Kent Academic Repository

Satish, Akul (2021) ***An investigation of our ability to control unwanted autobiograph***
memories of past morally relevant actions: EEG and behavioural evidence.
Doctor of Philosophy (PhD) thesis, University of Kent,.

Downloaded from

https://kar.kent.ac.uk/91156/ The University of Kent's Academic Repository KAR

The version of record is available from

https://doi.org/10.22024/UniKent/01.02.91156

This document version

UNSPECIFIED

DOI for this version

Licence for this version

CC BY-NC (Attribution-NonCommercial)

Additional information

Versions of research works

Versions of Record

If this version is the version of record, it is the same as the published version available on the publisher's web site. Cite as the published version.

Author Accepted Manuscripts

If this document is identified as the Author Accepted Manuscript it is the version after peer review but before type setting, copy editing or publisher branding. Cite as Surname, Initial. (Year) 'Title of article'. To be published in *Title of Journal* , Volume and issue numbers [peer-reviewed accepted version]. Available at: DOI or URL (Accessed: date).

Enquiries

If you have questions about this document contact ResearchSupport@kent.ac.uk. Please include the URL of the record in KAR. If you believe that your, or a third party's rights have been compromised through this document please see our Take Down policy (available from https://www.kent.ac.uk/guides/kar-the-kent-academic-repository#policies).

An investigation of our ability to control unwanted autobiographical memories of past morally relevant actions: EEG and behavioural evidence

Akul Satish

A thesis submitted for the degree of Doctor of Philosophy (Ph.D.)

in Cognitive Psychology/Neuropsychology

School of Psychology

University of Kent, Canterbury

July 2021

Declaration

I, Akul Satish, declare that the work presented in this thesis is my own. The work presented is original completed under the supervision of Dr Zara Bergström. I have not been awarded a degree by submitting the work included in this thesis for higher degree at any other institution.

Acknowledgements

I would like to firstly thank my parents for their unconditional love, support, encouragement, guidance, and sacrifice. You both have invested time, money, and energy in me, believed in me, inspired me, and continually give me confidence to pursue my dreams. Without you both, this thesis would not have been possible and I am truly grateful.

My supervisor, Dr. Zara Bergström has been a source of inspiration, support, and guidance throughout my PhD. You have constantly encouraged my strengths and took extra effort to support me when I had my doubts and help me improve my weaknesses. You have been a wonderful mentor and provided me with new perspectives not only topics related to my thesis, but on memory and cognitive science in general. Thank you, it definitely has, and will continue to, help me become a better researcher.

I would also like to particularly thank Dr. Robin Hellerstedt who has been like a second supervisor to me throughout my PhD and has gone out of his way to help me improve several aspects of this project, and also provide personal support and guidance. Our conversations, especially during trips to conferences, has broadened my perspectives. A special thank you goes out to Dr. Joseph Brooks. You introduced me to the research world, inspired me to purse my goals, and provided a strong foundation in research that came in handy during my PhD. I would also like to thank Prof. Michael Anderson for his valuable feedback and inputs on the experimental design aspects of this project.

Thank you also to members of the Memory Lab, with special thanks to Dr. Louisa Salhi, Dr. Matthew Plummer, and Dr. Phot Dhamapeera who were readily helpful when I needed assistance with the practicalities of my PhD and have been wonderful peers and friends with whom I could share this journey.

A personal thank you goes to all my family members for their support during my PhD, with special thanks to my Ajji (grandmother) who keeps me grounded and humble and holds and shares a wealth of knowledge with me. A huge personal thank you also to my friends both at Kent and in India who are an endless source of laughter, fun, knowledge, support, arguments, and entertainment all bundled into one. You are all special to me and have definitely made this journey easier to navigate.

