood of new as well as old test items is consistent with this account. By this view, removing the emotional attributes from the retrieval cue resulted in the absence of the changes in criterion with emotion that have been reported consistently throughout this thesis up to this point.

Turning to the outcomes in the valence rating task, there were reliable, although small, differences between ratings. When words were encoded in a positive and neutral context they were rated as more positive than when they were encoded in a negative context or when they had not been shown at study. This suggests that, at least in the positive and neutral context, words changed in perceived emotionality. Given this outcome, is it surprising that no changes in criterion were observed? It might be argued that this outcome stands in opposition to the view that the reason for the null result for criterion change with emotion is because only neutral words were presented at test. If the study encounter in any way changed the emotional attributes of the words when presented at test it might be seen in the criterion data. These effects on the valence task are small, however, and were revealed in a direct assessment. These considerations likely form part of an explanation for any disconnect between these two sets of outcomes in this experiment.

To summarise, by employing a manipulation of emotion at study only, this experiment design offered a means of assessing the influence of emotion on memory and memory control at different stages (encoding and retrieval). The encoding manipulation did not influence memory for the words, nor did it influence the ability to remember and forget the words. This contrasts with what was been predicted based on prior studies. However, as per the explanations outlined in this discussion other factors (such as the control of visual complexity) and other experimental differences, most notably the method used to create associations between words and images, may have led to these inconsistent outcomes. The outcomes do at least suggest that again when controlling for critical potential confounds, an emotional manipulation during encoding does not influence directed forgetting. Moreover, in contrast, emotion did influence memory for background images, and there was also a DF effect for the images (even though the R/F instructions were directed to the words only).

Perhaps the most significant finding in this experiment is that response criterion was not affected by the emotionality of the encoding context. The pre-experiment considerations anticipated that a null outcome here – no change in criterion with emotion – would argue for processes at retrieval (rather than those at encoding) being important. This outcome in this experiment is important, because of the consistent finding of a criterion change in prior studies. The outcomes in this experiment suggest that the emotional properties of retrieval cues are the primary driver for changes in response criterion with emotion. There are caveats, however: the reliance on a null result and the absence of a direct contrast at the time of retrieval and within the same experiment between conditions when emotion does/does not vary. These issues are picked up further in the General Discussion that follows.

# CHAPTER 5

# 5. GENERAL DISCUSSION

## 5.1 Objectives and Experimental Design

The overall objective of the experiments described in this thesis was to develop an understanding of when control over remembering and forgetting of emotional material can be exerted and how emotion influences memory judgements. The links between cognitive control and memory have been investigated extensively (Anderson & Green, 2001; Levy & Anderson, 2002; Rizio & Dennis, 2013), however there is little consensus about the linkage between emotion and control (Chen et al., 2012; Dehli & Brennen, 2008; Depue et al., 2006; Gallant et al., 2017, 2018; Hauer & Wessel, 2006). Of particular importance for the work here, it is unclear whether emotion influences the ability to exert control over remembering and forgetting emotional material (Barnier et al., 2004; Berger et al., 2018; Dehli & Brennen, 2008; Otani et al., 2012; Taylor et al., 2018; Yang et al., 2016).

As described in the Introduction (see section 1.4.6.1) and in several discussions throughout this thesis, published findings of directed forgetting and emotion are inconsistent (Bailey & Chapman, 2012; Brandt et al., 2013; Gallant & Yang, 2014; Hauswald et al., 2011; McNally et al., 1998; Myers et al., 1998; Payne & Corrigan, 2007; Wessel & Merckelbach, 2006).

In the DF item-method, which was used consistently in this thesis, participants receive a list of study items and are instructed after each item to either remember or forget it (MacLeod, 1975, 1999; Muther, 1965). In a subsequent test, their memory for all items is tested. The commonly observed DF effect is superior memory for items that participants were instructed to remember relative to those that they were instructed to forget. A combination of processes is proposed to be responsible for the DF effect in the item-method, which are encapsulated in two primary accounts; the selective rehearsal account (Bjork, 1972; MacLeod, 1975) and the active inhibition account (Anderson & Hanslmayr, 2014; Basden et al., 1993). According to the selective rehearsal account TBR items are rehearsed and encoded more extensively than TBF items. According to the active inhibition account, inhibition processes are activated when instructed to forget, which leads to relatively better memory for TBR items. The accuracy of either account is difficult to establish because of the marked heterogeneity in published findings, and the experiments in this thesis were not designed to distinguish between these accounts (although see the discussions of the ERP encoding analyses described below and in section 3.3.4). The experiments were designed, however, to contribute to an understanding of how emotion, memory and memory control interact, by exerting a consistent set of controls over stimulus properties. These are described below, following a brief recapitulation of other matters addressed in these experiments: the separation between criterion and sensitivity, the links between emotion, control, recollection and familiarity, and the similarities and differences between interactions between positive and negative valence and memory.

### 5.1.1 Word and Word Set Properties

One important factor is stimulus properties, which are a potential confound if they differ with valence (Kensinger & Kark, 2018). One critical property is semantic relatedness, which tends to be greater for sets of emotional words compared to sets of neutral words (Buchanan et al., 2006). Semantic relatedness refers to the strength of semantic relationships between words, and can be assessed in several ways (Landauer et al., 1998; Mandera et al., 2017; Nelson et al., 2004). It has been shown that when a list of words has a high degree of semantic relatedness this can facilitate memory, because relatedness may improve accessibility of memory (Tulving & Pearlstone, 1966). This general point has also been made specifically for emotional words (Buchanan, 2007; Maratos et al., 2000; Talmi & Moscovitch, 2004). These considerations suggest that it could be easier to remember, and (perhaps) harder to forget, emotional words, unless semantic relatedness is equated between word sets. Dougal and Rotello (2007) examined memory for emotional and neutral words using a Remember/Know (R/K) task. They matched the words for semantic relatedness and found no differences in memory sensitivity between emotional and neutral words. In addition, although no differences were observed for Know responses, Remember responses were increased for negative relative to positive and neutral words. These findings suggest that relatedness is an important consideration when interpreting findings in studies where memory for emotional material is examined, but different observations of the effect of emotion on memory have been reported elsewhere. For example, Minnema and Knowlton (2008) found enhanced memory for negative relative to positive and neutral words despite controlling for semantic relatedness.

In addition to controlling for relatedness, several other factors were equalised across valence categories in the experiments in this thesis. These were: word length and frequency across all three valence categories, and arousal between negative and positive words in Experiments 1-4 and visual complexity and arousal between negative and positive images in Experiment 5. Because of these attempts to exert the same degree of control over stimulus properties in word sets (and images) across all of the experiments in this thesis, it is possible to have a level of confidence in similarities and differences across the experiments that is not possible when contrasting other work where different levels and kinds of control have been exercised.

### 5.1.2 Measurement and Design

Differences in sensitivity as measured by d-prime (*d'*) might not be accurate when response criterion varies and single-point measures are used: there is a potential confound between measures of *d'* and response criterion, when response criterion varies across conditions. In this case, *d'* can be meaningfully compared only when the variances of the old and new items distributions are equal (Dougal & Rotello, 2007; Ratcliff, Sheu, & Gronlund, 1992; Yonelinas, 1994). However, empirically it is not typically the case that the distributions are comparable (Egan 1958, 1975; Ratcliff et al., 1992; Wixted, 2007; for a review of findings see Yonelinas, 1994), which makes it challenging to interpret changes in memory sensitivity across conditions. This issue has been addressed and discussed at length by Dougal and Rotello (2007), and addressed in this thesis by introducing confidence judgements at test to plot receiver operating characteristic (ROC) curves. ROCs provide an opportunity to separate estimates of sensitivity and criterion (Yonelinas, 1994; Yonelinas et al., 1996; Yonelinas & Parks, 2007), thereby offering a more accurate account of sensitivity measures when criterion changes across conditions. This is especially crucial when examining the effects of emotion on memory and memory control, because criterion tends to be more liberal for emotional material relative to neutral material (Bailey & Chapman, 2012; Dougal & Rotello, 2007; Kapucu et al., 2008; Ochsner, 2000; Windmann & Kutas, 2001).

Only a subset of directed forgetting (DF) studies have reported response criterion and/or focused on the importance of the role it may play in memory for emotional material (Bailey & Chapman, 2012; Berger et al., 2018; Gallant et al., 2018; Hauswald et al., 2011; Marchewka et al., 2016; Yang et al., 2012). In general, these studies have reported a relatively more liberal response criterion for emotional material relative to neutral material. However, with respect to the effects of emotion on memory sensitivity, the findings are inconsistent. Where some have reported equivalent levels of memory sensitivity for emotional and neutral material (Gallant et al., 2018; Hauswald et al., 2011; Yang et al., 2012), others have reported superior memory for neutral relative to emotional material (Bailey & Chapman, 2012; Berger et al., 2018) or superior memory for emotional relative to neutral material (Marchewka et al., 2016).

Besides disentangling measures of sensitivity and criterion, ROCs also allow extraction of estimates of the contributions of recollection (R) and familiarity (F) to recognition judgements (Koen et al., 2017; Yonelinas et al., 1996). Estimates of recollection and familiarity were derived from ROCs in Experiment 2. There are different means of acquiring estimates of recollection and familiarity and overlapping but not identical assumptions are employed in different cases. In order to assess the generality of the findings in Experiment 2, different ways of achieving this were employed in subsequent experiments. In Experiment 3 the R/K procedure was used, which offers a subjective means of estimating recollection and familiarity. In Experiment 4 event-related potentials (ERPs) were employed. The reason for doing this was to examine the contributions of recollection and familiarity in a task requiring only old/new recognition memory judgements (Rugg et al., 1998; Rugg & Curran, 2007; Sanquist et al., 1980). Changing task demands may influence in what manner participants approach a task and rely on recollection and familiarity. For example, old/new recognition judgements are assumed to rely heavily on familiarity (Johansson et al., 2004; Tulving, 1985b).

### 5.1.3 Valence Manipulations

The experiments described here involved three valence categories: negative, positive and neutral. The rationale for using both negative and positive material is partly because of some evidence that negative and positive valence have different effects on memory (Kensinger & Kark, 2018), and memory control (e.g. Li et al., 2017; Minnema & Knowlton, 2008; Otani et al., 2012). Superior memory for negative material relative to positive and neutral material has been reported (Minnema & Knowlton, 2008; Otani et al., 2012), although other data show superior memory for both negative and positive material relative to neutral material (Gallant & Yang, 2014; Li et al., 2017; Payne & Corrigan, 2007). With respect to memory control, some researchers have reported smaller DF effects for negative relative to positive material (e.g.

Minnema & Knowlton, 2008; Otani et al., 2012), whereas others have reported equivalent DF effects for these two valence categories (e.g. Bailey & Chapman, 2012; Payne & Corrigan, 2007). The experiments in this thesis offer a way to provide additional data that are germane to this question. If changes in memory are equivalent for positive and negative material relevant to neutral material, then a parsimonious explanation is that arousal is a substantive contributor to the observed effects (Buchanan et al., 2006; Cahill & McGaugh, 1995; Ochsner, 2000; Phelps & Sharot, 2008; Sharot et al., 2004). Different effects for positive and negative materials cannot be accommodated in this way.

Despite the evidence that is suggestive of differential effects of negative and positive material on memory (Kensinger & Kark, 2018; Li et al., 2017; Minnema & Knowlton, 2008; Otani et al., 2012), the majority of studies have focused on memory differences between negative and neutral material (e.g. Kensinger & Corkin, 2003; Xie et al., 2018; Yonelinas & Ritchey, 2015). One reason for this focus arises via a clinical perspective, as understanding the mechanisms of memory and how negative memories interrupt people's daily functioning is of practical importance (Harvey et al., 1998; Kuyken & Dalgleish, 1995; Williams & Broadbent, 1986). Disorders such as post-traumatic stress disorder (PTSD) are linked with failures to regulate memory for emotionally challenging events (Elzinga & Bremner, 2002). In short, the general finding is a diminished DF effect for negative material relative to neutral material for participants with depressive tendencies (Xie et al., 2018) or that suffer from PTSD (McNally et al., 1998). These outcomes reinforce the importance of understanding the effects of emotion on memory control in clinical populations (McNally et al., 2004; McNally et al., 1998), but do not take away from the argument that the influence of positive valence on memory is also worthy of further investigation.

### 5.1.4 Study/Test Separations

In the final experiment reported in this thesis an attempt was also made to separate effects of emotion at study and at test. The issue here is that when emotional material is used at study and test, distinguishing the effect of emotion on encoding and retrieval processes separately is difficult (Liu et al., 2017; Maratos & Rugg, 2001; Smith et al., 2004). For retrieval, the emotional attributes of a retrieval cue may trigger a sense of familiarity, which may influence memory judgements despite not being related to the presentation of the material in a prior study phase (Dougal & Rotello, 2007; Maratos & Rugg, 2001). For encoding, there are several arguments for how and why materials with different valences might be subject to different operations (Dolan et al., 1999; Kensinger & Corkin, 2004). A solution to this confound is to have an emotion manipulation only at one of the two stages. In line with the work of Liu et al. (2017) (see also Maratos & Rugg, 2001; Smith et al., 2004), in the final experiment reported in this thesis, this was achieved by manipulating the emotional encoding context at study in which neutral words were shown. In a subsequent recognition memory task, memory for the neutral words was tested.

### 5.1.5 Potential Confounds

#### 5.1.5.1 Multiple study-test cycles

The approach taken for the experimental designs in this thesis was to implement several study-test cycles. This increased the number of items per critical condition (particularly important for the ERP studies) and enabled maintenance of a level of memory performance sufficient to permit changes in sensitivity with experiment manipulations to be observed. Of course, one concern with this approach is the possibility that participants did not adhere to the directed forgetting instructions after the first study-test cycle.

Can we be confident then that the directed forgetting effect observed in this thesis consistently across experiments is robust? Several factors suggest that we can. First, the directed forgetting effect is observed consistently across experiments, which in and of itself suggests a robust effect. Second, participants were aware at the start of the experiment that memory would be tested for both TBR and TBF items, so the first study-test cycles is comparable to the others to some extent. Finally, analyses of the directed forgetting effect between the study-test cycles, in Experiment 1 (see Table 2.5), suggest that participants were performing in a similar manner: there were no reliable differences according to block.

As discussed in the introduction (see section 1.3.3.1), another possibility is that participants experience reduced motivation to retrieve TBF items (MacLeod, 1999), and this might change over multiple study-test cycles. Two observations are relevant here. First, MacLeod (1999) examined whether the motivation of participants influenced the directed forgetting effect by implementing a reward system for additional TBF items retrieved in a second memory test. A directed forgetting effect remained evident. Second, the results from block analyses in this thesis (no differences according to block) suggest that whatever participants were doing, they were adhering to the directed forgetting instructions.

#### 5.1.5.2 Gender differences

The majority of the participants across the experiments discussed in this thesis were females. This raises the possibility that gender differences are in play when examining emotion and memory control. Based on prior literature, for example, memory for details is enhanced in females whereas males tend to remember the 'gist' of events (Loprinzi & Frith, 2018). In addition, females tend to describe their memories using emotional terms more than males do (Loprinzi & Frith, 2018). Linking this with the effects of valence on memory, details of negative information tend to be better remembered than for positive and neutral material. In contrast, it is the general scope of positive information which is often better remembered (Bowen et al., 2018; Kensinger & Kark, 2018; Kensinger & Schacter, 2006). Taken together, these considerations raise the possibility that memory control for emotional material is different between males and females, and some of the results in this thesis are gender-specific. To the best of my knowledge, gender differences in directed forgetting for emotional materials has not been investigated. In addition, the majority of published studies report a broadly similar set of participants to those in the experiments in this thesis, although there have been some studies that included either a balanced set of participants (Hauswald et al., 2011; Nowicka et al., 2011; Xie et al., 2018; W. Yang et al., 2012), or a majority of males (Liu et al., 2017). Based on the unbalanced set of participants in this suite of experiments it may well be that the results presented in this thesis do not generalize equally to individuals identifying as men and women.

## 5.2 Summary of Findings and their Implications

### 5.2.1 Effects of Instruction and Valence

In all experiments, memory was superior for words that participants were instructed to remember compared to words that they were instructed to forget. This is consistent with the modal finding in directed forgetting studies (MacLeod, 1975, 1999). Strikingly, in none of the experiments in this thesis, did the directed forgetting effect vary with valence. These consistent outcomes are in line with some previous findings (Gallant et al., 2018; Gallant & Yang, 2014; Wessel & Merckelbach, 2006), but not with others (Bailey & Chapman, 2012; Berger et al., 2018; Hauswald et al., 2011; Payne & Corrigan, 2007; Taylor et al., 2018; but also see Brandt et al., 2013). Considering the consistency of this observation throughout this thesis, this strongly suggests that confounding factors may explain at least some of the inconsistencies observed in the literature. Theoretically, these findings also reject the assumption that, at least under certain circumstances, the ability to control memory for emotional material is not different from neutral material. This gives some indication that, clinically, when people experience pathological patterns of memory control this is not explained by simply an increased difficulty of forgetting. However, some explanation might be found in the direction of the quality of processing memories for negatively challenging events (Brewin, 2018). These are merely speculations, which based on the work presented here are not possible to either confirm or reject.

In Experiment 2 and 3 there was also a directed forgetting effect for recollection and familiarity. Previously in this thesis the prediction was made that memory control would be specifically difficult for recollection and not necessarily for familiarity. This prediction was made based on the assumption that under certain circumstances people with PTSD may re-experience an event when remembering occurs. Meaning that some recollection of contextual details is occurring and there is an inability to distinguish memories from the actual presence. This is of course when the inability to control memory becomes pathological, but it provides some indication that perhaps recollection is specifically difficult and necessary to control memory. However, these assumptions are purely speculative and extended research is necessary to understand memory control for recognition memory. Moreover, the findings in this study do not support this rationale. In contrast, the findings here strongly suggest that (i) the ability to control memory is equal for recollection and familiarity and (ii) memory control occurs for both recollection and familiarity.

The lack of an effect of emotion on directed forgetting may be due to the control of semantic relatedness that was exerted in these experiments. Minnema and Knowlton (2008) controlled for semantic relatedness and observed a smaller DF effect for negative words relative to positive and neutral words. Minnema and Knowlton (2008), however, used free association norms to measure the degree of relatedness, and ensured that any related words were put in different lists. The association norms were collected from subjective judgements from participants (Nelson et al., 2004). For the word lists employed in this thesis semantic relatedness was measured using an online tool ‘snaut’, which measures the degree of relatedness based on computational prediction models. These models predict a word, given the context (associated) words and calculate the semantic distance between word pairs (Mandera et al., 2017). The degree of relatedness was then matched between valence categories.

A limitation of the free association norms approach is that it is defined only for a subset of words, whereas this is not the case for the ‘snaut’ tool. Additionally, according to a comparison of various methods used for the measurement of semantic relatedness, the computational prediction model ‘snaut’ (together with Latent Semantic Analysis (LSA; Landauer et al., 1998)) is a superior method to measure the degree of relatedness compared to the human association norms method (Mandera et al., 2017). In addition, it is worth emphasising the consistency of the findings here: the absence of an effect of emotion on directed forgetting was observed in all experiments reported in this thesis.

There are also, however, examples of the absence of DF effects when relatedness has not been controlled for. Berger et al. (2018), for example (see also Gallant et al., 2018; Gallant & Yang, 2014; Taylor et al., 2018), used negative, positive and neutral words in a DF study and did not report any procedures involving the control of the level of semantic relatedness between emotional and neutral words. They did not find any effects of emotion on the DF effect. Of course, there is a possibility that, in this experiment and other similar ones, there are in fact no differences in relatedness according to valence.

Superior memory for emotional material was found in Experiment 1, which is consistent with findings in some other DF studies (Bailey & Chapman, 2012; Brandt et al., 2013; Hauswald et al., 2011; Liu et al., 2017; McNally et al., 1998; Myers et al., 1998; Payne & Corrigan, 2007), as well as in a variety of other paradigms in which memory has been assessed (Dougal & Rotello, 2007; Kensinger & Corkin, 2003; Talmi et al., 2007, 2008). This consistency, however, does not extend to some other studies in which memory was not enhanced for emotional material (Wessel & Merckelbach, 2006; Windmann & Kutas, 2001). Perhaps most importantly, and as anticipated earlier, there were differences in criterion with valence in Experiment 1 (discussed in greater detail below). In Experiment 2, there were three studies where paired valence categories were used (confidence judgements were also required), and sensitivity did not vary between emotional and neutral material. These outcomes are consistent with those of Dougal and Rotello (2007), as are the findings in two of the three studies in Experiment 3. In these studies, the R/K procedure was employed.

In the only divergence from this broadly consistent pattern, memory was superior for neutral relative to positive words in Experiment 3, and there was a trend in the same direction in Experiment 2, although this was not significant. This means that under both ROC and R/K methods, memory was somewhat superior for neutral relative to positive material. One reliable consistency, however, between these two experiments is the observed superior memory for negative relative to positive material. This observation may be a result of attentional biases being more pronounced for negative material for the purpose of attending to threatening stimuli in the environment (Kensinger & Corkin, 2004; Ochsner, 2000). In addition, negative material may benefit from prioritization of processes (Kensinger & Corkin, 2004), which has been referred to as natural selective attention (Dolan et al., 1999). Taken this together, this may lead to superior memory for negative material relative to positive material.

In Experiment 4 (where ERPs were acquired) and Experiment 5 (where emotion was manipulated at study only) there were no differences in sensitivity with valence. The consistency of Experiment 5 with other outcomes is arguably unsurprising. For Experiment 4, the most directly relevant contrast is with Experiment 1, because in both cases only recognition memory judgements were required. To re-iterate, in Experiment 1 sensitivity was superior for emotional relative to neutral material. One way of explaining the differences across the experiments is to note the sensitivity/criterion confound. This is also emphasized by the discrepancies in findings for response criterion. In Experiment 1, there was a relatively more conservative response criterion for positive relative to neutral and no differences between negative and neutral material. In Experiment 4 response criterion was liberal for emotional relative to neutral material.

Perhaps of equal or greater importance, the number of valence categories used in the two cases might be responsible for the different findings across experiments. In Experiment 1, an equal number of negative, positive and neutral words were used. The ratio between emotional and neutral words is therefore unequal, which is not the case in Experiment 4 (in each of the three studies – the two primary ones and the replication – only two valence categories were used). This suggests one explanation for the different findings across the experiments. In contrast to Experiment 1 (and Experiment 5, which is somewhat different in other ways), in Experiments 2, 3 and 4 only negative or positive words were contrasted with neutral words or with each other in paired contrasts.

How is this relevant to the design differences across experiments in this thesis? When it is the case that two-thirds of the words are emotional compared to one-third neutral words (Experiment 1), the emotional material may be relatively more salient. This may lead to emotional material receiving greater attention thereby resulting in memory enhancement for emotional material (Dewhurst & Parry, 2000), as was observed in Experiment 1. Another possibility is that due to a greater range in valence (the distance in the valence dimension between negative and positive) in Experiment 1, making valence an important dimension that influences memory, there is a greater effect of emotion on memory. In the case of the experiments described in this thesis, this is observed in Experiment 1.

### 5.2.2 Response Criterion

A consistent finding throughout this thesis, except in Experiment 1 and 5, was a more liberal response criterion for emotional material relative to neutral material. What considerations are relevant when looking to explain these outcomes? One answer, articulated earlier, is that emotional materials are associated with higher levels of familiarity (Kensinger & Corkin, 2003; Ochsner, 2000). If this is the case for old as well as new items, then it can manifest as differences in criterion with no change in memory sensitivity (Dougal & Rotello, 2007). This would be the case if the familiarity changes came about at study, test, or at both time points.

One way in which this possibility has been developed is to argue that emotional materials tend to be more interrelated, and as a result encourage more categorical, gist-based thinking (White et al., 2015; Windmann & Kutas, 2001). Gist has been identified as one basis for memory judgements (Budson, Todman, & Schacter, 2014; Koutstaal & Schacter, 1997; Reyna & Brainerd, 1995; Schacter, Israel, & Racine, 1999; Schacter et al., 1998; Schacter & Slotnick, 2004). In addition, at least for negative material, of relevance are considerations that an adaptive cognitive function is to prioritise potentially harmful events or information (Johansson et al., 2004; Windmann & Kutas, 2001). If negative emotional materials fit in this category, their level of salience might be amplified. If salience leads to an increased sense of familiarity (Kensinger & Corkin, 2003; Lavoie & O'Connor, 2013; Talmi et al., 2008), this can also explain criterion changes according to emotion.

Another possibility is that the differences in criterion do not come about because of differences in familiarity, but genuinely differences in criterion. When test items are intermixed this might seem implausible, but perhaps the salience argument above is relevant. If salient material is particularly relevant, then it might be beneficial to judge that material to be old. These salience considerations also draw support from the fact that there were also differences in response criterion between negative and positive valence when these two were contrasted together. Response criterion was more liberal for negative relative to positive words in Experiments 2 and 3. These outcomes can be accommodated by a salience account, and similar effects have been observed by Bailey and Chapman (2012).