Abstract

We are often faced with reminders of unpleasant memories from our past that triggers a discomforted state of being. Attempts to avoid such reminders would be futile. Fortunately, a large body of research indicates that the unwanted memories and emotions elicited by these reminders can be controlled to improve our mental state. However, our understanding of this memory control is based on simpler memories such as words or pictures. So, the present thesis aimed to explore our ability to control memories of complex autobiographical memories, and its consequences on memory and emotion. Memories of past morally wrong actions were specifically observed as they are theorised to be particular motivators to engage control processes. In the first experiment (Chapter 3), evidence from real-time behavioural, ERP, and EEG oscillation measures indicated that autobiographical memories can be successfully avoided using direct suppression. Moreover, morally wrong memories were rated as more difficult to suppress than morally right memories. Importantly, repeated attempts at memory control improved the ability to reduce unwanted intrusions of memory contents (Chapter 3), and ameliorated the immediate negative emotions associated with morally wrong memories (Chapter 5). Both strategies of substituting negative thoughts with positive imagery, and upward counterfactual thinking, provided a distinct advantage over direct suppression for regulation of immediate emotions. The experiment reported in Chapter 4 found that suppression was not effective at beating the EEG-based forensic concealed-information test for either immoral and personal, or lab-created, autobiographical memories. The thesis thus provides new evidence that suppression may not be effective at avoiding incrimination from this forensic test for real-world emotional autobiographical memories. Overall, memory control can be useful for reducing intrusive thoughts and negative emotions associated with unpleasant autobiographical memories of past immoral actions. The thesis has implications for theories of memory control, emotion regulation and forensic practice.

Table of Contents

List of Figures

List of Tables

Chapter 1 – General Introduction & Literature Review

> ‘I sometimes find, and I am sure you know the feeling, that I simply have too many thoughts and memories crammed into my mind’… ‘At these times,’ said Dumbledore, indicating the stone basin, ‘I use the Pensieve. One simply siphons the excess thoughts from one’s mind, pours them into the basin, and examines them at one’s leisure…’ (Rowling, 2000, p. 597)

Perhaps if we lived in the wizarding world of Harry Potter, disturbing memories could be cast away with a flick of a wand. Alas, in our reality, the ability to cope with unwanted memories can be challenging, especially for events that are particularly intrusive and riddled with unpleasant emotions. Emerging evidence indicates that conscious control can be used to limit the unwelcome influences of such memories, and this memory control is a hallmark for good mental health (Engen & Anderson, 2018; Mary et al., 2020). Most of the extant memory control literature investigates memories of relatively simple stimuli such as words or pictures (Anderson & Hanslmayr, 2014). The current thesis aims to delineate the neurocognitive mechanisms involved in controlling personal autobiographical memories and the consequences of this control on the intrusiveness of, and emotions associated with, these memories.

In this chapter, a review of the extant understanding of neural and cognitive mechanisms of controlling memories is first discussed (e.g., Anderson et al., 2004; Benoit & Anderson, 2012; Levy & Anderson, 2012). The present research used electroencephalography (EEG) to understand the neurocognitive mechanisms underlying memory control, therefore results from EEG-based memory control studies are discussed in more detail. An emerging research area on memory control is its role in regulating negative emotions (Engen & Anderson, 2018), and this issue was also explored in the current studies.

Therefore, the literature pertaining this issue is reviewed after discussing the neurocognitive mechanisms of memory control.

The theory of motivated forgetting indicates that memories that are particularly intrusive and unpleasant are more susceptible to memory control (see Anderson & Hanslmayr, 2014). Autobiographical memories of our past immoral acts are strongly related to the sense of self, usually associated with feelings of guilt, remorse, and/or shame (Stanley, Henne, Iyengar, Sinnott-Armstrong, & Brigard, 2017). Therefore, it is reasonable to argue that such memories can potentially be strong *motivators* for engaging memory control processes. Research on the motivated forgetting of immoral actions has gained some interest in social psychology but not cognitive psychology research (Stanley & De Brigard, 2019). The present thesis therefore specifically focusses on our ability to control memories of past immoral actions using electrophysiological and behavioural methods.

Moreover, in the forensic science domain, EEG is used to detect concealed information in suspects of a crime. Some research has found that memory control strategies can be used to prevent such forensic detection of memories of a laboratory created mock crime memory (Bergström, Anderson, Buda, Simons, & Richardson-Klavehn, 2013; Hu, Bergström, Bodenhausen, & Rosenfeld, 2015), but it is not clear if this control can be successfully applied on personal memories that real-world criminals would be attempting to control. Therefore, another objective in the present thesis was to investigate if people can suppress memories of their personal immoral actions in the concealed-information paradigm. Thus, prior research that is relevant for memory control in the concealed information test is discussed before providing an overview of the rest of the thesis.

1.1 Mechanisms of memory control

When faced with reminders of disturbing events, not only the contents of these memories but also the associated negative emotions can automatically come to conscious experience and lead to an unpleasant state of being. A large body of research indicates that these reminders provide an opportunity to limit retrieval of the unwanted memory and consequently feel better. This ability to voluntarily engage cognitive control processes and reduce conscious experience of such memories is known as memory control (see Anderson & Hanslmayr, 2014; Engen & Anderson, 2018). As Engen and Anderson (2018) explain, engagement of memory control in response to a reminder is thought to be achieved through at least two distinct cognitive mechanisms: a) Direct memory suppression, involving purging contents of the memory from conscious awareness in an attempt to completely inhibit retrieval; and b) thought substitution, involving actively diverting attention to another thought, such as retrieving an alternate memory trace or thoughts in response to the reminder, consequently removing the original memory from conscious awareness. There is considerably more research on the mechanism of direct suppression compared to thought substitution, and the former is therefore reviewed in more detail throughout this chapter.

These memory control mechanisms are traditionally studied in the lab using the think/no-think paradigm (Anderson & Green, 2001; Anderson et al., 2004), which simulates the process of encountering reminders and controlling memories associated to these reminders in a laboratory setting. Participants typically first learn cue-target pairs such as a pair of weakly related words (e.g., ordeal-roach), wherein one stimulus (ordeal) acts as a reminder of the other stimulus (roach). Then they complete the think/no-think task. In each trial, a reminder (e.g., ordeal) from the learned pairs is presented in either red or green. Participants are instructed to actively remember and keep in mind the associated word (e.g., roach) if the reminder is displayed in green (think condition), but if it is shown in red (no-think condition), they are instructed to purge the associated word from consciousness (direct suppression) or retrieve an alternate memory trace or thought (thought substitution). Reminders are shown repeatedly in the same colour so that participants complete the task multiple times for each memory. Following the TNT task, participants complete a surprise memory test of all cue-target pairs, wherein the reminders (e.g., ordeal) are presented again and participants are asked to recall and report all associated words (e.g., roach in ordeal-roach word pair), ignoring previous instructions.

Recall performance in the final test is often found to be enhanced for think compared to no-think items, indicating that the benefit of repeated reminders of a memory is eliminated as a result of memory control. Importantly, when baseline items that are not shown in the think/no-think phase are also added, recall accuracy is lower for no-think than baseline memories (see Anderson & Levy, 2009). This reduced accuracy in memory for no-think items is often known as *suppression-induced forgetting* (Anderson & Hanslmayr, 2014). The suppression-induced forgetting effect is found for memories of words (Anderson et al., 2004), pictures (e.g., Benoit et al., 2015), aversive images (e.g., Gagnepain et al., 2017), and also autobiographical memories (see later sections of the thesis for a more in-depth review; e.g., Noreen et al., 2016). This effect could indicate a weakened association between a particular reminder and the memory, but there is also evidence for suppression-induced forgetting when memories are tested with independent cues (e.g., Bergstrom et al., 2009b). Moreover, suppression-induced forgetting is found regardless of direct suppression or thought substitution instructions, suggesting that both strategies may be effective for forgetting. Taken together, these results indicate that unwanted memories may become inaccessible after attempts at memory control when faced with reminders.