The more liberal criterion for emotional material is consistent with what has been observed in a subset of DF studies (Bailey & Chapman, 2012; Berger et al., 2018; Gallant et al., 2018; Hauswald et al., 2011; Marchewka et al., 2016; Yang et al., 2012), and with other studies outside the DF literature (e.g. Dougal & Rotello, 2007). A rather unexpected finding, however, in Experiment 1, contrasting negative, positive and neutral words, was a relatively more liberal response criterion for neutral words relative to positive words. Moreover, there were no differences between negative and neutral words. As discussed above, one reason for these discrepancies might be the sensitivity/criterion confound, which is the difficulty of interpreting sensitivity measures when response criterion changes across condition and single point measures only are employed (Dougal & Rotello, 2007).

In what might be one of the most substantive findings in this thesis, in Experiment 5 there were no differences in response criterion across emotional and neutral conditions. The rationale for Experiment 5 was to remove the emotional attributes of the retrieval cue in order to separate effects of emotion on encoding and retrieval processes. The null result for response criterion in this case suggests that what happens during retrieval processing is what matters when it comes to criterion changes for emotional material. One of the arguments developed above is that familiarity changes are responsible for the differences in estimates of criterion across valence categories. If these familiarity changes were products of the retrieval cues themselves, then the outcome of Experiment 5 makes sense, in so far as the neutral words (all of the test stimuli) did not elicit differences of this kind.

In the most similar published study, Liu et al. (2017) found a diminished DF effect for words embedded in a negative context. This differs from the findings in this experiment and they did not report effects of emotion on response criterion. Based on a hits and false alarm calculation for criterion from their data, however, there is the suggestion of a relatively more liberal criterion for words in a neutral (0.28) relative to a negative (0.33) context. This pattern of data – and of course this is only an observation – is the opposite to what is observed in comparison to Experiment 5 (see Table 4.1). One explanation for the inconsistent findings in sensitivity is that whereas Liu et al. (2017) did not include matching the images for visual complexity, in Experiment 5 the negative, positive and neutral images were matched on visual complexity. As described earlier the degree of complexity differs between emotional and neutral images and may therefore influence the effects of emotion on memory.

#### 5.2.2.1 Directed Forgetting and Criterion

Participants also adopted a more liberal response criterion for words that they were instructed to remember relative to words that they were instructed to forget. This was observed consistently across the 5 reported experiments. White et al. (2015) proposed that the extent to which items in a list can be separated into categories, for example emotional or neutral material, influences response criterion (see Introduction, section 1.4.3). They only found differences in response criterion when the proportion of words that could be affiliated with distinct categories was different. When 'categorical' levels were matched there were no differences in response criterion. Linking this with present findings, it is in principle possible that TBR words shared a common category relative to TBF words, which resulted in a more liberal response criterion for TBR words. However, the designation of the directed forgetting instructions to words was counterbalanced. Thus, each word was followed by a remember instruction for half the participants whereas the other half of participants were instructed to forget the word. This makes it unlikely that the categorical similarities were different between TBR and TBF words and influenced response criterion.

A somewhat more likely interpretation – at least intuitively – is that TBR words received enhanced encoding relative to TBF words (Basden et al., 1993; MacLeod, 1975), resulting in an increased sense of familiarity and therefore leading to an increase in the likelihood of perceiving a stimulus as old rather than new. As a result, this may have led to a more lenient response criterion. This interpretation is compatible with the observed DF effects across all 5 experiments.

### 5.2.3 Recollection and Familiarity

The influence of emotion on the processes of recollection and familiarity differed between Experiments 2 (ROCs) and 3 (R/K procedure). In Experiment 2, recollection was superior for neutral relative to negative words, whereas in Experiment 3 there were no differences between emotional and neutral words for recollection. The differences between neutral and negative words in Experiment 2 are not in line with findings that remembering or recollection is greater for negative material (Ochsner, 2000; Sharot et al., 2007). As a result of extensive encoding (Hamann, 2001) and negative material being accompanied with remembering greater details (Bowen et al., 2018), arguably memory retrieval based on recollection should be increased for negative material. In Experiment 3, however, recollection was superior for negative relative to positive words. These outcomes fit with the view that general details of negative material tend to be better remembered relative to those of positive material (Kensinger & Kark, 2018). Alternatively, or perhaps as well, it may be that negative material receives prioritised processing, perhaps because of its salience, and hence more extensive encoding compared to positive material (Kensinger & Corkin, 2004). The same reasoning explaining the differences in memory sensitivity can be applied here; enhanced attention and encoding process are allocated towards negative material relative to positive (Johansson et al., 2004; Windmann & Kutas, 2001), which explains increased recollection for negative relative to positive material. In addition, in Experiment 3 valence interacted with recollection; recollection was superior for negative TBR words relative to positive TBR words and no differences were observed for TBF words. This suggests that whatever different processing negative material is subjected to, it results in negative material being easier to remember but not differentially resistant to forgetting.

With respect to familiarity, consistent across Experiments 2 and 3 was superior familiarity for negative relative to positive words. However, in contrast to Experiment 2, familiarity was also superior for emotional words relative to neutral words in Experiment 3. This is not surprising if emotional material has a higher degree of associations with personal experiences, which may lead to increased associated familiarity, as proposed by Kensinger and Corkin (2003). These findings are consistent with the findings for response criterion; a more liberal response criterion for emotional relative to neutral words and for negative relative to positive words. Presumably, familiarity influences response criterion and the more familiar an item seems the more liberal the response criterion that is adopted (Yonelinas, 1994). This has also been suggested by Windmann and Kutas (2001), who argued that the more liberal criterion for emotional material was linked with increased familiarity-based memory judgements for emotional relative to neutral material. In contrast, the absence of variability between emotional and neutral material in Experiment 2 is consistent with others observing little evidence of variability in familiarity due to emotion (Sharot et al., 2007; Yonelinas & Ritchey, 2015). The differences in familiarity according to negative and positive material in Experiment 2 and 3 also indicate that the contributions of recollection and familiarity differ between negative and positive material. Somewhat similar findings have been reported by other studies between negative and positive material (Ochsner, 2000), although the comparison of valence was not restricted to only negative and positive material in these cases.

The R/K procedure was used in Experiment 3 to examine the generality of findings in Experiment 2, where confidence judgements were employed. The relationship between recollection and familiarity in the two methods is based on somewhat distinct but also somewhat overlapping assumptions. In both methods (at least in the way they are used here) it is assumed that recollection and familiarity are independent (Yonelinas & Jacoby, 1995): the likelihood of one process being engaged is not influenced by the engagement of the other process. In the way that recollection and familiarity are computed here, it is also assumed that judgements are based on recollection when that occurs, and familiarity when it is above threshold and recollection fails. There is no scope in these models for recollection and familiarity signals to sum in any way to support a judgement, as has been suggested by Wixted and colleagues (Wixted, 2007; Wixted & Mickes, 2014; Wixted & Stretch, 2004; Yonelinas & Jacoby, 1995). In the R/K procedure, Remember judgements are assumed to reflect recollection. However, based on the independence assumption, Know judgements are an underestimate of familiarity because some Remember responses may actually have some degree of familiarity that would have elicited a Know response if recollection had failed (Yonelinas & Jacoby, 1995). For this reason, a correction was applied to the data (the independence R/K calculation) as described earlier in section 2.3.

In the ROC method, the highest level of confidence judgements is assumed to reflect recollection, with the remainder of the confidence judgements reflecting familiarity (Yonelinas & Parks, 2007). However, it has been argued that recollection may also be reflected by medium confidence judgements (Wixted, 2007). If this is the case, then discrepancies in the findings between Experiment 2 and 3, at least for recollection, are not surprising. This could also be the case for familiarity if high confidence judgements could also reflect familiarity-based judgements. Hence, one way of explaining the discrepancies between the two experiments is because both methods have somewhat different assumptions. This is also illustrated by the direct comparison using both methods to estimate recollection and familiarity in Experiments 2 and 3. This was done by analysing the ROC data in Experiment 2 using the R/K method and analysing the data from Experiment 3 using the ROC method. The rationale for doing this was to examine whether different methods would result in similar patterns of data. However, this was not the case, suggesting that at least one of these assumptions does not hold, although a caveat is that there were different participants in each experiment.

According to Parks and Yonelinas (2007), however, the assumptions underlying ROCs and the R/K procedure are accurate. They locate the reasons for any discrepancies across outcomes with the instructions provided to participants in the R/K procedure (Parks & Yonelinas, 2007). They argued that instructing participants to only respond Remember when they were able to provide contextual information when asked, results in responses actually based on recollection (e.g. Yonelinas et al., 1996). In Experiment 3, however, participants were simply instructed to respond Remember whenever they could remember any contextual details without the possibility of being asked for this information, which was also the case in other studies that argue against the similarities of the assumptions underlining these methods (Rotello et al., 2005; Wixted, 2007). This is another explanation for the discrepancies observed between Experiment 2 and 3. A comparison of Experiment 2 and 3 and previous literature illustrates that using different methods that are based on somewhat distinct assumptions make it challenging to compare findings across studies. Perhaps of most importance here, the present findings reinforce that negative and positive valence have distinct effects on memory, and therefore need to be considered separately.

### 5.2.4 ERPs, Encoding and Retrieval Processes

#### 5.2.4.1 Encoding Processes

There were no reliable subsequent memory effects, which is arguably not too surprising considering that previously reliable effects have been observed for the most part when using memory tasks that rely heavily on recollection, and where the separation between ‘remembered’ and ‘forgotten’ items is not just based on the accuracy of ‘old’ judgements (Friedman, de Chastelaine, Nessler, & Malcolm, 2010; Otten et al., 2001; Otten & Rugg, 2001; Otten et al., 2010).

To examine the encoding processes linked to instructions to remember and forget, and to determine whether these were modulated by emotion, another set of analyses was conducted. These were conducted on the ERPs elicited by the remember and forget cues that followed each item during the study phase. In the replication of study 1, there was a greater relative positivity for TBR cues relative to TBF cues over posterior scalp. This posterior positivity for TBR cues has also been reported in previous DF studies (Bailey & Chapman, 2012; Gallant et al., 2018; Hauswald et al., 2011). Furthermore, in the replication of study 1 and in study 2, there was a greater relative positivity over posterior scalp for TBR cues that was preceded by an emotional word relative to a neutral word. This effect had an onset of around 300ms and extended for the remainder of the recording epoch, at least in the replication of study 1. This suggests that emotion influences the encoding processes linked with TBR cues and not with TBF cues, which is consistent with the view that following TBF cues little effortful processing occurs. Previously, this greater positivity has been linked to enhanced attention allocation to emotional material (Hauswald et al., 2011; Kok, 1997), and is also consistent with observations in other studies (Brandt et al., 2013; Gallant & Dyson, 2016; Liu et al., 2017). However, in study 2, there was also a greater relative positivity over anterior scalp for TBF cues that were preceded by a neutral word relative to a positive word. The positivity over anterior scalp has been linked to inhibitory control processes (Yang et al., 2012), and has been argued to indicate that inhibitory processes are partially responsible for the DF effect (Anderson & Hanslmayr, 2014).

While these findings indicate that TBR and TBF cues elicit different activity and vary according to emotion, it is not clear on the basis of a simple pairwise comparison if this reflects simply more of the same kind of processing in one case rather than the other, qualitatively different processing in the two cases, or in fact active encoding processing in only one case. Thus, the finding of a difference is an important first step, but it does not permit strong claims to be made, for example, about the merits of a selective rehearsal or an active inhibition account (Anderson & Hanslmayr, 2014; Basden et al., 1993; MacLeod, 1975). Moreover, there is some indication based on these findings that negative and positive material relative to neutral material influence encoding processes linked with DF instructions differently. Whereas negative and positive materiel seemed to influence TBR encoding processes similarly, this was not the case for TBF cues.

#### 5.2.4.2 Retrieval Processes – Recollection and Familiarity

There were no reliable mid-frontal or left-parietal ERP old/new effects in the three studies. In study 1 the ERP old/new effects, while not reliable, also showed some polarity reversals that differ from what is typically observed (Friedman & Johnson, 2000; Rugg & Curran, 2007; Vilberg & Rugg, 2008). This motivated the replication study, where, while remaining non-significant, the old/new effects were qualitatively similar to effects observed previously (Rugg & Curran, 2007; Sanquist et al., 1980; Yonelinas, 2002). This was also the case for ERP study 2. In light of these outcomes the ERP data from test cannot provide any insights into the ways in which recollection and familiarity vary according to directed forgetting manipulations and emotion. Possible reasons for these null outcomes are discussed in Chapter 3.2.3.

## 5.3 Future Directions

The findings reported here on the links between emotion, memory and memory control provide some new insights, and replicate some previous findings but not others. In the considerations for future directions articulated below, three possibilities are identified. Not because they are the only issues that merit further investigation, rather that they are matters that could be taken forward relatively straightforwardly in a way that builds on the findings that have been described. These issues are: the role(s) played by semantic relatedness, encoding processes engaged by remember and forget cues, and changes in response criterion with emotion.

Overall, the outcomes reported here highlight that semantic relatedness among materials is an important factor that requires attention if the intent is to attribute confidently changes in memory and memory control to emotion as opposed to factors that variables correlate with changes in emotion. The limited studies that have controlled for semantic relatedness (Minnema & Knowlton, 2008) have still reported differences in DF effects according to emotion, which is not consistent with the outcomes reported in all of the relevant experiments in this thesis. In these experiments, however, care was taken to control for semantic relatedness throughout. A sensible extension of the work described here would be to manipulate relatedness within the same experiment designs, in an approach similar to that employed by Dougal and Rotello (2007).

Another complementary approach would be to examine whether the degree of semantic relatedness was indeed different between emotional and neutral words in prior studies where this was not reported: the degree of semantic relatedness may or may not have been different between emotional and neutral materials. This observation does not only apply to semantic relatedness. Other word properties, notably word frequency, have not been controlled for consistently. In the case of studies that have used images, the same reasoning applies. Most studies have not included any control of visual complexity between emotional and neutral images (Nowicka et al., 2011; Otani et al., 2012; Payne & Corrigan, 2007), as was done in Experiment 5. Again, whether there were actual differences in visual complexity is hard to say when these measures have not been reported.

A meta-analysis of previous studies has the potential to shed some light on whether stimulus properties differed between emotional and neutral material and how these might relate to the links between emotion and memory control. This might lead to a clearer picture of the reasons for the variability reported in the existing literature.

The ERP data in this thesis that provide new insights are those effects elicited by the remember and forget cues. These data highlighted the sensitivity of ERPs to encoding processes, offered indications that some encoding processes vary with emotion, and suggested that TBF instructions do not simply result in the absence of encoding operations. The key findings that suggested these outcomes were in separate experiments, and in each case two valence categories were used: negative vs neutral and positive vs neutral. Because the key outcomes were not the same in each of these experiments, they suggest that the encoding operations afforded positive and negative materials are not the same. These data points therefore argue strongly for a direct investigation of this possibility by including positive and negative material in the same experiment. As elsewhere in this thesis, when negative and positive materials are contrasted together, differences in memory, response criterion and recollection and familiarity arise. Analyzing ERPs elicited by remember and forget cues offers a means of identifying the precursors of these outcomes that are seen in the behavioral outcomes in the test phases of experiments.

In Experiment 5, response criterion did not change across the emotional and neutral conditions. In the majority of the preceding experiments there was a consistently more liberal criterion for emotional contents. The most likely reason for the absence of a comparable change in Experiment 5 is the fact that neutral words were employed throughout, with the emotion manipulation occurring via the use of background images at study that varied with emotion. The outcomes therefore suggest that processes engaged during retrieval rather than during encoding are what result in differences in response criterion. The design of Experiment 5, however, is markedly different from that of the preceding experiments, so in principle other unidentified elements of the design might be responsible for the null outcome for criterion. There are two designs that merit consideration. First, the use of a valence manipulation at test only – perhaps via the use of incidental background images that differ in valence. Second, an extension of Experiment 5 where words from more than one valence category are employed. If factors at test are in fact responsible for criterion changes, then there should be a liberal criterion for emotional words, but no differences in criterion according to study background.

## 5.4 Concluding Remarks

The work in this thesis highlights the importance of careful experimental design in order to compare findings between studies and further the understanding of the links between emotion, memory and memory control. The approach taken here was to exert tight control over stimulus properties and to maintain that consistently across a set of studies designed to address an overlapping set of questions. Among many findings, perhaps the key ones to highlight are: (i) when semantic relatedness is controlled for, there is little evidence that directed forgetting varies with emotion; (ii) via demonstrations using ROCs, emotion influences criterion and not memory sensitivity consistently – a conclusion that could not have been drawn on the basis of many prior studies in which only single point measures were employed; (iii) emotion influences the encoding operations that are engaged by remember and via forget cues; (iv) when emotion is manipulated only at study, there were no changes in response criterion. At the very least, the findings reported here can be seen as a set of pointers to guide the direction of the subsequent research into the links between emotion, memory and memory control.

**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

Aggleton, J. P. (2012). Multiple anatomical systems embedded within the primate medial temporal lobe: Implications for hippocampal function. *Neuroscience and Biobehavioral Reviews*, *36*(7), 1579–1596. https://doi.org/10.1016/j.neubiorev.2011.09.005

Aggleton, J. P. (2014). Looking beyond the hippocampus: Old and new neurological targets for understanding memory disorders. *Proceedings of the Royal Society B: Biological Sciences*, *281*, 1–9. https://doi.org/10.1098/rspb.2014.0565

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.

Anderson, M. C., Bunce, J. G., & Barbas, H. (2016). Prefrontal hippocampal pathways underlying inhibitory control over memory. *Neurobiology of Learning and Memory*, *134*, 145–161. https://doi.org/10.1016/j.nlm.2015.11.008

Anderson, M. C., & Green, C. (2001). Suppressing unwanted memories by executive control. *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), 279–282. https://doi.org/10.1016/j.tics.2014.03.002

Anderson, M. C., & Spellman, B. A. (1995). On the status of inhibitory mechanisms in cognition: memory retrieval as a model case. *Psychological Review*, *102*(1), 68–100. https://doi.org/10.1037/0033-295X.102.1.68

Azimian-Faridani, N., & Wilding, E. L. (2004). An event-related potential study of the revelation effect. *Psychonomic Bulletin and Review*, *11*(5), 926–931.

https://doi.org/10.3758/BF03196723

Baddeley, A. D., Thomson, N., & Buchanan, M. (1975). Word length and the structure of short-term memory. *Journal of Verbal Learning and Verbal Behavior*, *14*(6), 575–589. https://doi.org/10.1016/S0022-5371(75)80045-4

Baddeley, A., Lewis, V., Eldridge, M., & Thomson, N. (1984). Attention and retrieval from long-term memory. *Journal of Experimental Psychology: General*, *113*(4), 518–540. https://doi.org/10.1037/0096-3445.113.4.518

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

Bancroft, T. D., Hockley, W. E., & Farquhar, R. (2013). The longer we have to forget the more we remember: The ironic effect of postcue duration in item-based directed forgetting. *Journal of Experimental Psychology: Learning Memory and Cognition*, *39*(3), 691–699. https://doi.org/10.1037/a0029523

Barber, S. J., & Mather, M. (2012). Forgetting in context: The effects of age, emotion, and social factors on retrieval-induced forgetting. *Memory and Cognition*, *40*(6), 874–888. https://doi.org/10.3758/s13421-012-0202-8

Barnier, A., Hung, L., & Conway, M. (2004). Retrieval induced forgetting of emotional and unemotional autobiographical memories. *Cognition & Emotion*, *18*(4), 457–477. https://doi.org/10.1080/0269993034000392

Basden, B. H., Basden, D. R., & Gargano, G. J. (1993). Directed forgetting in implicit and explicit memory tests: A comparison of methods. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*(3), 603–616. https://doi.org/10.1037/0278-7393.19.3.603

Berger, N., Crossman, M., & Brandt, K. R. (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

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

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, DC: Winston.

Blaney, P. H. (1986). Affect and memory: a review. *Psychological Bulletin*, *99*(2), 229–246. https://doi.org/10.1037/0033-2909.99.2.229

Blix, I., & Brennen, T. (2012). Retrieval-induced forgetting after trauma: A study with victims of sexual assault. *Cognition and Emotion*, *26*(2), 321–331. https://doi.org/10.1080/02699931.2011.570312

Bowen, H. J., Kark, S. M., & Kensinger, E. A. (2018). NEVER forget: negative emotional valence enhances recapitulation. *Psychonomic Bulletin and Review*, *25*(3), 870–891. https://doi.org/10.3758/s13423-017-1313-9

Bowen, H. J., Spaniol, J., Patel, R., & Voss, A. (2016). A diffusion model analysis of decision biases affecting delayed recognition of emotional stimuli. *PLoS ONE*, *11*(1), 1–20. https://doi.org/10.1371/journal.pone.0146769

Bower, G. H., Monteiro, K. P., & Gilligan, S. G. (1978). Emotional mood as a context for learning and recall. *Journal of Verbal Learning and Verbal Behavior*, *17*(5), 573–585. https://doi.org/10.1016/S0022-5371(78)90348-1

Bower, G., & Mayer, J. (1989). In search of mood-dependent retrieval. *Journal of Social Behavior and Personality*, *2*, 1–38. https://doi.org/10.1080/02699939108411044

Bradley, M. M., Greenwald, M. K., Petry, M. C., & Lang, P. J. (1992). Remembering pictures: pleasure and arousal in memory. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *18*(2), 379–390. https://doi.org/10.1037//0278-7393.18.2.379

Brandt, K. R., Nielsen, M. K., & 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

Brewin, C. (2018). Memory and forgetting. *Current Psychiatry Reports*, *20*(87), 2–8. https://doi.org/https://doi.org/10.1007/s11920-018-0950-7

Bridger, E. K., Bader, R., Kriukova, O., Unger, K., & Mecklinger, A. (2012). The FN400 is functionally distinct from the N400. *NeuroImage*, *63*(3), 1334–1342. https://doi.org/10.1016/j.neuroimage.2012.07.047

Buchanan, T. W. (2007). Retrieval of emotional memories. *Psychological Bulletin*, *133*(5), 761–779. https://doi.org/10.1037/0033-2909.133.5.761

Buchanan, T. W., Etzel, J. A., Adolphs, R., & Tranel, D. (2006). The influence of autonomic arousal and semantic relatedness on memory for emotional words. *International Journal of Psychophysiology*, *61*(1), 26–33. https://doi.org/10.1016/j.ijpsycho.2005.10.022

Budson, A. E., Todman, R. W., & Schacter, D. L. (2014). Gist memory in Alzheimer's disease: Evidence from categorized pictures. *Neuropsychology*, *20*(1), 113–122. https://doi.org/10.1037/0894-4105.20.1.113

Bulevich, J. B., Roediger, H. L., Balota, D. A., & Butler, A. C. (2006). Failures to find suppression of episodic memories in the think/no-think paradigm. *Memory and Cognition*, *34*(8), 1569–1577. https://doi.org/10.3758/BF03195920

Burgess, N., Hockley, W. E., & Hourihan, K. L. (2017). The effects of context in item-based directed forgetting: Evidence for "one-shot" context storage. *Memory and Cognition*, *45*(5), 745–754. https://doi.org/10.3758/s13421-017-0692-5

Burgess, P. W., & Shallice, T. (1996). Confabulation and the control of recollection. *Memory*, *4*(4), 359–412. https://doi.org/http://dx.doi.org/10.1080/096582196388906

Cahill, L., & McGaugh, J. L. (1995). A novel demonstration of enhanced memory associated with emotional arousal. *Consciousness and Cognition*, *4*, 410–421. https://doi.org/10.1006/ccog.1995.1048

Carretié, L., Tapia, M., López-Martín, S., & Albert, J. (2019). EmoMadrid: An emotional pictures database for affect research. *Motivation and Emotion*, *43*, 929–939. https://doi.org/10.1007/s11031-019-09780-y

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

Chiu, C., Yeh, Y. Y., Huang, C. L., Wu, Y. C., Chiu, Y. C., & Lin, C. C. (2010). Unintentional memory inhibition is weakened in non-clinical dissociators. *Journal of Behavior Therapy and Experimental Psychiatry*, *41*(2), 117–124. https://doi.org/10.1016/j.jbtep.2009.11.003

Clayton, N. S., & Dickinson, A. (1998). Episodic-like memory during cache recovery by scrub jays. *Nature*, *395*, 272–274. https://doi.org/10.1038/26216

Coles, M. G. H., & Rugg, M. D. (1995). Event-related brain potentials: an introduction. In M. D. Rugg & M. G. H. Coles (Eds.), *Electrophysiology of mind. Event-related brain potentials and cognition.* (pp. 1-26). New York: Oxford University Press.

Cortese, M. J., Khanna, M. M., & Hacker, S. (2010). Recognition memory for 2,578 monosyllabic words. *Memory*, *18*(6), 595–609. https://doi.org/10.1080/09658211.2010.493892

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, *11*(6), 671–684. https://doi.org/10.1016/S0022-5371(72)80001-X

Curran, T. (1999). The electrophysiology of incidental and intentional retrieval: ERP old/new. *Neuropsychologia*, *37*, 771–785. https://doi.org/10.1016/S0028-3932(98)00133-X

Curran, T. (2000). Brain potentials of recollection and familiarity. *Memory & Cognition*, *28*(6), 923–938. https://doi.org/10.3758/BF03209340

Curran, T. (2004). Effects of attention and confidence on the hypothesized ERP correlates of recollection and familiarity. *Neuropsychologia*, *42*, 1088–1106. https://doi.org/10.1016/j.neuropsychologia.2003.12.011

Curran, T., & Cleary, A. M. (2003). Using ERPs to dissociate recollection from familiarity in picture recognition. *Cognitive Brain Research*, *15*(2), 191–205. https://doi.org/10.1016/S0926-6410(02)00192-1

Cuthbert, B. N., Schupp, H. T., Bradley, M. M., Birbaumer, N., & Lang, P. J. (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. (2008). Does retrieval-induced forgetting occur for emotional stimuli? *Cognition & 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. *Psychological Science*, *17*(5), 441–447. https://doi.org/10.1111/j.1467-9280.2006.01725.x

Dewhurst, S. A., & Parry, L. A. (2000). Emotionality, distinctiveness, and recollective experience. *European Journal of Cognitive Psychology*, *12*(4), 541–551. https://doi.org/10.1080/095414400750050222

Doidge, A. N., Evans, L. H., Herron, J. E., & Wilding, E. L. (2017). Separating content-specific retrieval from post-retrieval processing. *Cortex*, *86*, 1–10. https://doi.org/10.1016/j.cortex.2016.10.003

Dolan, R. J., Lane, R. D., & Chua, P. M. (1999). Common effects of emotional valence, arousal and attention on neural activation during visual processing of pictures. *Neuropsychologia*, *37*(9), 989–997. https://doi.org/10.1016/S0028-3932(99)00017-2

Dolcos, F., & Cabeza, R. (2002). Event-related potentials of emotional memory: Encoding pleasant, unpleasant and nuetral pictures. *Cognitive, Affective, & Behavioral Neuroscience*, *2*(3), 252–263. https://doi.org/10.3758/CABN.2.3.252

Dolcos, F., & Denkova, E. (2008). Neural correlates of encoding emotional memories : a review of functional neuroimaging evidence. *Cell Science Reviews*, *5*(2), 78–122. https://doi.org/10.2117/psysoc.2012.80

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

Donchin, E. (1966). A multivariate approach to the analysis of average evoked potentials. *IEEE Transactions on Biomedical Engineering*, *13*(3), 131–139. https://doi.org/10.1109/TBME.1966.4502423

Dougal, S., & Rotello, C. M. (2007). “Remembering” emotional words is based on response bias, not recollection. *Psychonomic Bulletin & Review*, *14*(3), 423–429. https://doi.org/10.3758/BF03194083

Driver, J., & Frith, C. (2000). Shifting baselines in attention research. *Nature Neuroscience*, *1*, 147–148. https://doi.org/10.1038/35039083

Duarte, A., Ranganath, C., Winward, L., Hayward, D., & Knight, R. T. (2004). Dissociable neural correlates for familiarity and recollection during the encoding and retrieval of pictures. *Cognitive Brain Research*, *18*(3), 255–272. https://doi.org/10.1016/j.cogbrainres.2003.10.010

Dunn, J. C. (2001). Dual-Process models of the Remember/Know paradigm. *Psychonomic Bulletin & Review*, 1–35.

Dunn, J. C. (2004). Remember–Know: A matter of confidence. *Psychological Review*, *111*(2), 524–542. https://doi.org/10.1037/0033-295X.111.2.524

Düzel, E., Vargha-Khadem, F., Heinze, H. J., & Mishkin, M. (2001). Brain activity evidence for recognition without recollection after early hippocampal damage. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(14), 8101–8106. https://doi.org/10.1073/pnas.131205798

Düzel, Emrah, Yonelinas, A. P., Mangun, G. R., Heinze, H.-J., & Tulving, E. (1997). Event-related brain potential correlates of two states of conscious awareness in memory. *Proceedings of the National Academy of Sciences*, *94*, 5973–5978.

Dzulkifli, M. A., Sharpe, H. L., & Wilding, E. L. (2004). Separating item-related electrophysiological indices of retrieval effort and retrieval orientation. *Brain and Cognition*, *55*(3), 433–443. https://doi.org/10.1016/j.bandc.2004.03.004

Easton A, Eacott MJ. 2008. A new working definition of episodic memory: replacing 'when' with 'which'. In Handbook of episodic memory (ed. E. Dere et al.), Vol. 18, pp. 185–196. Elsevier, Amsterdam.

Eftekhari, A., Zoellner, L. A., & Vigil, S. A. (2009). Patterns of emotion regulation and psychopathology. *Anxiety, Stress, and Coping*, *22*(5), 571–586. https://doi.org/10.1080/10615800802179860

Egan, J. P. (1958). Recognition memory and the operating characteristic (Tech. Note AFCRC-TN-58–51). Bloomington: Indiana University, Hearing and Communication

Laboratory.

Egan, J. P. (1975). Signal detection theory and ROC analysis. New York: Academic Press

Eich, E. (1995). Searching for mood dependent memory. *Psychological Science*, *6*(2), 67–75. https://doi.org/10.1111/j.1467-9280.1995.tb00309.x

Eichenbaum, H., Otto, T., & Cohen, N. J. (1994). Two functional components of the hippocampal memory system. *Behavioral and Brain Sciences*, *17*(3), 449–472. https://doi.org/10.1017/S0140525X00035391

Eichenbaum, H., Yonelinas, A. P., & Ranganath, C. (2007). The medial temporal lobe and recognition memory. *Annual Review of Neuroscience*, *30*, 123–152. https://doi.org/10.1146/annurev.neuro.30.051606.094328

Eimer, M. (1993). Effects of attention and stimulus probability on ERPs in a Go/Nogo task. *Biological Psychology*, *35*, 123–138. https://doi.org/10.1016/0301-0511(93)90009-W

Elzinga, B. M., & Bremner, J. D. (2002). Are the neural substrates of memory the final common pathway in posttraumatic stress disorder (PTSD)? *Journal of Affective Disorders*, *70*, 1–17. https://doi.org/10.1016/S0165-0327(01)00351-2

Engen, H. G., & Anderson, M. C. (2018). Memory control: A fundamental mechanism of emotion regulation. *Trends in Cognitive Sciences*, *22*(11), 982–995. https://doi.org/10.1016/j.tics.2018.07.015

Eustache, F., Desgranges, B., Laville, P., Guillery, B., Lalevée, C., Schaeffer, S., de la Sayette, V., Iglesias, S., Baron, J. C., & Viader, F. (1999). Episodic memory in transient global amnesia: encoding, storage, or retrieval deficit? *Journal of Neurology, Neurosurgery, and Psychiatry*, *66*(2), 148–154. https://doi.org/10.1136/jnnp.66.2.148

Fawcett, J. M., & Taylor, T. L. (2008). Forgetting is effortful: Evidence from reaction time probes in an item-method directed forgetting task. *Memory & Cognition*, *36*(6), 1168–1181. https://doi.org/10.3758/MC.36.6.1168

Friedman, D., de Chastelaine, M., Nessler, D., & Malcolm, B. (2010). Changes in familiarity and recollection across the lifespan: An ERP perspective. *Brain Research*, *1310*, 124–141. https://doi.org/10.1016/j.brainres.2009.11.016

Friedman, D., & Johnson, R. J. (2000). Event-related potential (ERP) studies of memory encoding and retrieval: A selective review. *Microscopy Research and Technique*, *51*, 6–28. https://doi.org/10.1002/1097-0029(20001001)51:1<6::AID-JEMT2>3.0.CO;2-R

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

Gallant, S. N., Dyson, B. J., & Yang, L. (2017). Local context effects during emotional item directed forgetting in younger and older adults. *Memory*, *25*(8), 1129–1138. https://doi.org/10.1080/09658211.2016.1274036

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

Gallant, S. N., & Yang, L. (2014). Positivity effect in source attributions of arousal-matched emotional and non-emotional words during item-based directed forgetting. *Frontiers in Psychology*, *5*, 1–10. https://doi.org/10.3389/fpsyg.2014.01334

Gardiner, J. M. (1988). Functional aspects of recollective experience. *Memory & Cognition*, *16*(4), 309–313. https://doi.org/10.3758/BF03197041

Gardiner, J. M. (2001). Episodic memory and autonoetic consciousness: a first-person approach. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *356*, 1351–1361. https://doi.org/10.1098/rstb.2001.0955

Garnefski, N., & Kraaij, V. (2007). The cognitive emotion regulation questionnaire: Psychometric features and prospective relationships with depression and anxiety in adults. *European Journal of Psychological Assessment*, *23*(3), 141–149. https://doi.org/10.1027/1015-5759.23.3.141

Gillund, G., & Shiffrin, R. M. (1984). A retrieval model for both recognition and recall. *Psychological Review*, *91*(1), 1–67. https://doi.org/10.1037/0033-295X.91.1.1

Glanzer, M., & Bowles, N. (1976). Analysis of the word-frequency effect in recognition memory. *Journal of Experimental Psychology: Human Learning and Memory*, *2*(1), 21–31. https://doi.org/10.1037/0278-7393.2.1.21

Glanzer, M., Hilford, A., Kim, K., & Adams, J. K. (1999). Slope of the receiver-operating characteristic in recognition memory. *Journal of Experimental Psychology: Learning Memory and Cognition*, *25*(2), 500–513. https://doi.org/10.1037/0278-7393.25.2.500

Greenhouse, S. W., & Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika*, *24*(2), 95–112. https://doi.org/10.1007/BF02289823

Groh-Bordin, C., Zimmer, H. D., & Mecklinger, A. (2005). Feature binding in perceptual priming and in episodic object recognition: Evidence from event-related brain potentials. *Cognitive Brain Research*, *24*(3), 556–567. https://doi.org/10.1016/j.cogbrainres.2005.03.006

Gross, J. J., & John, O. P. (2003). Emotion Regulation Questionnaire. *NeuroImage*, *48*(10), 9–9. https://doi.org/10.1037/0022-3514.85.2.348

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

Hanslmayr, S., Klimesch, W., Sauseng, P., Gruber, W., Doppelmayr, M., Freunberger, R., Pecherstorfer, T., & Birbaumer, N. (2007). Alpha phase reset contributes to the generation of ERPs. *Cerebral Cortex*, *17*, 1–8. https://doi.org/10.1093/cercor/bhj129

Harlow, I. M., & Donaldson, D. I. (2013). Source accuracy data reveal the thresholded nature of human episodic memory. *Psychonomic Bulletin and Review*, *20*(2), 318–325.

https://doi.org/10.3758/s13423-012-0340-9

Harris, C. B., Sharman, S. J., Barnier, A. J., & Moulds, M. L. (2010). Mood and retrieval-induced forgetting of positive and negative autobiographical memories. *Applied Cognitive Psychology*, *24*, 399–413. https://doi.org/10.1002/acp

Harvey, A. G., Bryant, R. A., & Dang, S. T. (1998). Autobiographical memory in acute stress disorder. *Journal of Consulting and Clinical Psychology*, *66*(3), 500–506. https://doi.org/10.1037//0022-006x.66.3.500

Hauer, B. J. A., & Wessel, I. P. (2006). Retrieval-induced forgetting of autobiographical memory details. *Cognition and Emotion*, *20*(3–4), 430–447. https://doi.org/10.1080/02699930500342464

Hauswald, A., Schulz, H., Iordanov, T., & Kissler, J. (2011). ERP dynamics underlying successful directed forgetting of neutral but not negative pictures. *Social Cognitive and Affective Neuroscience*, *6*(4), 450–459. https://doi.org/10.1093/scan/nsq061

Heathcote, A., Raymond, F., & Dunn, J. (2006). Recollection and familiarity in recognition memory: Evidence from ROC curves. *Journal of Memory and Language*, *55*, 495–514. https://doi.org/10.1016/j.jml.2006.07.001

Hellerstedt, R., & Johansson, M. (2014). Electrophysiological correlates of competitor activation predict retrieval-induced forgetting. *Cerebral Cortex*, *24*(6), 1619–1629. https://doi.org/10.1093/cercor/bht019

Hendry, L., & Tehan, G. (2005). An item/order trade-off explanation of word length and generation effects. *Memory*, *13*(3–4), 364–371. https://doi.org/10.1080/09658210344000341

Henson, R. (2005). A mini-review of fMRI studies of human medial temporal lobe activity associated with recognition memory. *The Quarterly Journal of Experimental Psychology*, *58*(3–4), 340–360. https://doi.org/10.1080/02724990444000113

Henson, R. N. A., Rugg, M. D., Shallice, T., Josephs, O., & Dolan, R. J. (1999). Recollection and familiarity in recognition memory: An event-related functional magnetic resonance imaging study. *Journal of Neuroscience*, *19*(10), 3962–3972. https://doi.org/10.1523/jneurosci.19-10-03962.1999

Herbert, C., Junghofer, M., & Kissler, J. (2008). Event related potentials to emotional adjectives during reading. *Psychophysiology*, *45*(3), 487–498. https://doi.org/10.1111/j.1469-8986.2007.00638.x

Herron, J. E., & Wilding, E. L. (2004). An electrophysiological dissociation of retrieval mode and retrieval orientation. *NeuroImage*, *22*(4), 1554–1562. https://doi.org/10.1016/j.neuroimage.2004.04.011

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

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

Higham, P. A. (2002). Strong cues are not necessarily weak: Thomson and Tulving (1970) and the encoding specificity principle revisited. *Memory & Cognition*, *30*(1), 67–80. https://doi.org/10.3758/BF03195266

Hintzman, D. L. (1986). “Schema Abstraction” in a multiple-trace memory model. *Psychological Review*, *93*(4), 411–428.

Ho, S. M. Y., Cheng, J., & Dai, D. W. T. (2017). The effect of positive and negative memory bias on anxiety and depression symptoms among adolescents. *Journal Clinical Psychology*, *74*, 1509–1525. https://doi.org/10.1002/jclp.22597

Hockley, W. E., Ahmad, F. N., & Nicholson, R. (2016). Intentional and incidental encoding of item and associative information in the directed forgetting procedure. *Memory and*

*Cognition*, *44*(2), 220–228. https://doi.org/10.3758/s13421-015-0557-8

Hsieh, L. T., Hung, D. L., Tzeng, O. J. L., Lee, J. R., & Cheng, S. kuen. (2009). An event-related potential investigation of the processing of Remember/Forget cues and item encoding in item-method directed forgetting. *Brain Research*, *1250*, 190–201. https://doi.org/10.1016/j.brainres.2008.11.016

Jackson, A. F., & Bolger, D. J. (2014). The neurophysiological bases of EEG and EEG measurement: A review for the rest of us. *Psychophysiology*, *51*(11), 1061–1071. https://doi.org/10.1111/psyp.12283

Jacoby, L. L., & Dallas, M. (1981). On the relationship between autobiographical memory and perceptual learning. *Journal of Experimental Psychology: General*, *110*(3), 306–340. https://doi.org/10.1037/0096-3445.110.3.306

Jasper, H. H., (1958). The ten-twenty electrode system of the International Federation. *Electroencephalography and Clinical Neurophysiology, 10*, 371-375.

Johansson, M., Mecklinger, A., & Treese, A. (2004). Recognition memory for emotional and neutral faces: An event-related potential study. *Journal of Cognitive Neuroscience*, *16*(10), 1840–1853. https://doi.org/10.1162/0898929042947883

Joormann, J., & Gotlib, I. H. (2010). Emotion regulation in depression: relation to cognitive inhibition. *Cognition & Emotion*, *24*(2), 281–298. https://doi.org/10.1080/02699930903407948

Kapucu, A., Rotello, C. M., Ready, R. E., & Seidl, K. N. (2008). Response bias in “remembering” emotional stimuli: A new perspective on age differences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(3), 703–711. https://doi.org/http://dx.doi.org/10.1037/0278-7393.34.3.703

Katkov, M., Romani, S., & Tsodyks, M. (2014). Word length effect in free recall of randomly assembled word lists. *Frontiers in Computational Neuroscience*, *8*, 8–11.

https://doi.org/10.3389/fncom.2014.00129

Keele, S. W. (1972). Attention demands of memory retrieval. *Journal of Experimental Psychology*, *93*(2), 245–248. https://doi.org/10.1037/h0032460

Kensinger, E. A., & Corkin, S. (2003). Memory enhancement for emotional words: Are emotional words more vividly remembered than neutral words? *Memory & Cognition*, *31*(8), 1169–1180. https://doi.org/10.3758/BF03195800

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 of the United States of America*, *101*(9), 3310–3315. https://doi.org/10.1073/pnas.0306408101

Kensinger, E. A., & Kark, S. M. (2018). Emotion and Memory. In J. T. Wixted (Ed.), *Steven's Handbook of Experimental Psychology and Cognitive Neuroscience.* (Fourth Edi, Issue 1, pp. 1–26). John Wiley & Sons, Inc. https://doi.org/10.1126/science.261.5119.369

Kensinger, E. A., & Schacter, D. L. (2006). Amygdala activity is associated with the succesful encoding of item, but not source, information for positive and negative stimuli. *Journal of Neuroscience*, *26*(9), 2564–2570. https://doi.org/10.1523/JNEUROSCI.5241-05.2006

Kinsbourne, M., & George, J. (1974). The mechanism of the word-frequency effect on recognition memory. *Journal of Verbal Learning and Verbal Behavior*, *13*(1), 63–69. https://doi.org/10.1016/S0022-5371(74)80031-9

Klem, G., Luders, H., Jasper, H., & Elger, C. (1958). The ten-twenty electrode system of the International Federation. *Electroencephalography and Clinical Neurophysiology*, *10*(2), 371–375. https://doi.org/10.1016/0013-4694(58)90053-1

Koen, J. D., Barrett, F. S., Harlow, I. M., & Yonelinas, A. P. (2017). The ROC Toolbox: A toolbox for analyzing receiver-operating characteristics derived from confidence ratings. *Behavior Research Methods*, *49*(4), 1399–1406. https://doi.org/10.3758/s13428-016-0796-z

Kok, A. (1986). Effects of degradation of visual stimuli on components of the event-related potential (ERP) in Go/NoGo reaction tasks. *Biological Psychology*, *23*, 21–38. https://doi.org/10.1016/0301-0511(86)90087-6

Kok, A. (1997). Event-related-potential (ERP) reflections of mental resources: A review and synthesis. *Biological Psychology*, *45*(1–3), 19–56. https://doi.org/10.1016/S0301-0511(96)05221-0

Koutstaal, W., & Schacter, D. L. (1997). Gist-based false recognition of pictures in older and younger adults. *Journal of Memory and Language*, *37*, 555–583. https://doi.org/10.1006/jmla.1997.2529

Kutas, M., & Dale, A. (1997). Electrical and magnetic readings of mental functions. In M. D. Rugg (Ed.), *Cognitive Neuroscience. Studies in Cognition* (pp. 197 – 242). Cambridge, MA: MIT Press.

Kuyken, W., & Dalgleish, T. (1995). Autobiographical memory and depression. *British Journal of Clinical Psychology*, *34*(1), 89–92. https://doi.org/10.1111/j.2044-8260.1995.tb01441.x

Lambert, A. J., Good, K. S., & Kirk, I. J. (2010). Testing the repression hypothesis: Effects of emotional valence on memory suppression in the think - No think task. *Consciousness and Cognition*, *19*(1), 281–293. https://doi.org/10.1016/j.concog.2009.09.004

Landauer, T. K., Foltz, P. W., & Laham, D. (1998). An introduction to latent semantic analysis. *Discourse Processes*, *25*(2), 259–284. https://doi.org/10.1080/01638539809545028

Lavoie, M. E., & O'Connor, K. P. (2013). Effect of emotional valence on episodic memory stages as indexed by event-related potentials. *World Journal of Neuroscience*, *03*(04), 250–262. https://doi.org/10.4236/wjns.2013.34034

Lecompte, D. C. (1995). Recollective experience in the revelation effect: Separating the contributions of recollection and familiarity. *Memory & Cognition*, *23*(3), 324–334. https://doi.org/10.3758/BF03197234

Levine, L. J., & Edelstein, R. S. (2009). Emotion and memory narrowing: A review and goal-relevance approach. *Cognition & Emotion*, *23*(5), 833–875. https://doi.org/10.1080/02699930902738863

Levy, B. J., & Anderson, M. C. (2002). Inhibitory processes and the control of memory retrieval. *Trends in Cognitive Sciences*, *6*(7), 299–305. https://doi.org/10.1016/S1364-6613(02)01923-X

Leynes, P. A., & Phillips, M. C. (2008). Event-related potential (ERP) evidence for varied recollection during source monitoring. *Journal of Experimental Psychology: Learning Memory and Cognition*, *34*(4), 741–751. https://doi.org/10.1037/0278-7393.34.4.741

Li, G., Wang, L., & Han, Y. (2017). Directed forgetting of negative performed actions is difficult: A behavioural study. *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

Loprinzi, P., & Frith, E. (2018). The role of sex in memory function: considerations and recommendations in the context of exercise. *Journal of Clinical Medicine*, *7*(6), 132. https://doi.org/10.3390/jcm7060132

Luck, S. J. (2014). *An introduction to the event-related potential technique*. MIT Press LTD.

https://doi.org/10.1118/1.4736938

MacLeod, C. M. (1975). Long-term recognition and recall following directed forgetting. *Journal of Experimental Psychology: Human Learning and Memory*, *104*(3), 271–279. https://doi.org/10.1037/0278-7393.1.3.271

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., & 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), 132–142. https://doi.org/10.1037/0278-7393.22.1.132

Macmillan, N. A., & Creelman, C. D. (1990). Response bias: Characteristics of detection theory, threshold theory, and “nonparametric” indexes. *Psychological Bulletin*, *107*(3), 401–413. https://doi.org/10.1037/0033-2909.107.3.401

Macmillan, N. A., & Creelman, C. D. (2005). Detection theory: A user’s guide. New York: Cambridge University Press

Madan, C. R., Bayer, J., Gamer, M., Lonsdorf, T. B., & Sommer, T. (2018). Visual complexity and affect: Ratings reflect more than meets the eye. *Frontiers in Psychology*, *8*, 1–19. https://doi.org/10.3389/fpsyg.2017.02368

Madan, C. R., Caplan, J. B., Lau, C. S. M., & Fujiwara, E. (2012). Emotional arousal does not enhance association-memory. *Journal of Memory and Language*, *66*(4), 695–716. https://doi.org/10.1016/j.jml.2012.04.001

Madan, C. R., Fujiwara, E., Caplan, J. B., & Sommer, T. (2017). Emotional arousal impairs association-memory: Roles of amygdala and hippocampus. *NeuroImage*, *156*, 14–28. https://doi.org/10.1016/j.neuroimage.2017.04.065

Madan, C. R., Scott, S. M. E., & Kensinger, E. A. (2019). Positive emotion enhances association-memory. *Emotion*, *19*(4), 733–740. https://doi.org/10.1037/emo0000465

Madan, C. R., Shafer, A. T., Chan, M., & Singhal, A. (2017). Shock and awe: Distinct effects of taboo words on lexical decision and free recall. *Quarterly Journal of Experimental Psychology*, *70*(4), 793–810. https://doi.org/10.1080/17470218.2016.1167925

Mandera, P., Keuleers, E., & Brysbaert, M. (2017). Explaining human performance in psycholinguistic tasks with models of semantic similarity based on prediction and counting: A review and empirical validation. *Journal of Memory and Language*, *92*, 57–78. https://doi.org/10.1016/j.jml.2016.04.001

Mandler, G. (1980). Recognizing: The judgment of previous occurrence. *Psychological Review*, *87*(3), 252–271. https://doi.org/10.1037/0033-295X.87.3.252

Maratos, E. J., Allan, K., & Rugg, M. D. (2000). Recognition memory for emotionally negative and neutral words: An ERP study. *Neuropsychologia*, *38*(11), 1452–1465. https://doi.org/10.1016/S0028-3932(00)00061-0

Maratos, E. J., & Rugg, M. D. (2001). Electrophysiological correlates of the retrieval of emotional and non-emotional context. *Journal of Cognitive Neuroscience*, *13*(7), 877–891. https://doi.org/10.1162/089892901753165809

Marchewka, A., Wypych, M., Michałowski, J. M., 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*, 1–14. https://doi.org/10.3389/fnhum.2016.00378

Marchewka, A., Wypych, M., Moslehi, A., Riegel, M., Michalowski, J. M., & Jednoróg, K. (2016). Arousal rather than basic emotions influence long-term recognition memory in humans. *Frontiers in Behavioral Neuroscience*, *10*, 1–10. https://doi.org/10.3389/fnbeh.2016.00198

Mather, M. (2007). Emotional arousal and memory binding: An object-based framework.