It is worth noting that two other strategies can be used to intentionally forget: Retrieval-induced forgetting and directed forgetting (see Anderson, 2005). Retrieval induced forgetting occurs when particular memory (e.g., apple) is retrieved from a category (e.g., fruits), then other category-related memories that are not retrieved (e.g., orange) are forgotten due to inhibition/interference (Anderson, Bjork, & Bjork, 2000). Directed forgetting involves providing explicit instructions (in contrast to suppression-induced forgetting) to either remember or forget certain items from a list of items, and usually, the “to-be-forgotten” items are subsequently forgotten in contrast to the “to-be-remembered” items (see Sahakyan et al., 2013 for further explanation). The think/no-think paradigm was used in the present thesis because it provides a unique opportunity to measure how difficult certain memories are to suppress and consequently forget, which would reveal specific motivations for us to engage memory control processes (see later sections for more detail).

1.1.1 Neurobiological models of direct suppression and thought substitution

Neuroimaging of participants’ brain activity when they are actively retrieving or suppressing memories in the think/no-think task has helped researchers develop neurobiological models of memory control. Much like the processes involved in stopping motor actions, stopping unwanted thoughts from coming to mind engages brain regions responsible for cognitive control that in-turn influence regions responsible for memory processes (see Anderson & Levy, 2009). Specifically, hemodynamic activity (measured by fMRI) in the hippocampus and other medial temporal lobe (MTL) regions is reduced during memory suppression attempts (Anderson et al., 2004; Benoit & Anderson, 2012; Gagnepain, Henson, & Anderson, 2014). Furthermore, suppression attempts also increases activation in the right dorsolateral and ventrolateral prefrontal cortices (right DLPFC and right VLPFC), that are important for volitional control processes (Benoit & Anderson, 2012; Depue et al., 2007), suggesting that enhanced engagement of prefrontal regions during memory control attempts suppresses memory retrieval processes mediated by MTL systems.

An important premise of memory suppression theories is that avoidance of retrieval is a direct consequence of top-down inhibitory processes. One criticism of this theory is that change in hippocampal activity could be interpreted as being greater in think trials compared to no-think trials, indicating a retrieval-driven rather than an inhibition-driven effect. However, when a baseline condition is introduced, the hippocampal activity is reduced compared to both the think and the baseline conditions, and moreover, this reduction predicts the suppression-induced forgetting effect (Depue et al., 2007; Levy & Anderson, 2012), consistent with an inhibitory account.

Yet another issue is that a reduction in hippocampal activity for no-think memories could simply suggest absence of retrieval during suppression trials without confirming the role of active inhibitory processes. Crucially, there is reliable evidence indicating that the prefrontal regions and medial temporal lobe regions together orchestrate the inhibition of memory traces from entering awareness. Firstly, the hippocampal reductions are negatively correlated with increase in rDLPFC activity (Depue et al., 2010; Depue et al., 2007). Two particularly convincing pieces of evidence elucidate the role of active inhibitory processes in retrieval avoidance: 1) Effective connectivity analyses of these brain regions found that the rDLPFC modulates hippocampal activity in a top-down fashion (Benoit et al., 2015; Gagnepain et al., 2014), and this rDLPFC negative coupling predicts suppression-induced forgetting (Benoit et al., 2015). Intra-cranial oscillation data derived from these brain regions also indicate a negative rDLPFC → hippocampal coupling during suppression attempts, providing evidence for top-down inhibitory mechanisms involved in retrieval avoidance; 2) The concentration of inhibitory GABA (γ-aminobutyric acid) neurotransmitters in the hippocampus predicted the individual differences in the ability to suppress memories, wherein higher hippocampal GABA concentrations predicted greater rDLPFC → hippocampal coupling during suppression attempts (Schmitz, Correia, Ferreira, Prescot, &

Anderson, 2017). Therefore, these results indicate that a targeted inhibition of memory traces by volitional control processes leads to purging these contents from awareness.