*Perspectives on Psychological Science*, *2*(1), 33–52. https://doi.org/10.1111/j.1745-6916.2007.00028.x

Mather, M., & Knight, M. (2008). The emotional harbinger effect: Poor context memory for cues that previously predicted something arousing. *Emotion*, *8*(6), 850–860. https://doi.org/10.1037/a0014087

McCullough, A. M., & Yonelinas, A. P. (2013). Cold-pressor stress after learning enhances familiarity-based recognition memory in men. *Neurobiology of Learning and Memory*, *106*, 11–17. https://doi.org/10.1016/j.nlm.2013.06.011

McGaugh, J. L. (2000). Memory - a century of consolidation. *Science*, *287*, 248–251. https://doi.org/10.1126/science.287.5451.248

McNally, R., Clancy, S., Barrett, H., & Parker, H. (2004). Inhibiting retrieval of trauma cues in adults reporting histories of childhood sexual abuse. *Cognition & Emotion*, *18*(4), 479–493. https://doi.org/10.1080/02699930341000400

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), 596–601. https://doi.org/http://dx.doi.org/10.1037/0021-843X.107.4.596

Mecklinger, A. (2010). The control of long-term memory: Brain systems and cognitive processes. *Neuroscience and Biobehavioral Reviews*, *34*(7), 1055–1065. https://doi.org/10.1016/j.neubiorev.2009.11.020

Mecklinger, A., Frings, C., & Rosburg, T. (2012). Response to Paller et al.: The role of familiarity in making inferences about unknown quantities. *Trends in Cognitive Sciences*, *16*(6), 315–316. https://doi.org/10.1016/j.tics.2012.04.009

Minnema, M. T., & Knowlton, B. J. (2008). Directed forgetting of emotional words. *Emotion*, *8*(5), 643–652. https://doi.org/10.1037/a0013441

Morris, C. D., Bransford, J. D., & Franks, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning and Verbal Behavior*, *16*(5), 519–533. https://doi.org/10.1016/S0022-5371(77)80016-9

Moscovitch, M. (1992). Memory and working-with-memory: A component process model based on modules and central systems. *Journal of Cognitive Neuroscience*, *4*(3), 257–267. https://doi.org/10.1162/jocn.1992.4.3.257

Moulds, M. L., & Bryant, R. A. (2002). Directed forgetting in acute stress disorder. *Journal of Abnormal Psychology*, *111*(1), 175–179. https://doi.org/10.1037/0021-843X.111.1.175

Moulds, M. L., & Kandris, E. (2006). The effect of practice on recall of negative material in dysphoria. *Journal of Affective Disorders*, *91*(2–3), 269–272. https://doi.org/10.1016/j.jad.2006.01.003

Murdock, B. B. (1982). A theory for the storage and retrieval of item and associative information. *Psychological Review*, *89*(6), 609–626. https://doi.org/10.1037/0033-295X.89.6.609

Murray, B. D., Anderson, M. C., & Kensinger, E. A. (2015). Older adults can suppress unwanted memories when given an appropriate strategy. *Psychology and Aging*, *30*(1), 9–25. https://doi.org/10.1037/a0038611

Murray, J. G., Howie, C. A., & Donaldson, D. I. (2015). The neural mechanism underlying recollection is sensitive to the quality of episodic memory: Event related potentials reveal a some-or-none threshold. *NeuroImage*, *120*, 298–308. https://doi.org/10.1016/j.neuroimage.2015.06.069

Muther, W. S. (1965). Erasure or partitioning in short-term memory. *Psychonomic Science*, *3*, 429–430. https://doi.org/10.3758/BF03343215

Myers, L. B., Brewin, C. R., & Power, M. J. (1998). Repressive coping and the directed forgetting of emotional material. *Journal of Abnormal Psychology*, *107*(1), 141–148.

https://doi.org/10.1037/0021-843x.107.1.141

Nelson, D. L., Evoy, C. L. M. C., & Schreiber, T. A. (2004). The university of south florida free association, rhyme, and word fragment norms. *Behavior Research Methods, Instruments, & Computers*, *36*(3), 402–407. https://doi.org/10.3758/BF03195588

Nørby, S. (2018). Forgetting and emotion regulation in mental health, anxiety and depression. *Memory*, *26*(3), 342–363. https://doi.org/10.1080/09658211.2017.1346130

Nowicka, A., Jednoróg, K., Wypych, M., & Marchewka, A. (2009). Reversed old/new effect for intentionally forgotten words: An ERP study of directed forgetting. *International Journal of Psychophysiology*, *71*(2), 97–102. https://doi.org/10.1016/j.ijpsycho.2008.06.009

Nowicka, A., Marchewka, A., Jednoróg, K., Tacikowski, P., & Brechmann, A. (2011). Forgetting of emotional information is hard: An fMRI study of directed forgetting. *Cerebral Cortex*, *21*(3), 539–549. https://doi.org/10.1093/cercor/bhq117

Ochsner, K. N. (2000). Are affective events richly recollected or simply familiar? The experience and process of recognizing feelings past. *Journal of Experimental Psychology. General*, *129*(2), 242–261. https://doi.org/10.1037/0096-3445.129.2.242

Otani, H., Libkuman, T. M., Goernert, P. N., Kato, K., Migita, M., Freehafer, S. E., & Landow, M. P. (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

Otten, L. J., Henson, R. N. A., & Rugg, M. D. (2001). Depth of processing effects on neural correlates of memory encoding: Relationship between findings from across- and within-task comparisons. *Brain*, *124*(2), 399–412. https://doi.org/10.1093/brain/124.2.399

Otten, L. J., & Rugg, M. D. (2001). Electrophysiological correlates of memory encoding are task-dependent. *Cognitive Brain Research*, *12*(1), 11–18. https://doi.org/10.1016/S0926-6410(01)00015-5

Otten, Leun J., Quayle, A. H., & Puvaneswaran, B. (2010). Prestimulus subsequent memory effects for auditory and visual events. *Journal of Cognitive Neuroscience*, *22*(6), 1212–1223. https://doi.org/10.1162/jocn.2009.21298

Paller, K. A. (1990). Recall and stem-completion priming have different electrophysiological correlates and are modified differentially by directed forgetting. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *16*(6), 1021–1032. https://doi.org/10.1037/0278-7393.16.6.1021

Paller, K. A., Kutas, M., & Mayes, A. R. (1987). Neural correlates of encoding in an incidental learning paradigm. *Electroencephalography and Clinical Neurophysiology*, *67*(4), 360–371. https://doi.org/10.1016/0013-4694(87)90124-6

Paller, K. A., McCarthy, G., & Wood, C. C. (1988). ERPs predictive of subsequent recall and recognition performance. *Biological Psychology*, *26*(1–3), 269–276. https://doi.org/10.1063/1.3583461

Paller, K. A., Voss, J. L., & Boehm, S. G. (2007). Validating neural correlates of familiarity. *Trends in Cognitive Sciences*, *11*(6), 243–250. https://doi.org/10.1016/j.tics.2007.04.002

Paller, K. A., & Wagner, A. D. (2002). Observing the transformation of experience into memory. *Trends in Cognitive Sciences*, *6*(2), 93–102. https://doi.org/10.1016/S1364-6613(00)01845-3

Parks, C. M., & Yonelinas, A. P. (2007). Moving beyond pure signal-detection models: Comment on wixted (2007). *Psychological Review*, *114*(1), 188–201. https://doi.org/10.1037/0033-295X.114.1.188

Payne, K. 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

Paz-Caballero, M. D., Menor, J., & Jiménez, J. M. (2004). Predictive validity of event-related potentials (ERPs) in relation to the directed forgetting effects. *Clinical Neurophysiology*, *115*(2), 369–377. https://doi.org/10.1016/j.clinph.2003.09.011

Pazzaglia, A. M., Dube, C., & Rotello, C. M. (2013). A critical comparison of discrete-state and continuous models of recognition memory: Implications for recognition and beyond. *Psychological Bulletin*, *139*(6), 1173–1203. https://doi.org/10.1037/a0033044

Perrin, F., Pernier, J., & Bertrand, O. (1989). Spherical splines for scalp potential and current density mapping. *Electroencephalography and Clinical Neurophysiology*, *72*, 184–187. https://doi.org/https://doi.org/10.1016/0013-4694(89)90180-6

Phelps, E. A., & Sharot, T. (2008). How (and why) emotion enhances the subjective sense of recollection. *Current Directions in Psychological Science*, *17*(2), 147–152. https://doi.org/10.1111/j.1467-8721.2008.00565.x

Picton, T. W., Bentin, S., Berg, P., Donchin, E., Hillyard, S. A., Johnson, R., Miller, G. A., Ritter, W., Ruchkin, D. S., Rugg, M. D., & Taylor, M. J. (2000). Guidelines for using human event-related potentials to study cognition: Recording standards and publication criteria. *Psychophysiology*, *37*, 127–152. https://doi.org/10.1111/1469-8986.3720127

Pierce, B. H., & Kensinger, E. A. (2011). Effects of emotion on associative recognition: valence and retention interval matter. *Emotion*, *11*(1), 139–144. https://doi.org/10.1037/a0021287

Pratto, F., & John, O. P. (1991). Automatic vigilance: The attention-grabbing power of negative social information. *Journal of Personality & Social Psychology*, *61*(3), 380–391. https://doi.org/10.1037/0022-3514.61.3.380

Rajaram, S. (1993). Remembering and knowing: Two means of access to the personal past. *Memory & Cognition*, *21*(1), 89–102. https://doi.org/10.3758/BF03211168

Ratcliff, R., Sheu, C., & Gronlund, S. D. (1992). Testing global memory models using ROC curves. *Psychological Review*, *99*(3), 518–535. https://doi.org/10.1037/0033-

295x.99.3.518

Reisberg, D., & Heuer, F. (1992). Remembering the details of emotional events. In E. Winograd & U. Neisser (Eds.), *Affect and accuracy in recall* (pp. 162–190). Cambridge University Press. https://doi.org/10.1017/CBO9780511664069.009

Reyna, V. F., & Brainerd, J. (1995). Fuzzy-trace theory: an interim synthesis. *Learning and Individual Differences*, *7*(1), 1–75. https://doi.org/10.1016/1041-6080(95)90031-4

Richards, J. M., & Gross, J. J. (2000). Emotion regulation and memory: the cognitive costs of keeping one's cool. *Journal of Personality and Social Psychology*, *79*(3), 410–424. https://doi.org/10.1037/0022-3514.79.3.410

Ritchey, M., Dolcos, F., & Cabeza, R. (2008). Role of amygdala connectivity in the persistence of emotional memories over time: An event-related fMRI investigation. *Cerebral Cortex*, *18*(11), 2494–2504. https://doi.org/10.1093/cercor/bhm262

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.

Rotello, C. M., Macmillan, N. A., & Reeder, J. A. (2004). Sum-Difference theory of remembering and knowing: A two-dimensional signal-detection model. *Psychological Review*, *111*(3), 588–616. https://doi.org/10.1037/0033-295X.111.3.588

Rotello, C. M., Macmillan, N. A., Hicks, J. L., & Hautus, M. J. (2006). Interpreting the effects of response bias on remember-know judgments using signal detection and threshold models. *Memory and Cognition*, *34*(8), 1598–1614.

Rotello, C. M., Macmillan, N. A., Reeder, J. A., & Wong, M. (2005). The remember response: subject to bias, graded, and not a process-pure indicator of recollection. *Psychonomic Bulletin & Review*, *12*(5), 865–873. https://doi.org/10.3758/BF03196778

Rugg, M. D. (1995a). Memory and consciousness: A selective review of issues and data.

*Neuropsychologia*, *33*(9), 1131–1141. https://doi.org/10.1016/0028-3932(95)00053-6

Rugg, M. D. (1995b). ERP studies of memory. In M. D. Rugg & M. G. H. Coles (Eds.), *Electrophysiology of mind. Event-related brain potentials and cognition.* (pp. 132-170). New York: Oxford University Press.

Rugg, M. D., & Coles, M. G. H. (1995). The ERP and cognitive psychology: conceptual issues. In M. D. Rugg & M. G. H. Coles (Eds.), *Electrophysiology of mind. Event-related brain potentials and cognition.* (pp. 27-39). New York: Oxford University Press

Rugg, M. D., & Curran, T. (2007). Event-related potentials and recognition memory. *Trends in Cognitive Sciences*, *11*(6), 251–257. https://doi.org/10.1016/j.tics.2007.04.004

Rugg, M. D., & Doyle, M. C. (1992). Event-related potentials and recognition memory for low-frequency and high-frequency words. *Journal of Cognitive Neuroscience*, *4*(1), 69–79. https://doi.org/10.1016/0013-4694(89)90045-X

Rugg, M. D., Mark, R. E., Walla, P., Schloerscheidt, A. M., Birch, C. S., & Allan, K. (1998). Dissociation of the neural correlates of implicit and explicit memory. *Nature*, *392*, 595–598. https://doi.org/10.1038/33396

Rugg, M. D., & Wilding, E. L. (2000). Retrieval processing and episodic memory. *Trends in Cognitive Sciences*, *4*(3), 108–115. https://doi.org/10.1016/S1364-6613(00)01445-5

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, *39*(6), 1161–1178. https://doi.org/10.1037/h0077714

Russo, R., & Grammatopoulou, N. (2003). Word length and articulatory suppression affect short-term and long-term recall tasks. *Memory and Cognition*, *31*(5), 728–737. https://doi.org/10.3758/BF03196111

Sanquist, T. F., Rohrbaugh, J. W., Syndulko, K., & Lindsley, D. B. (1980). Electrocortical signs of levels of processing: Perceptual analysis and recognition memory. *Psychophysiology*, *17*(6), 568–576. https://doi.org/10.1111/j.1469-8986.1980.tb02299.x

Sauseng, P., & Klimesch, W. (2008). What does phase information of oscillatory brain activity tell us about cognitive processes? *Neuroscience and Biobehavioral Reviews*, *32*, 1001–1013. https://doi.org/10.1016/j.neubiorev.2008.03.014

Sayers, B. M., Beagley, H. A., & Henshall, W. R. (1974). The mechanism of auditory evoked EEG responses. *Nature*, *247*, 481–483. https://doi.org/10.1038/247481a0

Schacter, D. L., Eich, J. E., & Tulving, E. (1978). Richard Semon's theory of memory. *Journal of Verbal Learning and Verbal Behavior*, *17*(6), 721–743. https://doi.org/10.1016/S0022-5371(78)90443-7

Schacter, D. L., Israel, L., & Racine, C. (1999). Suppressing false recognition in younger and older adults: The distinctiveness heuristic. *Journal of Memory and Language*, *24*, 1–24. https://doi.org/10.1006/jmla.1998.2611

Schacter, D. L., Norman, K. A., & Koutstaal, W. (1998). The cognitive neuroscience of constructive memory. *Annual Review of Psychology*, *49*, 289–318.

Schacter, D. L., & Slotnick, S. D. (2004). The cognitive neuroscience of memory distortion. *Neuron*, *44*, 149–160. https://doi.org/10.1016/j.neuron.2004.08.017

Schupp, H. T., Markus, J., Weike, A. I., & Hamm, A. O. (2003). Emotional facilitation of sensory processing in the visual cortex. *Psychological Science*, *14*(1), 7–13. https://doi.org/10.1111/1467-9280.01411

Semon, R. (1904). Die Mneme als erhaltendes Prinzip im Wechsel des organischen Geschehens. Leipzig: Wilhelm Engelmann.

Senkfor, A. J., & Van Petten, C. (1998). Who said what? An event-related potential investigation of source and item memory. *Journal of Experimental Psychology*, *24*(4), 1005–1025. https://doi.org/10.1037//0278-7393.24.4.1005

Sharot, T., Delgado, M. R., & Phelps, E. A. (2004). How emotion enhances the feeling of remembering. *Nature Neuroscience*, *7*(12), 1376–1380. https://doi.org/10.1038/nn1353

Sharot, T., Verfaelli, M., & Yonelinas, A. P. (2007). How emotion strengthens the recollective experience: A time-dependent hippocampal process. *PLoS ONE*, *2*(10), 1–10. https://doi.org/10.1371/journal.pone.0001068

Simpson, A. J., & Fitter, M. J. (1973). What is the best index of detectability? *Psychological Bulletin*, *80*(6), 481–488. https://doi.org/10.1037/h0035203

Smith, A. P. R., Dolan, R. J., & Rugg, M. D. (2004). Event-related potential correlates of the retrieval of emotional and nonemotional context. *Journal of Cognitive Neuroscience*, *16*(5), 760–775. https://doi.org/10.1162/089892904970816

Smith, M. E. (1993). Neurophysiological manifestations of recollective experience during recognition memory judgments. *Journal of Cognitive Neuroscience*, *5*(1), 1–13. https://doi.org/10.1162/jocn.1993.5.1.1

Snodgrass, J. G., & Corwin, J. (1988). Pragmatics of measuring recognition memory: applications to dementia and amnesia. *Journal of Experimental Psychology. General*, *117*(1), 34–50. https://doi.org/10.1037/0096-3445.117.1.34

Squire, L. R., & Wixted, J. T. (2011). The cognitive neuroscience of human memory since H.M. *Annual Review of Neuroscience*, *34*, 259–288. https://doi.org/10.1146/annurev-neuro-061010-113720

Squire, L. R. (2004). Memory systems of the brain: A brief history and current perspective. *Neurobiology of Learning and Memory*, *82*(3), 171–177. https://doi.org/10.1016/j.nlm.2004.06.005

Stretch, V., & Wixted, J. T. (1998). Decision rules for recognition memory confidence judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*(6), 1397–1410. https://doi.org/10.1037/0278-7393.24.6.1397

Talmi, D., Anderson, A. K., Riggs, L., Caplan, J. B., & Moscovitch, M. (2008). Immediate memory consequences of the effect of emotion on attention to pictures. *Learning &*

*Memory*, *15*(3), 172–182. https://doi.org/10.1101/lm.722908

Talmi, D., & Moscovitch, M. (2004). Can semantic relatedness explain the enhancement of memory for emotional words? *Memory and Cognition*, *32*(5), 742–751. https://doi.org/10.3758/BF03195864

Talmi, D., Schimmack, U., Paterson, T., & Moscovitch, M. (2007). The role of attention and relatedness in emotionally enhanced memory. *Emotion*, *7*(1), 89–102. https://doi.org/10.1037/1528-3542.7.1.89

Taylor, T. L., Cutmore, L., & Pries, L. (2018). Item-method directed forgetting: Effects at retrieval? *Acta Psychologica*, *183*, 116–123. https://doi.org/10.1016/j.actpsy.2017.12.004

Taylor, T. L., Quinlan, C. K., & Vullings, K. C. H. (2018). Decomposing item-method directed forgetting of emotional pictures: Equivalent costs and no benefits. *Memory and Cognition*, *46*(1), 132–147. https://doi.org/10.3758/s13421-017-0751-y

Teasdale, J. D., & Fogarty, S. J. (1979). Differential effects of induced mood on retrieval of pleasant and unpleasant events from episodic memory. *Journal of Abnormal Psychology*, *88*(3), 248–257. https://doi.org/10.1037/0021-843X.88.3.248

Tehan, G., & Tolan, G. A. (2007). Word length effects in long-term memory. *Journal of Memory and Language*, *56*(1), 35–48. https://doi.org/10.1016/j.jml.2006.08.015

Teplan, M. (2002). Fundamentals of EEG measurement. *Measurement Science Review*, *2*(2), 1–11. https://doi.org/10.1021/pr070350l

Thayer, R. E. (1978). Toward a psychological theory of multidimensional activation (arousal). *Motivation and Emotion*, *2*(1), 1–34. https://doi.org/10.1007/BF00992729

Tsivilis, D., Allan, K., Roberts, J., Williams, N., Downes, J. J., & El-Deredy, W. (2015). Old-new ERP effects and remote memories: the late parietal effect is absent as recollection fails whereas the early mid-frontal effect persists as familiarity is retained. *Frontiers in Human Neuroscience*, *9*, 1–11. https://doi.org/10.3389/fnhum.2015.00532

Tulving, E. (1972). Episodic and semantic memory. In E. Tulving & W. Donaldson (Eds.), *Organization of memory* (Vol. 1). https://doi.org/10.1017/S0140525X00047257

Tulving, E. (1985a). How many memory systems are there? *American Psychologist*, *40*(4), 385–398. https://doi.org/10.1037/0003-066X.40.4.385

Tulving, E. (1985b). Memory and consciousness. *Canadian Psychology*, *26*(1), 1–12. https://doi.org/10.1037/h0080017

Tulving, E., & Pearlstone, Z. (1966). Availability versus accessibility of information in memory for words. *Journal of Verbal Learning and Verbal Behavior*, *5*(4), 381–391. https://doi.org/10.1016/S0022-5371(66)80048-8

Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review*, *80*(5), 352–373. https://doi.org/10.1037/h0020071

Van Heuven, W. J. B., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). SUBTLEX-UK: A new and improved word frequency database for British English. *Quarterly Journal of Experimental Psychology*, *67*(6), 1176–1190. https://doi.org/10.1080/17470218.2013.850521

Van Hooff, J. C., Whitaker, T. A., & Ford, R. M. (2009). Directed forgetting in direct and indirect tests of memory: Seeking evidence of retrieval inhibition using electrophysiological measures. *Brain and Cognition*, *71*(2), 153–164. https://doi.org/10.1016/j.bandc.2009.05.001

Verde, M. F., Macmillan, N. A., & Rotello, C. M. (2006). Measures of sensitivity on a single hit rate and false alarm rate: the accuracy, precision, and robustness of d', Az, and A'. *Perception & Psychophysics*, *68*(4), 643–654.

Verde, M. F., & Rotello, C. M. (2003). Does familiarity change in the revelation effect? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*(5), 739–

746. https://doi.org/10.1037/0278-7393.29.5.739

Vilberg, K. L., & Rugg, M. D. (2008). Memory retrieval and the parietal cortex: A review of evidence from a dual-process perspective. *Neuropsychologia*, *46*(7), 1787–1799. https://doi.org/10.1016/j.neuropsychologia.2008.01.004

Vito, D. De, & Fenske, M. J. (2017). Suppressing memories of words and familiar objects results in their affective devaluation : Evidence from Think/No-think tasks. *Cognition*, *162*, 1–11. https://doi.org/10.1016/j.cognition.2017.01.020

Voss, J. L., Lucas, H. D., & Paller, K. A. (2012). More than a feeling: Pervasive influences of memory without awareness of retrieval. *Cognitive Neuroscience*, *3*(3–4), 193–226. https://doi.org/10.1080/17588928.2012.674935

Wagner, A. D., Schacter, D. L., Rotte, M., Koutstaal, W., Maril, A., Dale, A. M., Rosen, B. R., & Buckner, R. L. (1998). Building memories: Remembering and forgetting of verbal experiences as predicted by brain activity. *Science*, *281*, 1188–1191. https://doi.org/10.1126/science.281.5380.1188

Wang, Y., Mao, X., Li, B., Wang, W., & Guo, C. (2016). Dissociating the electrophysiological correlates between item retrieval and associative retrieval in associative recognition: From the perspective of directed forgetting. *Frontiers in Psychology*, *7*, 1–14. https://doi.org/10.3389/fpsyg.2016.01754

Warren, L. R. (1980). Evoked potential correlates of recognition memory. *Biological Psychology*, *11*, 21–35. https://doi.org/10.1016/0301-0511(80)90023-x

Warriner, A. B., Kuperman, V., & Brysbaert, M. (2013). Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behavior Research Methods*, *45*(4), 1191–1207. https://doi.org/10.3758/s13428-012-0314-x

Watkins, M. J., & Peynircioglu, Z. F. (1990). The revelation effect: when disguising test items induces recognition. *Journal of Experimental Psychology: Learning, Memory, and*

*Cognition*, *16*(6), 1012–1020. https://doi.org/10.1037/0278-7393.16.6.1012

Wessel, I., & Merckelbach, H. (2006). Forgetting "murder" is not harder than forgetting "circle": Listwise-directed forgetting of emotional words. *Cognition & Emotion*, *20*(1), 129–137. https://doi.org/10.1080/02699930500260195

Wheeler, M. A., Stuss, D. T., & Tulving, E. (1997). Toward a theory of episodic memory: the frontal lobes and autonoetic consciousness. *Psychological Bulletin*, *121*(3), 331–354. https://doi.org/10.1037/0033-2909.121.3.331

White, C. N., Kapucu, A., Bruno, D., Rotello, C. M., & Ratcliff, R. (2015). Memory bias for negative emotional words in recognition memory is driven by effects of category membership. *Cognition & Emotion*, *28*(5), 867–880. https://doi.org/10.1126/science.1249098.Sleep

Wilding, E. L. (1999). Separating retrieval strategies from retrieval success: An event- related potential study of source memory. *Neuropsychologia*, *37*(4), 441–454. https://doi.org/10.1016/S0028-3932(98)00100-6

Wilding, E. L. (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

Wilding, E. L., Doyle, M. C., & Rugg, M. D. (1995). Recognition memory with and without retrieval of context: An event-related potential study. *Neuropsychologia*, *33*(6), 743–767. https://doi.org/10.1016/0028-3932(95)00017-W

Wilding, E. L., & Nobre, A. C. (2001). Task-switching and memory retrieval processing: electrophysiological evidence. *NeuroReport*, *12*(16), 3613–3617. https://doi.org/10.1097/00001756-200111160-00048

Wilding, E. L., & Rugg, M. D. (1996). An event-related potential study of recognition memory with and without retrieval of source. *Brain*, *119*, 889–906.

https://doi.org/10.1093/brain/119.3.889

Williams, J. M. G., & Broadbent, K. (1986). Autobiographical memory in suicide attempters. *Journal of Abnormal Psychology*, *95*(2), 144–149. https://doi.org/10.1037//0021-843x.95.2.144

Windmann, S., & Kutas, M. (2001). Electrophysiological correlates of emotion-induced recognition bias. *Journal of Cognitive Neuroscience*, *13*(1998), 577–592. https://doi.org/10.1162/089892901750363172

Windmann, S., Urbach, T. P., & Kutas, M. (2002). Cognitive and neural mechanisms of decision biases in recognition memory. *Cerebral Cortex*, *12*(8), 808–817. https://doi.org/10.1093/cercor/12.8.808

Wixted, J. T. (2007). Dual-process theory and signal-detection theory of recognition memory. *Psychological Review*, *114*(1), 152–176. https://doi.org/10.1037/0033-295X.114.1.152

Wixted, J. T., & Mickes, L. (2014). A signal-detection-based diagnostic-feature-detection model of eyewitness identification. *Psychological Review*, *121*(2), 262–276. https://doi.org/10.1037/a0035940

Wixted, J. T., & Stretch, V. (2004). In defense of the signal detection interpretation of remember/know judgments. *Psychonomic Bulletin and Review*, *11*(4), 616–641. https://doi.org/10.3758/BF03196616

Woodruff, C. C., Hayama, H. R., & Rugg, M. D. (2006). Electrophysiological dissociation of the neural correlates of recollection and familiarity. *Brain Research*, *1100*(1), 125–135. https://doi.org/10.1016/j.brainres.2006.05.019

Xie, H., Jiang, D., & Zhang, D. (2018). Individuals with depressive tendencies experience difficulty in forgetting negative material: Two mechanisms revealed by ERP data in the directed forgetting paradigm. *Scientific Reports*, *8*(1), 1–14. https://doi.org/10.1038/s41598-018-19570-0

Yang, T., Lei, X., & Anderson, M. (2016). Decreased inhibitory control of negative information in directed forgetting. *International Journal of Psychophysiology*, *100*, 44–51. https://doi.org/10.1016/j.ijpsycho.2015.09.007

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

Yick, Y. Y., Buratto, L. G., & Schaefer, A. (2015). The effects of negative emotion on encoding-related neural activity predicting item and source recognition. *Neuropsychologia*, *73*, 48–59. https://doi.org/10.1016/j.neuropsychologia.2015.04.030

Yonelinas, A. P. (1994). Receiver-operating characteristics in recognition memory: evidence for a dual-process model. *Journal of Experimental Psychology*, *20*(6), 1341–1354. https://doi.org/10.1037//0278-7393.20.6.1341

Yonelinas, A. P. (1999). The contribution of recollection and familiarity to recognition and source-memory judgments: A formal dual-process model and an analysis of receiver operating characteristics. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*(6), 1415–1434. https://doi.org/10.1037//0278-7393.25.6.1415

Yonelinas, A. P. (2001a). Components of episodic memory: The contribution of recollection and familiarity. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *356*, 1363–1374. https://doi.org/10.1098/rstb.2001.0939

Yonelinas, A. P. (2001b). Consciousness, control and confidence: The 3 Cs of recognition memory. *Journal of Experimental Psychology. General*, *130*(3), 361–379.

Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language*, *46*(3), 441–517. https://doi.org/10.1006/jmla.2002.2864

Yonelinas, A. P., Dobbins, I., Szymanski, M. D., Dhaliwal, H. S., & King, L. (1996). Signal-detection, threshold, and dual-process models of recognition memory: ROCs and conscious recollection. *Consciousness and Cognition*, *5*(4), 418–441. https://doi.org/10.1006/ccog.1996.0026

Yonelinas, A. P., & Jacoby, L. L. (1995). The relation between remembering and knowing as bases for recognition: Effects of size congruency. *Journal of Memory and Language*, *34*, 622–643.

Yonelinas, A. P., Kroll, N. E. A., Dobbins, I. G., & Soltani, M. (1999). Recognition memory for faces: When familiarity supports associative recognition judgments. *Psychonomic Bulletin and Review*, *6*(4), 654–661. https://doi.org/10.3758/BF03212975

Yonelinas, A. P., Kroll, N. E. A., Dobbins, I., Lazzara, M., & Knight, R. T. (1998). Recollection and familiarity deficits in amnesia: Convergence of remember-know, process dissociation, and receiver operating characteristic data. *Neuropsychology*, *12*(3), 323–339.

Yonelinas, A. P., Otten, L. J., Shaw, K. N., & Rugg, M. D. (2005). Separating the brain regions involved in recollection and familiarity in recognition memory. *Journal of Neuroscience*, *25*(11), 3002–3008. https://doi.org/10.1523/JNEUROSCI.5295-04.2005

Yonelinas, A. P., & Parks, C. M. (2007). Receiver operating characteristics (ROCs) in recognition memory: a review. *Psychological Bulletin*, *133*(5), 800–832. https://doi.org/10.1037/0033-2909.133.5.800

Yonelinas, A. P., Parks, C. M., Koen, J. D., Jorgenson, J., & Mendoza, S. P. (2011). The effects of post-encoding stress on recognition memory: Examining the impact of skydiving in young men and women. *Stress*, *14*(2), 136–144. https://doi.org/10.1038/jid.2014.371

Yonelinas, A. P., & Ritchey, M. (2015). The slow forgetting of emotional episodic memories:

An emotional binding account. In *Trends in Cognitive Sciences* (Vol. 19, Issue 5, pp. 259–267). https://doi.org/10.1016/j.tics.2015.02.009

Yonelinas, A. P., Wei-Chun Wang, M. A., & Koen, J. D. (2010). Recollection and familiarity: Examining controversial assumptions and new directions. *Hippocampus*, *20*(11), 1178–1194. https://doi.org/10.1038/jid.2014.371

Zheng, Z., Lang, M., Wang, W., Xiao, F., & Li, J. (2018). Electrophysiological evidence for the effects of emotional content on false recognition memory. *Cognition*, *179*(16), 298–310. https://doi.org/S0010027718301653

## Appendix A

## Experiment 1 – Stimuli Analyses

Table A1

*Outcomes of Independent Sample T-tests for Assessments on Equality for Valence, Absolute Valence, Arousal, Word Length, Word Frequency and Semantic Relatedness for the three Lists of 96 Words between Valence Type (Negative, Positive and Neutral) in Experiment 1. SD = standard deviation*

| | Negative vs. Neutral | | | Negative vs. Positive | | | Positive vs. Neutral | | |
|---|---|---|---|---|---|---|---|---|---|
| | Negative | Neutral | | Negative | Positive | | Positive | Neutral | |
| | M (SD) | M (SD) | t-value (*p*) | M (SD) | M (SD) | t-value (*p*) | M (SD) | M (SD) | t-value (*p*) |
| Valence | 2.92 (0.63) | 5.45 (0.42) | -32.66 (<.001)*** | 2.92 (0.63) | 7.34 (0.49) | -54.41 (<.001)*** | 7.34 (0.49) | 5.45 (0.42) | 28.67 (<.001)*** |
| Absolute Valence | 2.08 (0.63) | 0.53 (0.32) | 21.52 (<.001)*** | 2.08 (0.63) | 2.34 (0.49) | -3.18 (.002)** | 2.34 (0.49) | 0.53 (0.32) | 30.37 (<.001)*** |
| Arousal | 4.67 (0.90) | 3.92 (0.77) | 6.25 (<.001)*** | 4.67 (0.90) | 4.62 (0.94) | 0.36 (.717) | 4.62 (0.94) | 3.92 (0.77) | 5.73 (<.001)*** |
| Word length | 6.96 (2.07) | 6.99 (2.03) | -0.11 (.916) | 6.96 (2.07) | 7.33 (2.38) | -1.17 (.245) | 7.33 (2.38) | 6.99 (2.03) | 1.08 (.283) |
| Word frequency | 2.76 (0.68) | 2.96 (0.79) | -1.70 (.091) | 2.76 (0.68) | 3.03 (0.67) | -2.59 (.010)** | 3.03 (0.67) | 2.96 (0.79) | 0.68 (.495) |
| Semantic distance | 0.88 (0.02) | 0.89 (0.03) | -0.57 (.572) | 0.88 (0.02) | 0.89 (0.03) | -1.01 (.313) | 0.89 (0.03) | 0.89 (0.03) | 0.43 (.670) |

*Notes*. ** p < .01, *** p < .001.

Table A2

*Outcomes of Independent Sample T-tests for Assessments on Equality for Arousal, Word Length, Word Frequency and Semantic Relatedness for the Study and Test Lists between Valence Type (Negative, Positive and Neutral) in Experiment 1. SD = standard deviation*

| | | | Negative vs Neutral | | | Negative vs Positive | | | Positive vs Neutral | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Negative | Neutral | | Negative | Positive | | Positive | Neutral | |
| | | | M (SD) | M (SD) | t-value (*p*) | M (SD) | M (SD) | t-value (*p*) | M (SD) | M (SD) | t-value (*p*) |
| Study lists | A | Arousal | 4.74 (0.77) | 3.89 (0.68) | 4.13 (.000)*** | 4.74 (0.77) | 4.52 (0.66) | 1.03 (.308) | 4.52 (0.66) | 3.89 (0.68) | 3.37 (.002)** |
| | | Word length | 6.67 (2.14) | 7.25 (1.33) | -1.14 (.262) | 6.67 (2.14) | 7.33 (2.57) | -0.98 (.333) | 7.33 (2.57) | 7.25 (1.33) | 0.14 (.888) |
| | | Word frequency | 2.97 (0.73) | 2.92 (0.79) | 0.25 (.805) | 2.97 (0.73) | 3.13 (0.61) | -0.82 (.418) | 3.13 (0.61) | 2.92 (0.79) | 1.05 (.302) |
| | | Semantic distance | 0.86 (0.03) | 0.87 (0.03) | -1.05 (.298) | 0.86 (0.03) | 0.88 (0.04) | -1.83 (.073) | 0.88 (0.04) | 0.87 (0.03) | 0.89 (.380) |
| | B | Arousal | 4.72 (1.01) | 3.99 (0.94) | 2.63 (.011)* | 4.72 (1.01) | 4.56 (1.03) | 0.57 (.571) | 4.56 (1.03) | 3.99 (0.94) | 2.01 (.050) |
| | | Word length | 7.29 (1.99) | 6.54 (2.25) | 1.23 (.227) | 7.29 (1.99) | 7.17 (2.30) | 0.20 (.841) | 7.17 (2.30) | 6.54 (2.25) | 0.95 (.345) |
| | | Word frequency | 2.78 (0.64) | 3.13 (0.83) | -1.67 (.102) | 2.78 (0.64) | 3.04 (0.79) | -1.29 (.205) | 3.04 (0.79) | 3.13 (0.83) | -0.39 (.701) |
| | | Semantic distance | 0.89 (0.02) | 0.90 (0.02) | -1.61 (.115) | 0.89 (0.02) | 0.89 (0.03) | -1.01 (.318) | 0.89 (0.03) | 0.90 (0.02) | -0.41 (.682) |
| Test lists | A | Arousal | 4.70 (0.86) | 3.95 (0.74) | 4.57 (.000)*** | 4.70 (0.86) | 4.53 (0.78) | 1.00 (.320) | 4.53 (0.78) | 3.95 (0.74) | 3.75 (.000)*** |
| | | Word length | 7.02 (2.17) | 7.38 (1.59) | -0.91 (.364) | 7.02 (2.17) | 7.44 (2.28) | -0.92 (.361) | 7.44 (2.28) | 7.38 (1.59) | 0.16 (.877) |
| | | Word frequency | 2.81 (0.66) | 2.98 (0.72) | -1.27 (.209) | 2.81 (0.66) | 3.10 (0.61) | -2.25 (.027)* | 3.10 (0.61) | 2.98 (0.72) | 0.83 (.409) |
| | | Semantic distance | 0.88 (0.03) | 0.87 (0.04) | 1.39 (.169) | 0.88 (0.03) | 0.88 (0.03) | 0.33 (.741) | 0.88 (0.03) | 0.87 (0.04) | 1.01 (.317) |
| | B | Arousal | 4.65 (0.96) | 3.88 (0.80) | 4.25 (.000)*** | 4.65 (0.96) | 4.72 (1.08) | -0.34 (.734) | 4.72 (1.08) | 3.88 (0.80) | 4.32 (.000)*** |
| | | Word length | 6.90 (1.98) | 6.90 (2.34) | 0.66 (.512) | 6.90 (1.98) | 7.23 (2.50) | -0.73 (.470) | 7.23 (2.50) | 6.90 (2.34) | 1.27 (.209) |
| | | Word frequency | 2.75 (0.71) | 2.93 (0.85) | -1.14 (.258) | 2.75 (0.71) | 2.96 (0.72) | -1.47 (.146) | 2.96 (0.72) | 2.93 (0.85) | 0.19 (.848) |
| | | Semantic distance | 0.89 (0.03) | 0.89 (0.03) | -0.56 (.579) | 0.89 (0.03) | 0.89 (0.03) | -1.08 (.283) | 0.89 (0.03) | 0.89 (0.03) | 0.55 (.586) |

*Notes.* * $p < .05$, ** $p < .01$, *** $p < .001$

## Appendix B

### Experiment 1 – Emotion Regulation and Memory Control of Emotional Material

Scatterplots illustrating the relationship between emotion regulation strategies (cognitive reappraisal and expressive suppression) and memory sensitivity separated across each critical condition for instructions (remember and forget) and for valence (negative, positive and neutral).

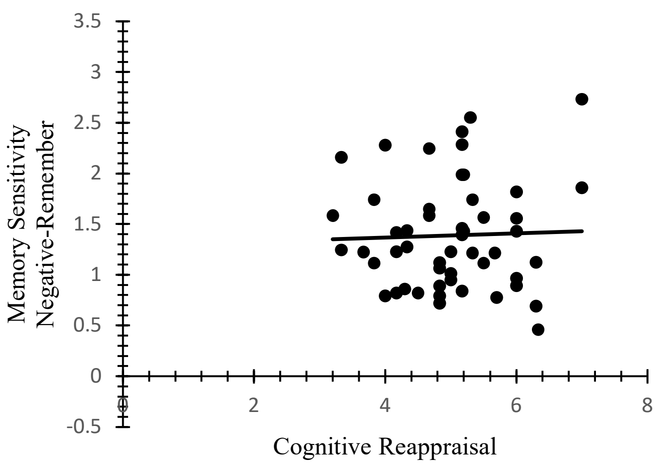



*Figure B1.* Scatterplot of the mean cognitive reappraisal scores against memory sensitivity in the remember condition for negative items

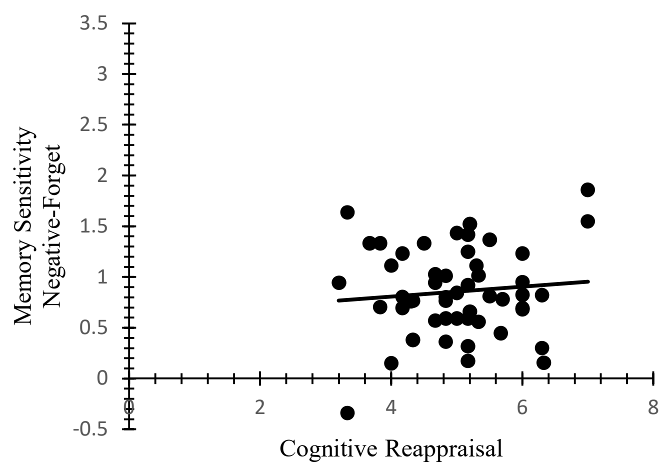



*Figure B2.* Scatterplot of the mean cognitive reappraisal scores against memory sensitivity in the forget condition for negative items

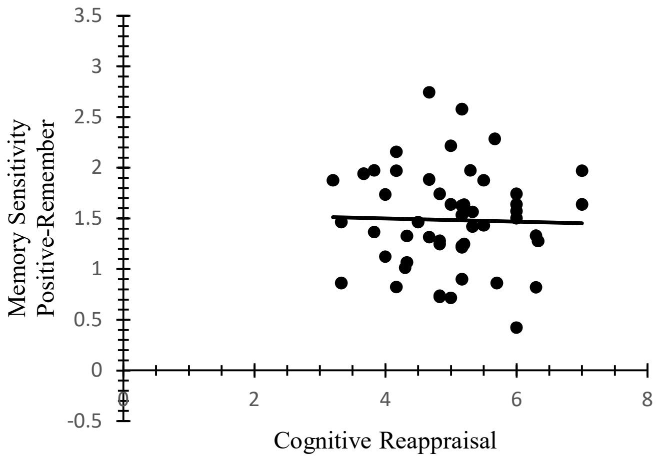



*Figure B3.* Scatterplot of the mean cognitive reappraisal scores against memory sensitivity in the remember condition for positive items

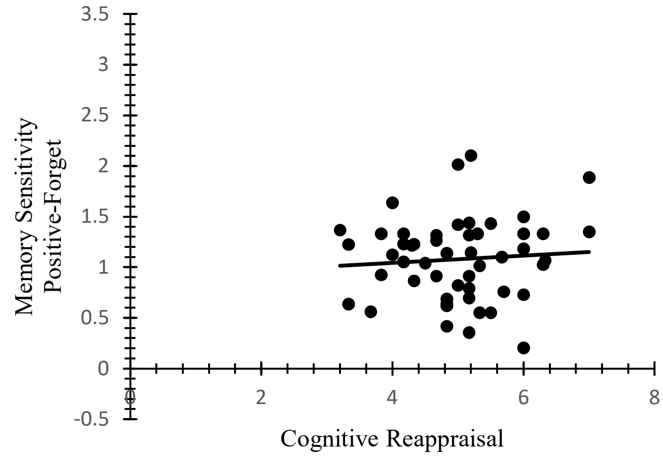



*Figure B4.* Scatterplot of the mean cognitive reappraisal scores against memory sensitivity in the forget condition for positive items

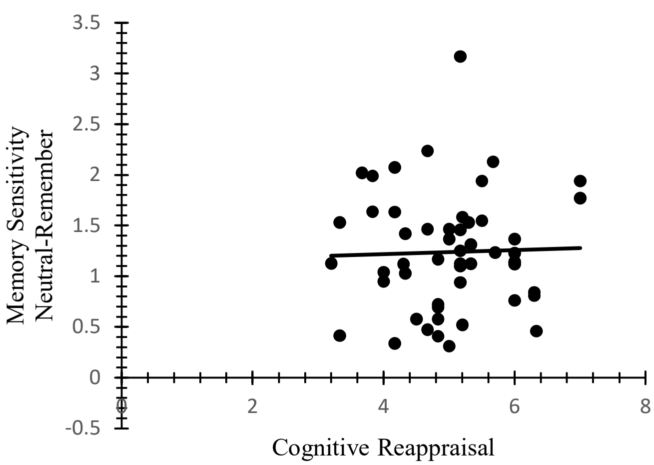



*Figure B5.* Scatterplot of the mean cognitive reappraisal scores against memory sensitivity in the remember condition for neutral items

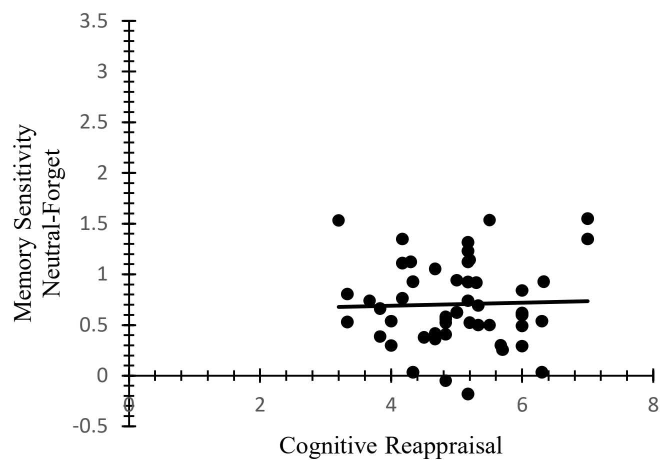



*Figure B6.* Scatterplot of the mean cognitive reappraisal scores against memory sensitivity in the remember condition for neutral items

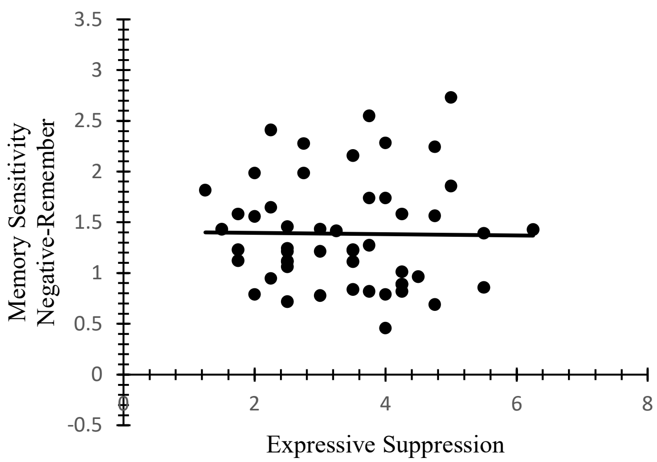



*Figure B7.* Scatterplot of the mean expressive suppression scores against memory sensitivity in the remember condition for negative items

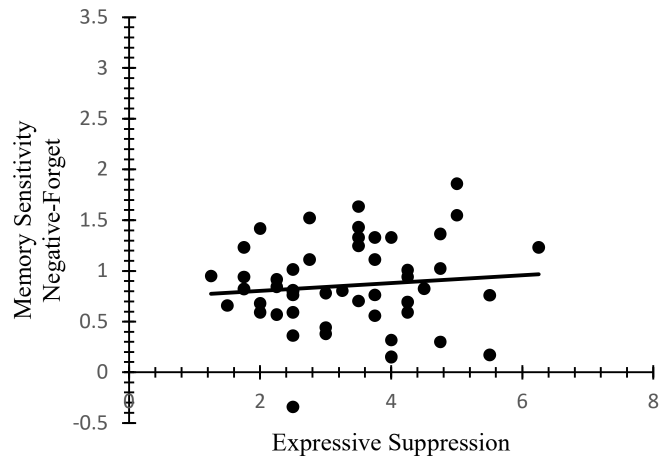



*Figure B8.* Scatterplot of the mean expressive suppression scores against memory sensitivity in the forget condition for negative items

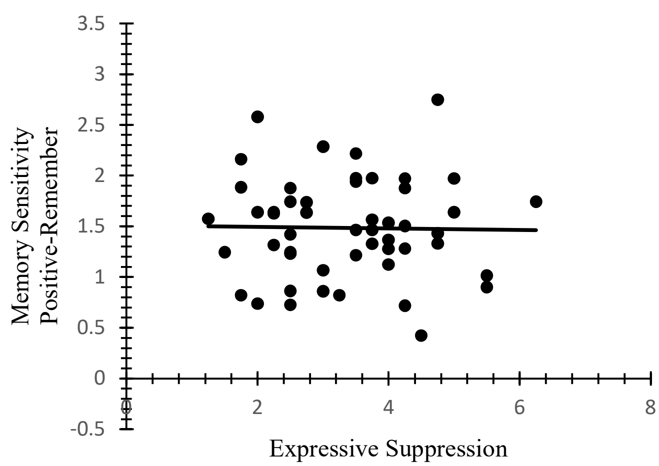



*Figure B9.* Scatterplot of the mean expressive suppression scores against memory sensitivity in the remember condition for positive items

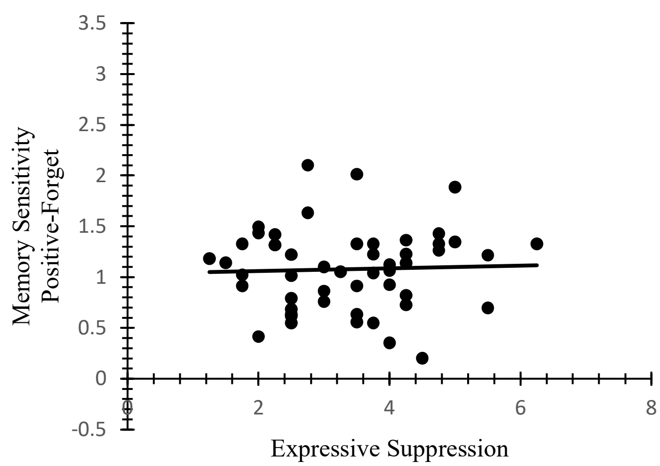



*Figure B10.* Scatterplot of the mean expressive suppression scores against memory sensitivity in the forget condition for positive items

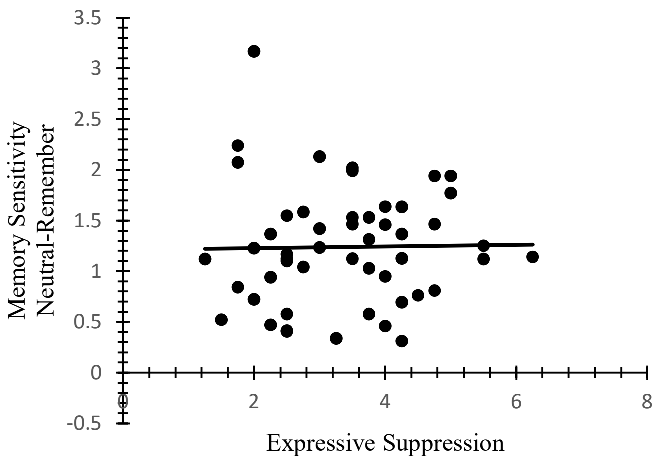



*Figure B11.* Scatterplot of the mean expressive suppression scores against memory sensitivity in the remember condition for neutral items

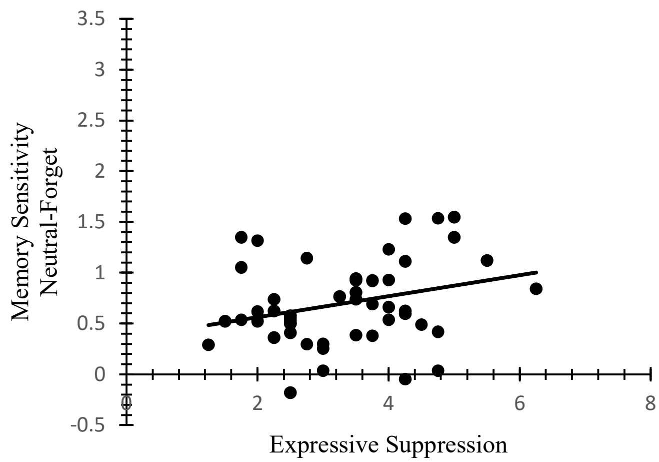



*Figure B12.* Scatterplot of the mean expressive suppression scores against memory sensitivity in the forget condition for neutral items

## Appendix C

### Experiment 2 – Stimuli Analyses

Table C1

*Outcomes of Independent Sample T-tests for Assessments on Equality for Valence, Absolute Valence, Arousal, Word Length, Word Frequency and Semantic Relatedness for the Lists of 180 Words (for Each Valence) between Valence Type (Negative, Positive and Neutral) for Studies 1, 2 and 3 in Experiment 2 and 3. SD = standard deviation*

| | Study 1 | | | Study 2 | | | Study 3 | | |
|---|---|---|---|---|---|---|---|---|---|
| | Negative | Neutral | | Negative | Positive | | Positive | Neutral | |
| | M (SD) | M (SD) | t-value (*p*) | M (SD) | M (SD) | t-value (*p*) | M (SD) | M (SD) | t-value (*p*) |
| Valence | 2.56 (0.63) | 5.12 (0.42) | -45.85 (<.001)*** | 2.56 (0.63) | 7.33 (0.57) | -75.46 (<.001)*** | 7.33 (0.57) | 5.12 (0.42) | 41.76 (<.001)*** |
| Absolute Valence | 2.44 (0.63) | 0.31 (0.30) | 41.19 (<.001)*** | 2.44 (0.63) | 2.33 (0.57) | 1.84 (.067) | 2.33 (0.57) | 0.31 (0.30) | 41.80 (<.001)*** |
| Arousal | 4.94 (0.92) | 4.18 (0.82) | 8.27 (<.001)*** | 4.94 (0.92) | 4.61 (0.98) | 3.28 (.001)** | 4.61 (0.98) | 4.18 (0.82) | 4.54 (<.001)*** |
| Word length | 7.37 (2.17) | 7.10 (2.06) | 1.22 (.223) | 7.37 (2.17) | 7.38 (2.23) | -0.02 (.981) | 7.38 (2.23) | 7.10 (2.06) | 1.23 (.220) |
| Word frequency | 2.72 (0.67) | 2.78 (0.82) | -0.72 (.474) | 2.72 (0.67) | 2.91 (0.67) | -2.69 (.007)** | 2.91 (0.67) | 2.78 (0.82) | 1.70 (.091) |
| Semantic distance | 0.89 (0.02) | 0.89 (0.02) | -0.20 (.843) | 0.89 (0.02) | 0.89 (0.03) | -0.11 (.915) | 0.89 (0.03) | 0.89 (0.02) | -0.06 (.954) |

*Notes.* ** *p* < .01, *** *p* < .001.

Table C2

*Outcomes of Independent Sample T-tests for Assessments on Equality for Arousal, Word Length, Word Frequency and Semantic Relatedness for the Study and Tests Lists (for Each Valence) between Valence Type (Negative and Neutral, and Negative and Positive) for Studies 1 and 2 in Experiment 2 and 3. SD = standard deviation*

| | | | Study 1 | | | Study 2 | | |
|---|---|---|---|---|---|---|---|---|
| | | | Negative | Neutral | | Negative | Positive | |
| | | | M (SD) | M (SD) | t-value (*p*) | M (SD) | M (SD) | t-value (*p*) |
| Study lists | A1 | Arousal | 4.87 (0.76) | 4.40 (0.63) | 2.62 (.011)* | 4.71 (0.67) | 4.54 (0.78) | 0.93 (.358) |
| | | Word length | 7.23 (2.14) | 6.47 (1.57) | 1.58 (.120) | 7.40 (2.51) | 7.27 (2.18) | 0.22 (.827) |
| | | Word frequency | 2.70 (0.50) | 2.96 (0.73) | -1.63 (.110) | 2.61 (0.57) | 2.84 (0.61) | -1.55 (.126) |
| | | Semantic distance | 0.89 (0.01) | 0.89 (0.02) | -0.17 (.862) | 0.88 (0.02) | 0.90 (0.03) | -1.89 (.064) |
| | A2 | Arousal | 5.09 (0.97) | 4.22 (0.82) | 3.73 (<.001)*** | 5.08 (1.02) | 4.61 (1.00) | 1.80 (.076) |
| | | Word length | 7.50 (2.18) | 7.33 (1.81) | 0.32 (.748) | 7.50 (1.93) | 7.23 (2.34) | 0.48 (.632) |
| | | Word frequency | 2.83 (0.73) | 2.60 (0.83) | 1.16 (.250) | 2.81 (0.67) | 3.01 (0.61) | -1.16 (.249) |
| | | Semantic distance | 0.88 (0.02) | 0.88 (0.03) | -0.17 (.868) | 0.89 (0.02) | 0.88 (0.02) | 1.28 (.206) |
| | B1 | Arousal | 5.09 (0.98) | 4.36 (0.93) | 2.97 (.004)** | 5.04 (0.92) | 4.60 (1.02) | 1.75 (.085) |
| | | Word length | 7.47 (2.11) | 7.47 (2.26) | 0.00 (1.000) | 6.97 (2.20) | 7.37 (2.50) | -0.66 (.513) |
| | | Word frequency | 2.58 (0.55) | 2.62 (0.67) | -0.24 (.814) | 2.68 (0.79) | 2.96 (0.70) | -1.44 (.156) |
| | | Semantic distance | 0.89 (0.02) | 0.90 (0.03) | -1.00 (.321) | 0.88 (0.02) | 0.88 (0.03) | -0.17 (.863) |
| | B2 | Arousal | 4.86 (0.79) | 4.25 (0.91) | 2.74 (.008)** | 5.00 (1.01) | 4.79 (0.92) | 0.86 (.393) |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | Word length | 7.63 (2.54) | 6.93 (2.29) | 1.12 (.267) | 7.40 (1.80) | 7.10 (1.56) | 0.69 (.492) |
| | | Word frequency | 2.72 (0.70) | 2.69 (0.86) | 0.16 (.877) | 2.64 (0.61) | 2.82 (0.70) | -1.04 (.30) |
| | | Semantic distance | 0.89 (0.02) | 0.90 (0.02) | -1.85 (.070) | 0.90 (0.02) | 0.90 (0.03) | 0.02 (.983) |
| | C1 | Arousal | 4.70 (1.03) | 3.89 (0.73) | 3.54 (.001)** | 4.95 (0.76) | 4.57 (0.96) | 1.72 (.090) |
| | | Word length | 7.10 (2.01) | 7.17 (2.31) | -0.12 (.905) | 7.57 (2.58) | 7.60 (2.28) | -0.05 (.958) |
| | | Word frequency | 2.74 (0.78) | 2.83 (0.96) | -0.41 (.684) | 2.89 (0.70) | 3.06 (0.77) | -0.86 (.393) |
| | | Semantic distance | 0.89 (0.02) | 0.89 (0.03) | 0.30 (.767) | 0.89 (0.02) | 0.89 (0.03) | -0.64 (.525) |
| | C2 | Arousal | 5.08 (0.94) | 3.95 (0.81) | 4.99 (<.001)*** | 4.59 (1.05) | 4.63 (1.11) | -0.14 (.886) |
| | | Word length | 7.13 (2.21) | 7.23 (2.03) | -0.18 (.856) | 7.43 (2.03) | 7.73 (2.48) | -0.51 (.610) |
| | | Word frequency | 2.83 (0.71) | 2.95 (0.77) | -0.66 (.510) | 2.67 (0.64) | 2.77 (0.65) | -0.57 (.573) |
| | | Semantic distance | 0.89 (0.02) | 0.88 (0.03) | 1.52 (.133) | 0.90 (0.03) | 0.88 (0.04) | 1.56 (.124) |
| Test lists | A | Arousal | 4.97 (0.87) | 4.31 (0.73) | 4.56 (<.001)*** | 4.87 (0.80) | 4.74 (0.84) | 0.83 (.406) |
| | | Word length | 7.37 (2.15) | 6.90 (1.73) | 1.31 (.193) | 7.51 (2.31) | 7.47 (2.40) | 0.12 (.908) |
| | | Word frequency | 2.77 (0.62) | 2.78 (0.80) | -0.10 (.920) | 2.71 (0.63) | 2.93 (0.61) | -1.90 (.060) |
| | | Semantic distance | 0.88 (0.02) | 0.88 (0.02) | -0.32 (.748) | 0.89 (0.02) | 0.89 (0.02) | 0.26 (.798) |
| | B | Arousal | 4.97 (0.89) | 4.30 (0.92) | 5.06 (<.001)*** | 5.02 (0.96) | 4.69 (0.97) | 1.86 (.065) |
| | | Word length | 7.55 (2.32) | 7.20 (2.27) | 0.84 (.405) | 7.18 (2.00) | 7.23 (2.07) | -0.13 (.893) |
| | | Word frequency | 2.65 (0.63) | 2.65 (0.78) | -0.02 (.983) | 2.66 (0.70) | 2.89 (0.70) | -1.78 (.078) |
| | | Semantic distance | 0.89 (0.02) | 0.90 (0.02) | -1.91 (.058) | 0.89 (0.02) | 0.89 (0.03) | -0.36 (.717) |
| | C | Arousal | 4.89 (0.99) | 3.92 (0.77) | 6.00 (<.001)*** | 4.77 (0.93) | 4.60 (1.03) | 0.97 (.336) |
| | | Word length | 7.12 (2.09) | 7.20 (2.15) | -0.22 (.830) | 7.50 (2.30) | 7.67 (2.36) | -0.39 (.696) |
| | | Word frequency | 2.78 (0.74) | 2.89 (0.86) | -0.75 (.458) | 2.78 (0.68) | 2.91 (0.72) | -1.01 (.314) |
| | | Semantic distance | 0.89 (0.02) | 0.89 (0.02) | -0.30 (.764) | 0.89 (0.02) | 0.89 (0.03) | 1.08 (.283) |

*Notes.* * $p < .05$, ** $p < .01$, *** $p < .001$.

Table C3

*Outcomes of Independent Sample T-tests for Assessments on Equality for Arousal, Word Length, Word Frequency and Semantic Relatedness for the Study and Test Lists (for Each Valence) between Valence Type (Positive and Neutral) for Study 3 in Experiment 2 and 3. SD = standard deviation*

| | | | Study 3 | | |
|---|---|---|---|---|---|
| | | | Positive | Neutral | |
| | | | M (SD) | M (SD) | t-value (*p*) |
| Study lists | A1 | Arousal | 4.45 (1.08) | 4.24 (0.83) | 0.84 (.405) |
| | | Word length | 7.13 (1.78) | 7.43 (2.18) | -0.59 (.561) |
| | | Word frequency | 2.88 (0.58) | 2.90 (0.57) | -0.10 (.920) |
| | | Semantic distance | 0.88 (0.02) | 0.89 (0.03) | -1.82 (.075) |
| | A2 | Arousal | 4.84 (1.09) | 4.22 (0.85) | 2.48 (.016)* |
| | | Word length | 7.77 (2.76) | 6.73 (1.82) | 1.71 (.092) |
| | | Word frequency | 2.12 (0.49) | 2.53 (0.82) | -2.40 (.019)* |
| | | Semantic distance | 0.89 (0.03) | 0.88 (0.03) | 1.65 (.105) |
| | B1 | Arousal | 4.49 (0.88) | 4.11 (0.75) | 1.79 (.078) |
| | | Word length | 6.93 (2.16) | 7.03 (1.81) | -0.19 (.847) |

| | | | | | |
|---|---|---|---|---|---|
| | | Word frequency | 3.07 (0.67) | 2.76 (0.94) | 1.50 (.140) |
| | | Semantic distance | 0.87 (0.03) | 0.88 (0.03) | -1.39 (.171) |
| | B2 | Arousal | 4.55 (0.97) | 4.28 (0.80) | 1.18 (.244) |
| | | Word length | 7.23 (1.85) | 6.57 (1.74) | 1.44 (.156) |
| | | Word frequency | 3.03 (0.57) | 2.78 (0.81) | 1.36 (.179) |
| | | Semantic distance | 0.89 (0.03) | 0.88 (0.02) | 1.11 (.272) |
| | C1 | Arousal | 4.57 (0.90) | 4.13 (0.72) | 2.10 (.040)* |
| | | Word length | 6.90 (2.29) | 7.80 (2.52) | -1.45 (.154) |
| | | Word frequency | 3.26 (0.54) | 2.72 (0.94) | 2.72 (.009)** |
| | | Semantic distance | 0.88 (0.03) | 0.89 (0.02) | -0.92 (.364) |
| | C2 | Arousal | 4.65 (0.96) | 4.15 (1.03) | 1.98 (.053) |
| | | Word length | 8.10 (2.41) | 7.13 (2.26) | 1.60 (.114) |
| | | Word frequency | 3.10 (0.55) | 2.96 (0.76) | 0.81 (.420) |
| | | Semantic distance | 0.91 (0.03) | 0.90 (0.02) | 1.44 (.156) |
| | A | Arousal | 4.65 (1.09) | 4.23 (0.83) | 2.35 (.020)* |
| | | Word length | 7.45 (2.32) | 7.08 (2.02) | 0.92 (.358) |
| | | Word frequency | 2.50 (0.66) | 2.72 (0.72) | -1.72 (.089) |
| | | Semantic distance | 0.89 (0.03) | 0.89 (0.02) | 1.03 (.304) |
| | B | Arousal | 4.52 (0.92) | 4.19 (0.77) | 2.09 (.038)* |
| Test lists | | Word length | 7.08 (2.00) | 6.80 (1.77) | 0.82 (.414) |
| | | Word frequency | 3.05 (0.62) | 2.77 (0.87) | 2.04 (.044)* |
| | | Semantic distance | 0.88 (0.03) | 0.88 (0.02) | -0.64 (.526) |
| | C | Arousal | 4.66 (0.93) | 4.10 (0.87) | 3.39 (.001)** |
| | | Word length | 7.60 (2.34) | 7.42 (2.33) | 0.43 (.668) |
| | | Word frequency | 3.18 (0.54) | 2.84 (0.86) | 2.58 (.011)* |
| | | Semantic distance | 0.89 (0.03) | 0.90 (0.02) | -1.61 (.110) |

Notes. * p < .05, ** p < .01.

## Appendix D

### Experiment 1, 2 & 3 – Differences in Word Frequency

The analyses of the word frequency data in Experiments 1 – 3 were completed over estimates obtained using the SUBTLEX-UK count variable, which is a widely used measure (Van Heuven et al., 2014). During the course of the PhD research, a decision was made to equate word frequency differently, by using values based on a log transformation, better reflecting the distribution of word frequencies among the words sets that were employed. This method was employed from Experiment 4 onwards. Re-analysing the stimulus-set data of the earlier experiments using log transformed values revealed reliable, although small, differences between frequencies (see Appendix A for Experiment 1 and Appendix C for Experiment 2 and 3).

How might these differences affect the interpretation of the results observed in these experiments? In recognition memory tests, sensitivity is commonly higher for low frequency words than for high frequency words (Glanzer & Bowles, 1976; MacLeod & Kampe, 1996; Rugg & Doyle, 1992). As a result, for any valence categories with reliable differences in word frequency, reliable sensitivity differences could be attributable to this variable rather than to valence. What the appendices show, however, is that there is no consistent relationship between the occurrence of differences in word frequencies and evidence for differences between sensitivities. To give two examples: in Experiment 2 the small differences between frequencies for positive and neutral words were not accompanied by differences in sensitivities. In Experiment 1, meanwhile, word frequencies estimated via the log correction were higher for positive than negative words and sensitivity was also higher. This outcome suggests that, if anything, the magnitude of the sensitivity advantage due to valence is an underestimate. Thus, overall, while it is not possible to rule out the potential influence of small differences between frequencies on some outcomes, there is no strong indication from the data sets overall that it has done so.

## Appendix E

## Experiment 2 – Example of the Old/New Recognition Test

An example of what the screen display looked like in the test phase, when participants were required to respond using a 6-point confidence rating scale.

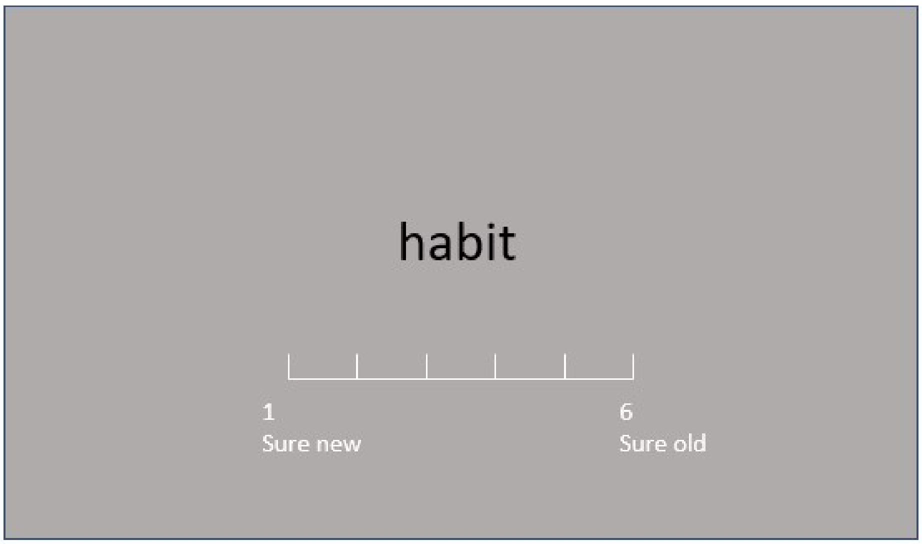



*Figure E1.* One trial in the test phase showing a word presented on the screen with the 6-point confidence rating scale.

## Appendix F

### Experiment 2 – Proportion of Responses on the Confidence Scale

Figure F1 illustrates the proportions of responses for each confidence levels for old and new words for each valence type (negative, positive and neutral) in each study. The data are collapsed across the TBR/TBF dimension and the figures show that the proportions of old responses to old words are greater than the proportions of new responses to old words. The data in the right-hand columns show that the reverse is true for new words, with a markedly greater proportion of new words attracting the two classes of 'new' judgements.

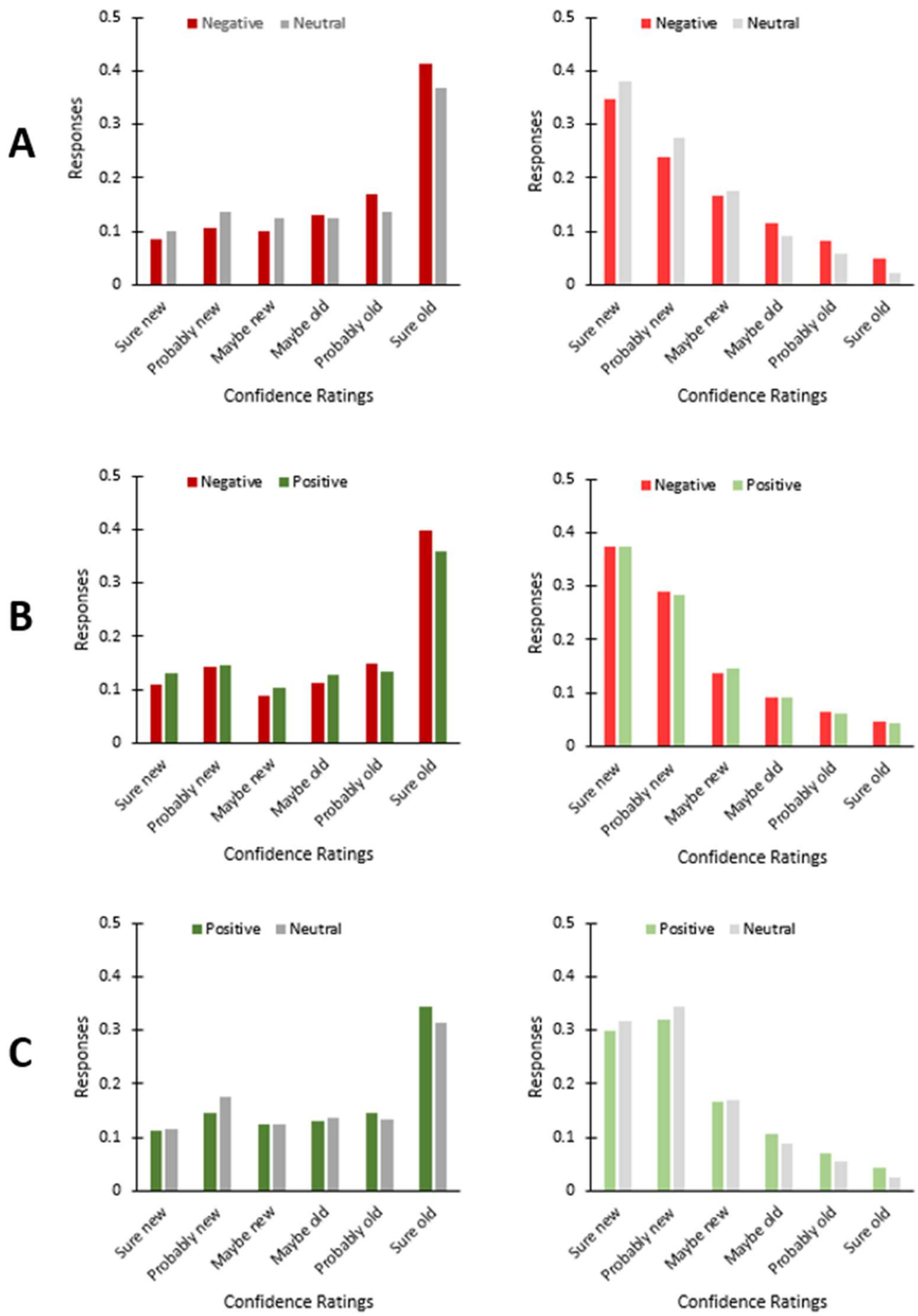



*Figure F1.* Response rates (proportions) for each confidence rating for old and new items across valence (negative, positive and neutral) in Experiment 2 for study 1 (A: negative vs. neutral), study 2 (B: negative vs. positive) and study 3 (C: positive vs. neutral). The left column illustrates responses to old words and the right column illustrates responses to new items.

## Appendix G

## Experiment 2: ANOVA results on $d_a$ analyses

Table G1

*Summary of Repeated Measures ANOVA Results for Comparisons between Instruction and Valence in Studies 1, 2 and 3*

| Measure (DV) | Study 1 | | | | Study 2 | | | | Study 3 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | *df* | *F* | *p* | $\eta_p^2$ | *df* | *F* | *p* | $\eta_p^2$ | *df* | *F* | *p* | $\eta_p^2$ |
| Instruction | 1, 43 | 4.89 | .032* | .10 | 1, 43 | 23.43 | <.000*** | .35 | 1, 43 | 16.39 | <.000*** | .28 |
| Valence | 1, 43 | 1.49 | .230 | .03 | 1, 43 | 7.80 | .008** | .15 | 1, 43 | 0.18 | .676 | .00 |
| Instruction x Valence Interactions | 1, 43 | 0.51 | .479 | .01 | 1, 43 | 0.08 | .775 | .00 | 1, 43 | 2.25 | .141 | .05 |

*Notes.* * $p < .05$, ** $p < .01$, *** $p < .001$.
df = degrees of freedom

## Appendix H

## z-Transformed ROC curves for Experiment 2 and 3

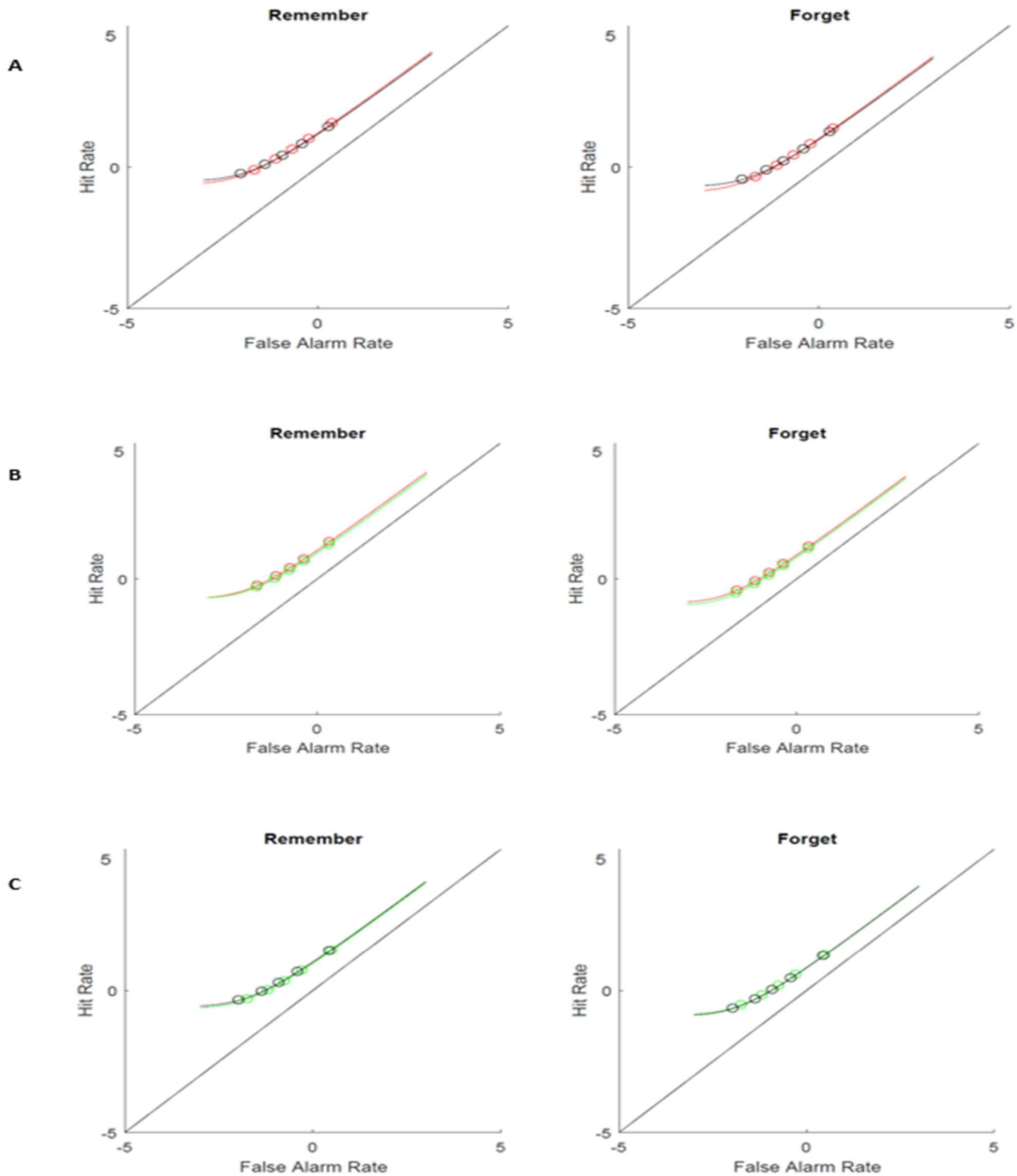



*Figure H1.* z-Transformed Receiver operating characteristic (zROC) data for study 1 (A), study 2 (B) and study 3 (C) in Experiment 2. The red curves denote the negative condition, the green curves denote the positive condition and the black curves denote the neutral condition. The curves indicate memory strength and the circles indicate the level of response confidence (response criterion).

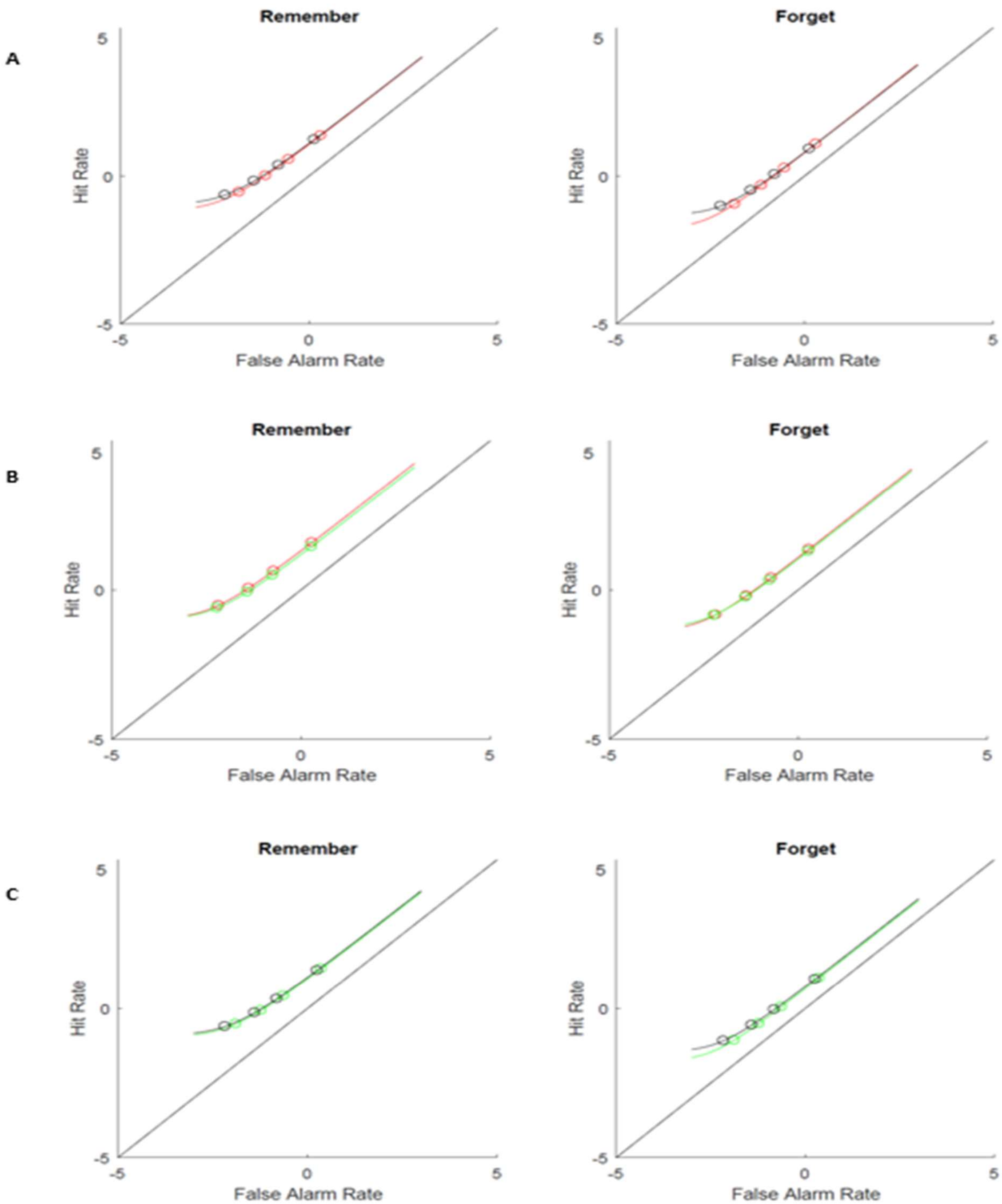



*Figure H2.* z-Transformed Receiver operating characteristic (zROC) data for study 1 (A), study 2 (B) and study 3 (C) in Experiment 3. The red curves denote the negative condition, the green curves denote the positive condition and the black curves denote the neutral condition. The curves indicate memory strength and the circles indicate the level of response confidence (response criterion).

## Appendix I

### Experiment 3 – Proportion of Responses on the Confidence Scale

Figure H1 illustrates the proportions of responses for each confidence levels for old and new words for each valence type (negative, positive and neutral) in each study. The data are collapsed across the TBR/TBF dimension and the figures show that the proportions of old (Remember or Know (sure old and maybe old)) responses to old words are greater than the proportions of new responses to old words. The data in the right-hand columns show that the reverse is true for new words, with a markedly greater proportion of new words attracting the two classes of 'new' judgements.

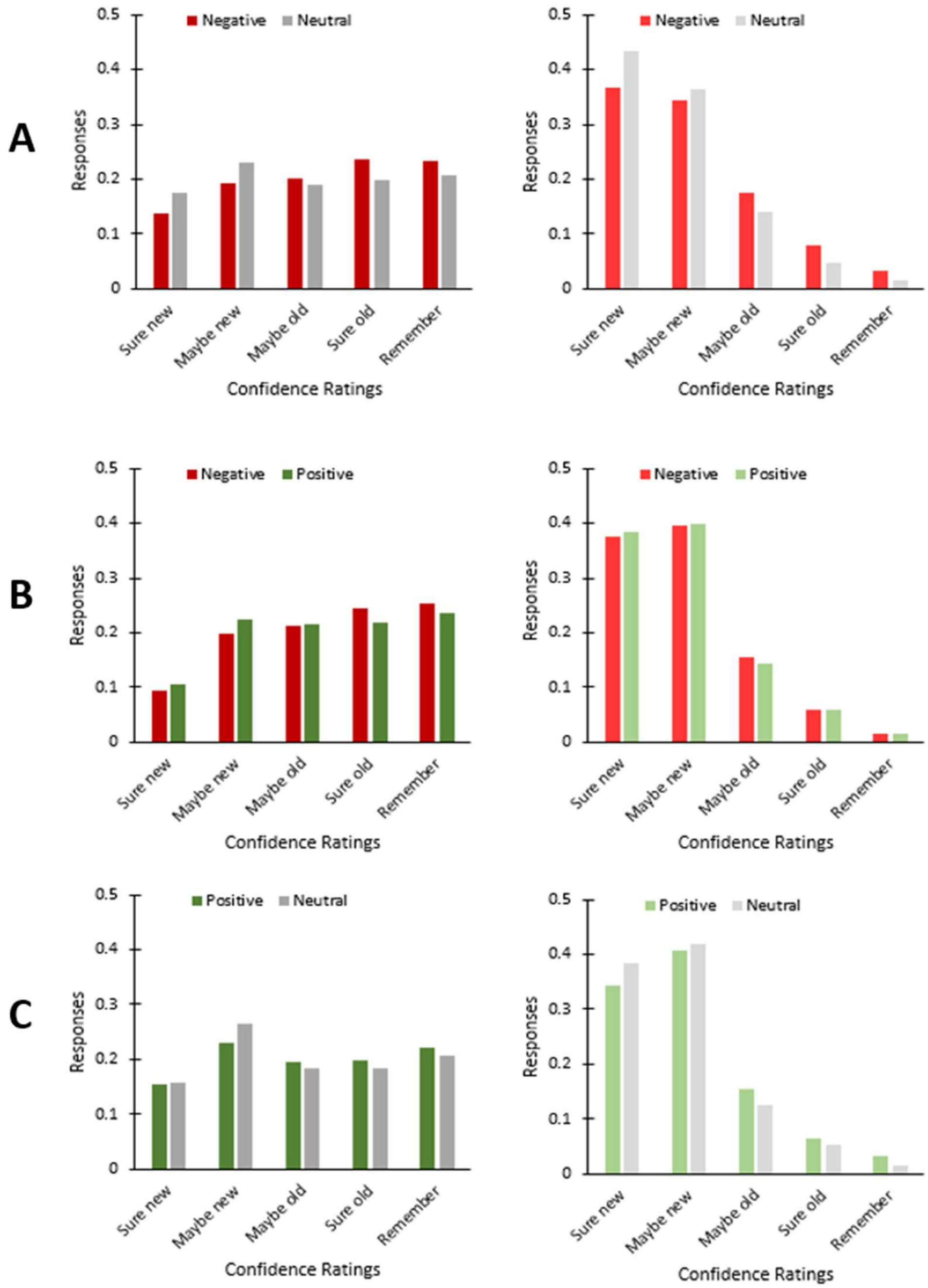



*Figure I1*. Response rates (proportions) for each response possibility (confidence ratings and remember response) for old and new items across valence (negative, positive and neutral) in Experiment 3 for study 1 (A: negative vs. neutral), study 2 (B: negative vs. positive) and study 3 (C: positive vs. neutral). The left column represents responses to old items and the right column represent responses to new items

## Appendix J

### Experiment 3 – Instructions Provided during the R/K Test Phase

Below are the instructions that were provided to the participants on how to respond and how to use the confidence scale appropriately in Experiment 3.

**Instructions Test phase:**
A test phase follows each study phase. In each test phase you will again be shown words one at a time on the screen. Some of these ('old' words) will have been shown in the study phase. Others ('new' words) will be shown at test for the first time in the experiment.
For each word please make one of five response options. The sheet beside you shows you which key corresponds to which response option.
1= Sure new
2= Maybe new
3= Maybe old
4= Sure old
5= Remember

You should make a Remember response if you can recall specific information about when you encountered the word in the preceding study phase. For example, you might remember where it was in the study list, or what you thought at the time you saw it. For some words you might believe that you saw them at study but not be able to remember specific details. If this is the case, then please make either a 'Sure old' or 'Maybe old' response, depending upon how confident you are that you studied the word. Similarly, if you see a word at test that you think you did not study, please make either a "Sure new' or 'Maybe new' judgment, depending upon how confident you are.

## Appendix K

## Experiment 3 – ROC Figures based on a 5-point Rating Scale

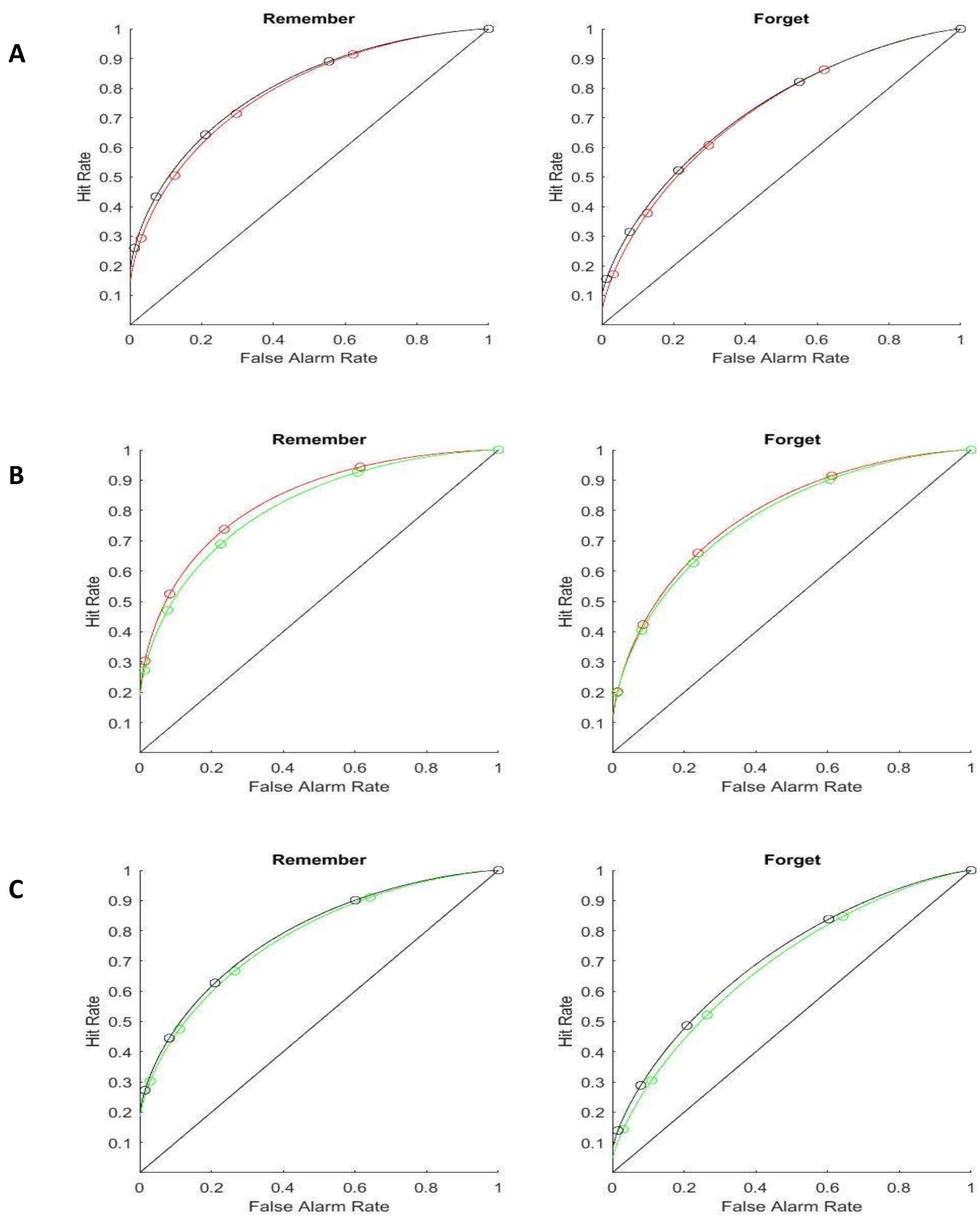



*Figure K1.* Receiver operating characteristic (ROC) data for study 1 (A), study 2 (B) and study 3 (C) in Experiment 3. The red curves denote the negative condition, the green curves denote the positive condition and the black curves denote the neutral condition. The curves indicate memory strength and the circles indicate the level of response confidence (response criterion).

## Appendix L

## Experiment 4 – Stimuli Analyses

Table L1

*Outcomes of Independent Sample T-tests for Assessments on Equality for Valence, Absolute Valence, Arousal, Word Length, Word Frequency and Semantic Relatedness for the Lists of 240 Words (for Each Valence) between Valence Type (Negative, Positive and Neutral) for Studies 1 and 2 in Experiment 4. SD = standard deviation*

| | Study 1 | | | Study 2 | | |
|---|---|---|---|---|---|---|
| | Negative | Neutral | | Positive | Neutral | |
| | M (SD) | M (SD) | t-value (*p*) | M (SD) | M (SD) | t-value (*p*) |
| Valence | 2.70 (0.67) | 5.12 (0.38) | -48.52 (<.001)*** | 7.28 (0.48) | 5.12 (0.38) | 54.96 (<.001)*** |
| Absolute Valence | 2.30 (0.67) | 0.27 (0.29) | 42.96 (<.001)*** | 2.28 (0.48) | 0.27 (0.29) | 55.78 (<.001)*** |
| Arousal | 4.85 (0.97) | 4.11 (0.76) | 9.27 (<.001)*** | 4.58 (0.94) | 4.11 (0.76) | 5.94 (<.001)*** |
| Word length | 7.42 (2.17) | 7.16 (2.17) | 1.31 (.193) | 7.51 (2.11) | 7.16 (2.17) | 1.77 (.077) |
| Word frequency | 2.69 (0.67) | 2.77 (0.81) | -1.29 (.198) | 2.83 (0.71) | 2.77 (0.81) | 0.78 (.436) |
| Semantic relatedness | 0.90 (0.02) | 0.90 (0.02) | -1.62 (.107) | 0.89 (0.03) | 0.90 (0.02) | -1.72 (.087) |

*Notes.* *** $p < .001$.

Table L2

*Outcomes of Independent Sample T-tests for Assessments on Equality for Arousal, Word Length, Word Frequency and Semantic Relatedness for the Study Lists (for Each Valence) between Valence Type (Negative and Neutral, and Negative and Positive) for Studies 1 and 2 in Experiment 4. SD = standard deviation*

| | | Study 1 | | | Study 2 | | |
|---|---|---|---|---|---|---|---|
| | | Negative | Neutral | | Positive | Neutral | |
| | | M (SD) | M (SD) | t-value (*p*) | M (SD) | M (SD) | t-value (*p*) |
| A1 | Arousal | 4.58 (0.94) | 3.92 (0.78) | 2.45 (.019)* | 4.63 (1.03) | 4.14 (0.98) | 1.55 (.131) |
| | Word length | 7.60 (2.01) | 6.95 (2.50) | 0.91 (.371) | 7.20 (2.04) | 6.55 (1.73) | 1.09 (.284) |
| | Word frequency | 2.60 (0.54) | 2.72 (0.92) | -0.49 (.629) | 3.11 (0.48) | 2.94 (1.02) | 0.70 (.489) |
| | Semantic relatedness | 0.88 (0.04) | 0.88 (0.02) | -0.24 (.809) | 0.90 (0.02) | 0.90 (0.03) | 0.58 (.569) |
| A2 | Arousal | 4.78 (1.12) | 4.14 (0.61) | 2.24 (.031)* | 4.42 (1.06) | 4.13 (0.89) | 0.92 (.365) |
| | Word length | 6.90 (2.20) | 6.90 (1.52) | 0.00 (1.000) | 7.45 (2.63) | 8.00 (2.36) | -0.70 (.490) |
| | Word frequency | 2.87 (0.89) | 2.84 (0.81) | 0.11 (.916) | 2.86 (0.77) | 2.73 (0.86) | 0.50 (.617) |
| | Semantic relatedness | 0.88 (0.03) | 0.89 (0.03) | -0.58 (.566) | 0.90 (0.03) | 0.89 (0.03) | 1.31 (.197) |
| B1 | Arousal | 4.92 (0.81) | 3.91 (0.52) | 4.70 (<.001)*** | 4.64 (0.90) | 4.11 (0.83) | 1.95 (.059) |
| | Word length | 7.15 (1.76) | 7.00 (2.68) | 0.21 (.835) | 7.80 (2.29) | 7.40 (2.74) | 0.50 (.619) |
| | Word frequency | 2.84 (0.72) | 2.90 (0.96) | -0.21 (.827) | 2.74 (0.64) | 2.93 (0.75) | -0.84 (.408) |
| | Semantic relatedness | 0.89 (0.03) | 0.90 (0.02) | -1.55 (.130) | 0.91 (0.03) | 0.91 (0.02) | 0.27 (.791) |
| B2 | Arousal | 4.94 (0.79) | 4.29 (0.89) | 2.47 (.018)* | 4.34 (0.98) | 4.23 (0.68) | 0.40 (.690) |
| | Word length | 6.80 (2.21) | 7.25 (1.37) | -0.77 (.445) | 7.50 (2.19) | 7.30 (2.92) | 0.25 (.808) |
| | Word frequency | 2.84 (0.72) | 2.98 (0.72) | -0.61 (.548) | 2.83 (0.83) | 2.50 (0.52) | 1.52 (.137) |
| | Semantic relatedness | 0.89 (0.02) | 0.90 (0.03) | -0.83 (.415) | 0.89 (0.03) | 0.90 (0.03) | -1.38 (.177) |
| C1 | Arousal | 4.91 (1.11) | 4.23 (0.74) | 2.25 (.030)* | 4.68 (1.07) | 4.10 (0.75) | 1.98 (.055) |
| | Word length | 7.00 (2.13) | 6.45 (1.61) | 0.92 (.362) | 7.85 (2.39) | 6.65 (2.03) | 1.71 (.095) |
| | Word frequency | 2.75 (0.61) | 2.80 (0.62) | -0.25 (.804) | 3.09 (0.70) | 2.87 (0.61) | 1.04 (.304) |
| | Semantic relatedness | 0.90 (0.03) | 0.90 (0.02) | -0.14 (.893) | 0.89 (0.03) | 0.90 (0.03) | -1.01 (.320) |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| C2 | Arousal | 4.79 (1.12) | 4.10 (1.01) | 2.05 (.047)* | 4.57 (0.93) | 4.12 (0.76) | 1.70 (.097) |
| | Word length | 8.60 (2.64) | 7.45 (2.06) | 1.53 (.133) | 7.85 (2.32) | 7.45 (2.46) | 0.53 (.600) |
| | Word frequency | 2.44 (0.67) | 2.69 (0.61) | -1.23 (.228) | 2.53 (0.62) | 2.77 (0.84) | -1.02 (.313) |
| | Semantic relatedness | 0.91 (0.03) | 0.90 (0.02) | 0.81 (.421) | 0.91 (0.03) | 0.91 (0.02) | -0.08 (.940) |
| D1 | Arousal | 4.73 (0.82) | 4.11 (0.71) | 2.56 (.014)* | 4.53 (0.80) | 3.99 (0.69) | 2.28 (.028)* |
| | Word length | 7.85 (2.37) | 8.45 (1.99) | -0.87 (.391) | 7.55 (1.67) | 6.70 (1.66) | 1.62 (.114) |
| | Word frequency | 2.79 (0.55) | 2.34 (0.84) | 1.99 (.054) | 2.67 (0.77) | 2.62 (0.83) | 0.19 (.852) |
| | Semantic relatedness | 0.91 (0.03) | 0.92 (0.02) | -1.88 (.067) | 0.87 (0.03) | 0.88 (0.04) | -0.51 (.613) |
| D2 | Arousal | 4.84 (0.95) | 4.27 (0.72) | 2.16 (.037)* | 4.58 (0.78) | 4.12 (0.56) | 2.13 (.040)* |
| | Word length | 7.75 (2.24) | 7.00 (1.75) | 1.18 (.246) | 7.30 (2.32) | 7.70 (2.39) | -0.54 (.594) |
| | Word frequency | 2.65 (0.66) | 2.81 (0.79) | -0.71 (.483) | 2.78 (0.77) | 2.48 (0.99) | 1.08 (.289) |
| | Semantic relatedness | 0.89 (0.03) | 0.90 (0.02) | -1.62 (.114) | 0.90 (0.03) | 0.90 (0.02) | -0.39 (.699) |
| E1 | Arousal | 5.01 (0.89) | 4.19 (0.75) | 3.14 (.003)** | 4.80 (0.89) | 4.18 (0.64) | 2.55 (.015)* |
| | Word length | 6.95 (2.39) | 6.40 (2.28) | 0.74 (.462) | 6.60 (1.54) | 7.30 (1.78) | -1.33 (.191) |
| | Word frequency | 2.80 (0.73) | 2.73 (0.77) | 0.31 (.761) | 2.98 (0.57) | 2.98 (0.57) | 0.94 (.352) |
| | Semantic relatedness | 0.90 (0.02) | 0.89 (0.03) | 0.17 (.864) | 0.90 (0.03) | 0.90 (0.03) | -.24 (.812) |
| E2 | Arousal | 4.85 (1.14) | 4.12 (0.91) | 2.24 (.031)* | 4.61 (1.08) | 4.30 (0.62) | 1.10 (.280) |
| | Word length | 7.15 (2.08) | 7.55 (2.63) | -0.53 (.597) | 7.70 (1.66) | 6.70 (1.66) | 1.65 (.108) |
| | Word frequency | 2.52 (0.60) | 2.81 (0.84) | -1.28 (.207) | 2.96 (0.69) | 2.89 (0.74) | 0.30 (.765) |
| | Semantic relatedness | 0.89 (0.02) | 0.88 (0.04) | 0.83 (.411) | 0.89 (0.03) | 0.89 (0.02) | 0.77 (.447) |
| F1 | Arousal | 5.02 (0.82) | 3.94 (0.66) | 4.59 (<.001)*** | 4.71 (0.94) | 4.19 (0.81) | 1.87 (.069) |
| | Word length | 7.45 (1.79) | 7.35 (2.46) | 0.15 (.884) | 7.75 (1.83) | 6.75 (1.92) | 1.69 (.100) |
| | Word frequency | 2.70 (0.68) | 2.62 (0.91) | 0.28 (.779) | 2.60 (0.73) | 2.96 (0.68) | - 1.64 (.109) |
| | Semantic relatedness | 0.89 (0.02) | 0.90 (0.02) | -1.87 (.070) | 0.88 (0.03) | 0.88 (0.02) | 0.32 (.750) |
| F2 | Arousal | 4.86 (1.18) | 4.16 (0.84) | 2.17 (.037)* | 4.44 (0.90) | 3.85 (0.89) | 2.10 (.042)* |
| | Word length | 7.85 (1.90) | 7.20 (2.50) | 0.93 (.361) | 7.55 (1.91) | 7.25 (1.65) | 0.53 (.598) |
| | Word frequency | 2.41 (0.53) | 3.02 (0.80) | -2.80 (.008)** | 2.76 (0.75) | 2.73 (0.89) | 0.10 (.919) |
| | Semantic relatedness | 0.90 (0.03) | 0.91 (0.03) | -0.48 (.633) | 0.90 (0.03) | 0.91 (0.03) | -1.91 (.064) |

*Notes.* * $p < .05$, ** $p < .01$, *** $p < .001$.

Table L3

*Outcomes of Independent Sample T-tests for Assessment on Equality for Arousal, Word Length, Word Frequency and Semantic Relatedness for the Test Lists (for Each Valence) between Valence Type (Negative and Neutral, and Negative and Positive) for Studies 1 and 2 in Experiment 4. SD = standard deviation*

| | | Study 1 | | | Study 2 | | |
|---|---|---|---|---|---|---|---|
| | | Negative | Neutral | | Positive | Neutral | |
| | | M (SD) | M (SD) | t-value (*p*) | M (SD) | M (SD) | t-value (*p*) |
| A | Arousal | 4.68 (1.02) | 4.03 (0.70) | 3.33 (.001)** | 4.52 (1.03) | 4.14 (0.93) | 1.76 (.082) |
| | Word length | 7.25 (2.11) | 6.93 (2.04) | 0.70 (.486) | 7.33 (2.33) | 7.28 (2.17) | 0.10 (.921) |
| | Word frequency | 2.73 (0.74) | 2.78 (0.86) | -0.24 (.808) | 2.99 (0.65) | 2.83 (0.94) | 0.85 (.398) |
| | Semantic relatedness | 0.88 (0.03) | 0.89 (0.02) | -0.23 (.820) | 0.90 (0.03) | 0.89 (0.03) | 0.48 (.633) |
| B | Arousal | 4.93 (0.79) | 4.10 (0.74) | 4.86 (<.001)*** | 4.49 (0.94) | 4.17 (0.75) | 1.68 (.097) |
| | Word length | 6.98 (1.98) | 7.13 (2.10) | -0.33 (.743) | 7.65 (2.21) | 7.35 (2.80) | 0.53 (.596) |
| | Word frequency | 2.84 (0.71) | 2.94 (0.84) | -0.57 (.572) | 2.79 (0.73) | 2.71 (0.67) | 0.47 (.639) |
| | Semantic relatedness | 0.89 (0.02) | 0.90 (0.03) | -1.66 (.101) | 0.90 (0.03) | 0.91 (0.02) | -1.89 (.063) |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| C | Arousal | 4.85 (1.10) | 4.17 (0.88) | 3.06 (.003)** | 4.63 (0.99) | 4.11 (0.75) | 2.64 (.010)** |
| | Word length | 7.80 (2.50) | 6.95 (1.89) | 1.71 (.091) | 7.85 (2.32) | 7.05 (2.26) | 1.56 (.123) |
| | Word frequency | 2.60 (0.65) | 2.75 (0.61) | -1.05 (.250) | 2.81 (0.71) | 2.82 (0.73) | -0.07 (.943) |
| | Semantic relatedness | 0.91 (0.03) | 0.90 (0.02) | 0.76 (.449) | 0.90 (0.03) | 0.90 (0.02) | -1.44 (.153) |
| D | Arousal | 4.78 (0.88) | 4.19 (0.71) | 3.35 (.001)** | 4.56 (0.78) | 4.02 (0.66) | 3.29 (.002)** |
| | Word length | 7.80 (2.29) | 7.73 (1.99) | 0.16 (.876) | 7.43 (2.00) | 7.30 (2.21) | 0.27 (.791) |
| | Word frequency | 2.72 (0.61) | 2.58 (0.84) | 0.85 (.386) | 2.72 (0.76) | 2.58 (0.93) | 0.74 (.462) |
| | Semantic relatedness | 0.90 (0.02) | 0.91 (0.02) | -1.64 (.105) | 0.89 (0.03) | 0.90 (0.03) | -1.62 (.110) |
| E | Arousal | 4.93 (1.01) | 4.16 (0.82) | 3.75 (<.001)*** | 4.70 (0.98) | 4.24 (0.63) | 2.53 (.013)* |
| | Word length | 7.05 (2.22) | 6.98 (2.50) | 0.14 (.887) | 7.15 (1.93) | 7.00 (1.73) | 0.37 (.715) |
| | Word frequency | 2.66 (0.67) | 2.77 (0.80) | -0.68 (.502) | 2.97 (0.63) | 2.84 (0.76) | 0.88 (.383) |
| | Semantic relatedness | 0.89 (0.02) | 0.89 (0.03) | 0.14 (.886) | 0.90 (0.03) | 0.89 (0.03) | 0.98 (.331) |
| F | Arousal | 4.94 (1.01) | 4.05 (0.75) | 4.48 (<.001)*** | 4.58 (0.92) | 4.02 (0.86) | 2.81 (.006)** |
| | Word length | 7.65 (1.83) | 7.28 (2.45) | 0.78 (.441) | 7.65 (1.85) | 7.00 (1.78) | 1.60 (.113) |
| | Word frequency | 2.56 (0.62) | 2.82 (0.87) | -1.56 (.122) | 2.68 (0.73) | 2.85 (0.79) | -1.00 (.323) |
| | Semantic relatedness | 0.90 (0.02) | 0.90 (0.03) | -1.08 (.282) | 0.89 (0.02) | 0.90 (0.02) | -1.82 (.072) |

*Notes.* * $p < .05$, ** $p < .01$, *** $p < .001$.

# Appendix M

## Experiment 4 – Effects of Valence on Dm Effects

Based on previous reports of analyses of subsequent memory effects and emotion in experiments involving a recognition memory test (Yick et al., 2015), mean amplitudes were targeted for two time windows: 200-400ms and 400-1000ms. Data from 30 electrode locations were grouped in 5 scalp regions: Frontal (F5, F3, F1, F2, F4, F6), Fronto-central (FC5, FC3, FC1, FC2, FC4, FC6), Central (C5, C3, C1, C2, C4, C6), Centro-parietal (CP5, CP3, CP1, CP2, CP4, CP6) and Parietal (P5, P3, P1, P2, P4, P6). For each time window repeated-measure ANOVAs were computed. They included the following factors: Response (correct vs incorrect), Valence (negative vs neutral or positive vs neutral), Anterior-Posterior (Frontal vs Fronto–Central vs Central vs Centro–parietal vs Parietal), Hemisphere (left vs right) and Site (Inferior vs Mid-lateral vs Superior).

### Results

Grand average waveform data are displayed below for the following electrodes: Fz, FCz, Cz, CPz and Pz for each valence condition, and separated by accuracy (response category). The scalp maps at the foot of each figure are computed on the basis of difference scores obtained by subtracting the mean amplitude measures for incorrect responses from those for the correct responses.

#### Effect of Valence on Dm Effects

In study 1 (see Figure L1) and in the replication of study 1 (see Figure L2) there were no differences according to valence. In study 2 (see Figure L3), however, there are marked differences for both positive and neutral words. The figure shows a greater relative positivity for incorrect responses than correct responses which onsets at around 250ms post-stimulus and continues for the duration of the recording epoch.

## Study 1

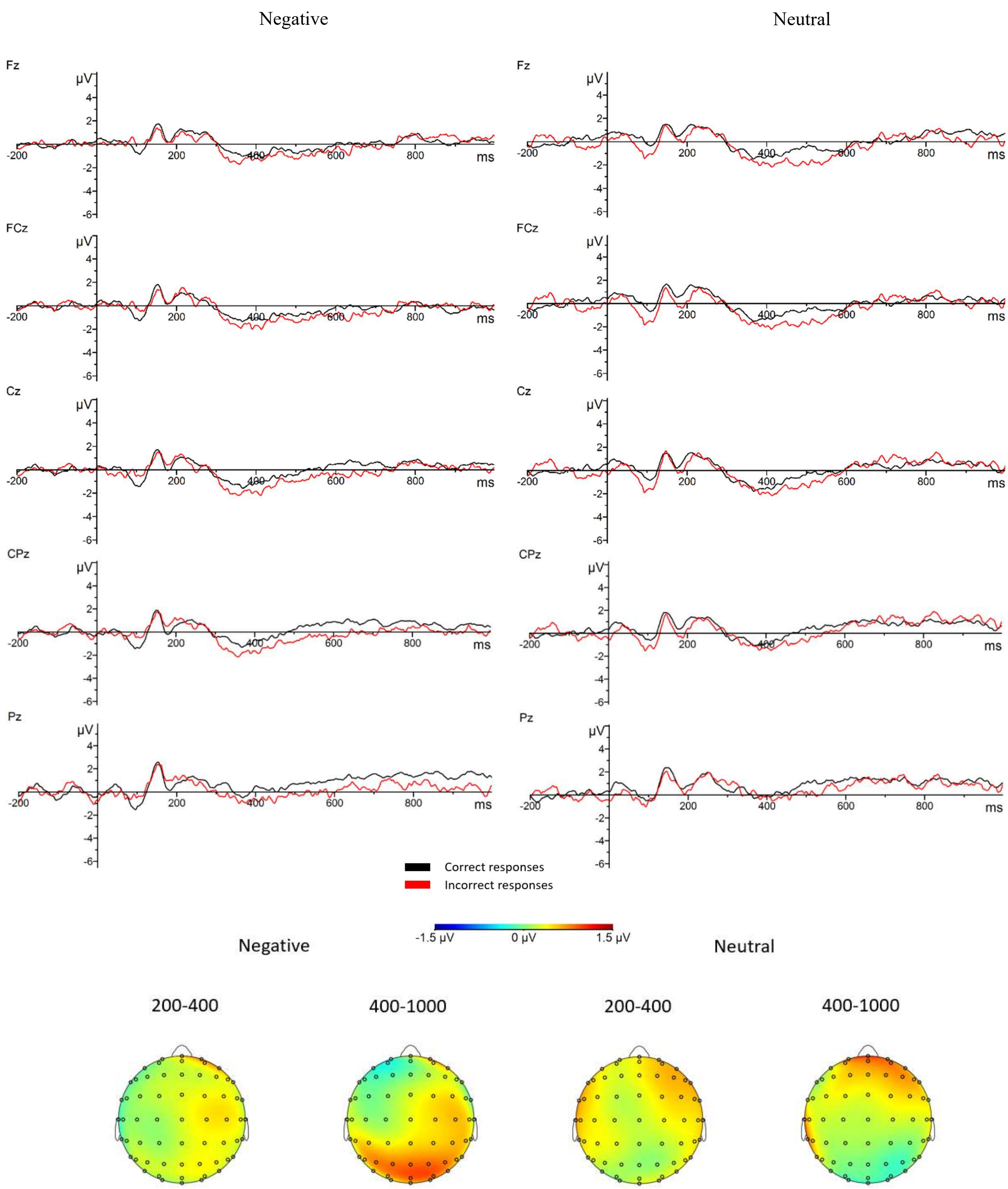



*Figure M1.* (A) Grand average ERPs for channels Fz, FCz, Cz, CPz and Pz for negative (left column) and neutral (right column) words, separated for correct and incorrect responses. (B) Scalp maps showing the differences between the neural activities elicited by correct and incorrect responses in each targeted condition: negative and neutral words for the time windows 200-400ms and 400-1000ms.

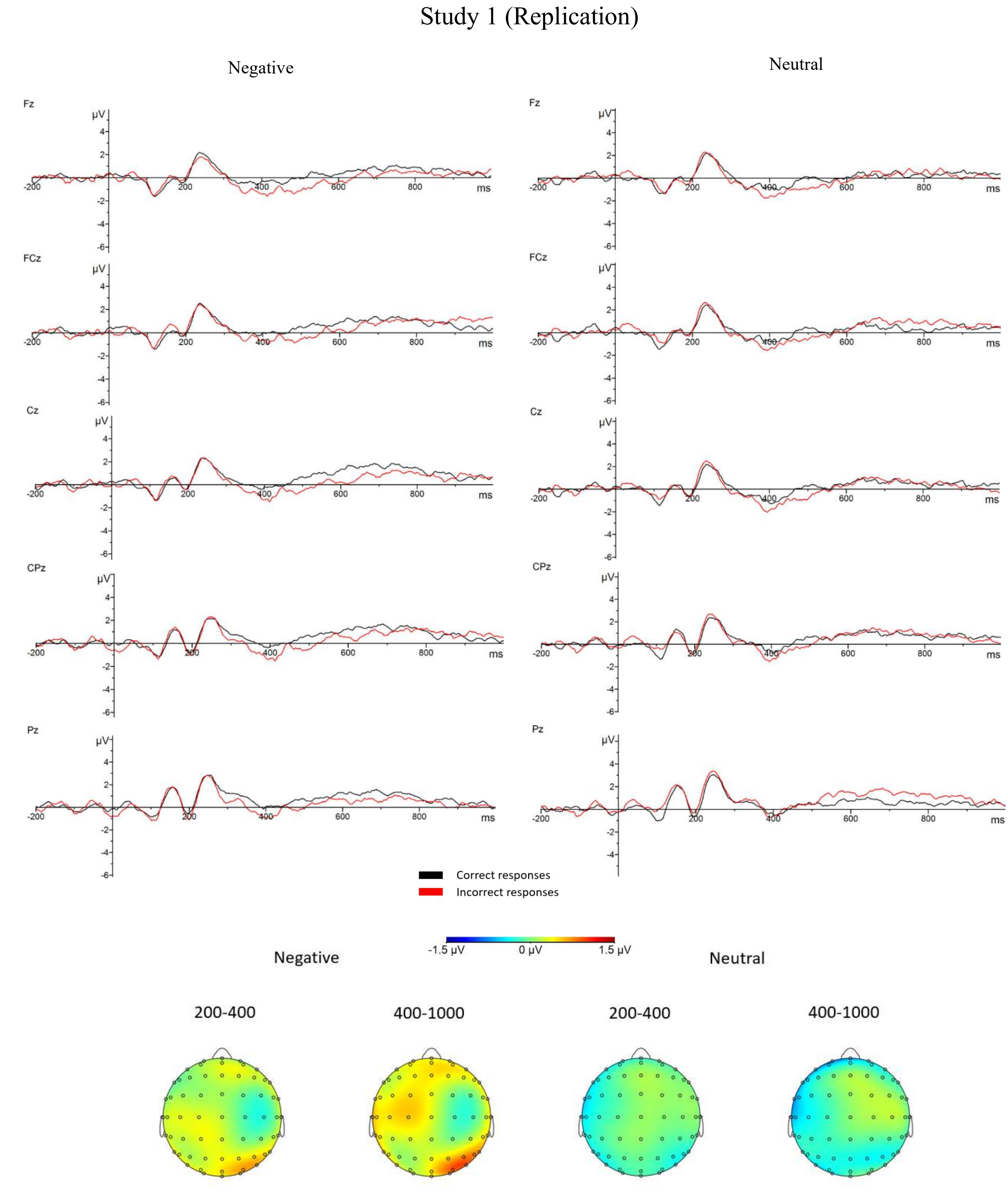



*Figure M2.* (A) Grand average ERPs for channels Fz, FCz, Cz, CPz and Pz for negative (left column) and neutral (right column) words, separated for correct and incorrect responses. (B) Scalp maps showing the differences between the neural activities elicited by correct and incorrect responses in each targeted condition: negative and neutral words for the time windows 200-400ms and 400-1000ms.

## Study 2

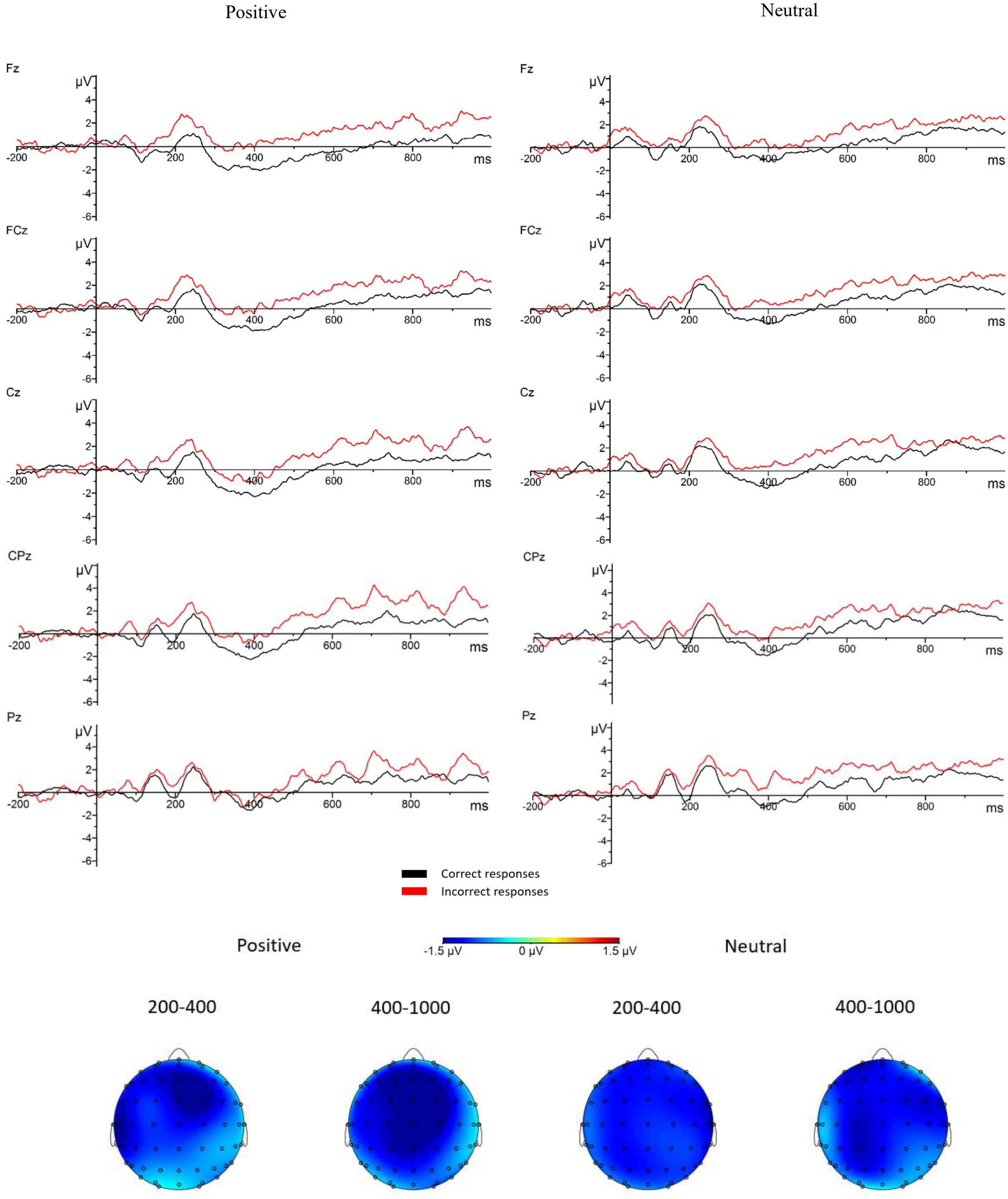



*Figure M3.* (A) Grand average ERPs for channels Fz, FCz, Cz, CPz and Pz for positive (left column) and neutral (right column) words, separated for correct and incorrect responses. (B) Scalp maps showing the differences between the neural activities elicited by correct and incorrect responses in each targeted condition: positive and neutral words for the time windows 200-400ms and 400-1000ms.

There were no interactions between valence and response category in all studies in either time windows (see Table L1 and L2). In study 2, there were main effects for response category in both time windows, due to a greater relative positivity for incorrect responses relative to correct responses with an onset of around 200ms and extended for the remainder of the recording epoch.

Table M1

*Summary of Repeated Measures ANOVA Results for Comparisons between Valence (Negative, Positive and Neutral) and Response (Correct and Incorrect Responses) in the 200-400ms time window for Study 1, Replication Study 1 and Study 2*

| | Study 1 | | | | Study 1 (Replication) | | | | Study 2 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Measure (DV) | *df* | *F* | *p* | $\eta_p^2$ | *df* | *F* | *p* | $\eta_p^2$ | *df* | *F* | *p* | $\eta_p^2$ |
| Valence | 1,15 | 0.02 | .898 | .00 | 1,15 | 0.01 | .928 | .00 | 1,15 | 3.27 | .091 | .18 |
| Response | 1,15 | 1.19 | .293 | .07 | 1,15 | 0.00 | .976 | .00 | 1,15 | 13.53 | .002** | .47 |
| Valence*Response | 1,15 | 0.90 | .765 | .01 | 1,15 | 0.54 | .476 | .03 | 1,15 | 0.02 | .896 | .00 |

*Note.* ** $p < .01$.

df = degrees of freedom

Table M2

*Summary of Repeated Measures ANOVA Results for Comparisons between Valence (Negative, Positive and Neutral) and Response (Correct and Incorrect Responses) in the 400-1000ms time window for Study 1, Replication Study 1 and Study 2*

| | Study 1 | | | | Study 1 (Replication) | | | | Study 2 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Measure (DV) | *df* | *F* | *p* | $\eta_p^2$ | *df* | *F* | *p* | $\eta_p^2$ | *df* | *F* | *p* | $\eta_p^2$ |
| Valence | 1,15 | 0.00 | .968 | .00 | 1,15 | 0.02 | .905 | .00 | 1,15 | 1.79 | .201 | .11 |
| Response | 1,15 | 0.32 | .580 | .02 | 1,15 | 0.05 | .834 | .00 | 1,15 | 28.60 | .000*** | .66 |
| Valence*Response | 1,15 | 0.01 | .920 | .00 | 1,15 | 0.71 | .414 | .05 | 1,15 | 0.01 | .946 | .00 |

*Note.* *** $p < .001$.

df = degrees of freedom

# Appendix N

## Experiment 5 – Stimulus Analyses of the Images

Table N1

*Outcomes of Independent Sample T-tests for Assessments on Equality for Arousal, Absolute Valence, Edge Density and Feature Congestion for the Images in the Study Phase between Valence Type (Negative, Positive and Neutral) in Experiment 5. SD = standard deviation*

| | Negative vs. Neutral | | | Negative vs. Positive | | | Positive vs. Neutral | | |
|---|---|---|---|---|---|---|---|---|---|
| | Negative | Neutral | | Negative | Positive | | Positive | Neutral | |
| | M (SD) | M (SD) | t-value (*p*) | M (SD) | M (SD) | t-value (*p*) | M (SD) | M (SD) | t-value (*p*) |
| Arousal | 1.09 (0.31) | -0.03 (0.36) | 18.31 (<.001)*** | 1.09 (0.31) | 1.03 (0.40) | 0.92 (.361) | 1.03 (0.40) | -0.03 (0.36) | 15.29 (<.001)*** |
| Absolute Valence | 1.27 (0.36) | 0.22 (0.14) | 20.94 (<.001)*** | 1.27 (0.36) | 1.22 (0.31) | 0.97 (.332) | 1.22 (0.31) | 0.22 (0.14) | 25.97 (<.001)*** |
| Edge Density | 0.04 (0.03) | 0.03 (0.02) | 1.13 (.261) | 0.04 (0.03) | 0.04 (0.02) | 0.48 (.633) | 0.04 (0.02) | 0.03 (0.02) | 0.82 (.413) |
| Feature Congestion | 3.28 (1.38) | 3.29 (1.15) | -0.01 (.990) | 3.28 (1.38) | 3.56 (1.12) | -1.20 (.231) | 3.56 (1.12) | 3.29 (1.15) | 1.32 (.190) |

*Notes*. *** p < .001.