Thought substitution on the other hand operates in an opposing fashion. Rather than asking participants to “push the memory out of mind”, i.e., *direct suppression*, in turn inhibiting the memory trace, participants are asked to retrieve an alternate memory, i.e., *thought substitution*. Thus, thought substitution involves diverting retrieval processes from the unwanted memory to a substitute thought or memory trace, consequently purging the unwanted memory from awareness. The earliest neuroimaging evidence for the difference in these mechanisms is that the event-related-potential (ERP) marker for retrieval was absent for direct suppression but present for thought substitution instructions (Bergström, de Fockert, & Richardson-Klavehn, 2009b), indicating that retrieval is occurring during thought substitution but not suppression (see section 2 for a more detailed review of EEG methods in think/no-think research).

Research using fMRI further elucidated the difference in brain systems involved in these two memory control mechanisms. Two groups of participants attempted the think/no-think task wherein one group used direct suppression and the other used thought substitution to control their memories (Benoit & Anderson, 2012). Hemodynamic activity using fMRI was measured during these attempts at memory control and was compared to activity during retrieval. Forgetting of no-think compared to think and baseline memories in the surprise test was observed regardless of strategy used to control memories. Intriguingly, direct suppression led to the reduction of hippocampal hemodynamic activity whereas thought substitution had no significant reduction in hippocampal activity for no-think compared to think items. Interestingly, thought substitution led to increased activity in the left caudal pre-frontal cortex (left cPFC) and mid ventrolateral pre-frontal cortex (mid-VLPFC). These two regions were found to be more strongly connected for participants who had more difficulty in retrieving the substitute memory while maintaining the to-be-forgotten memory out of consciousness, whereas these two regions had weaker coupling for participants who were easily able to think of the substitute.

These left prefrontal brain regions were found to be involved in selective retrieval of a weaker memory when interfering with already present stronger memory traces (Wimber, Rutschmann, Greenlee, & Bäuml, 2009). Thus, the evidence implicates that thought substitution occurs due to competition-resolution by the left prefrontal cortex areas. Moreover, hippocampal activity was increased under conditions of high competition and correlated with the left prefrontal regions. These results imply that the unwanted memory may not be accessible at later attempts of retrieval because the competition is selectively weakened by the interfering memory (see Storm & Nestojko, 2010). Another explanation is that the substituted thought is strengthened and thus enters and remains in conscious awareness, so this substituted memory trace can block access to the competing unwanted memory (Bergström et al., 2009b; Levy & Anderson, 2002).

Importantly, this literature indicates that conscious control can be exerted over unwanted memory traces that can render them inaccessible when faced with a reminder in the future. This inhibition is driven by the prefrontal cortex, which exerts a top-down influence onto the hippocampus and other regions in the memory network, leading to the unwanted memory traces inaccessible during latter attempts at retrieval.

1.1.2 The role of unwanted intrusions in engaging memory control

The think/no-think paradigm also provides an opportunity to test the intrusiveness of a memory, which is an important unwanted influence of aversive memories where those memories are retrieved involuntarily. Therefore, the paradigm not only tests whether unwanted memories are later forgotten as a consequence of repeated memory control, but also provides a way to investigate the neuro-cognitive processes and behavioural outcomes immediately after each memory control attempt and therefore track the success or failures at control of memories over repeated attempts.

One perspective of involuntary retrieval has focused on *incidental* retrieval, for instance by asking participants to note down when a memory randomly comes to mind while they are doing everyday activities (e.g., Berntsen, 2010). This line of research investigates the mechanism of memories that enter conscious awareness without explicit intention to retrieve that memory. Sometimes these memories can be of positive episodes and lead to a pleasant state of being, but they can also be unwanted and lead to unpleasant emotional states (Watson, Berntsen, Kuyken, & Watkins, 2012). Critically though, Anderson and colleagues argue that although this retrieval process is unintentional and has applications in its own right, a more pragmatic definition of memory intrusions would refer to retrieval that occurs despite the person's best attempts at preventing it (Hellerstedt et al., 2016; van Schie & Anderson, 2017). Therefore, they argue that the think/no-think task provides a direct way to measure memory intrusions.

Levy and Anderson, (2012) introduced a phenomenological rating scale after each reminder (think or no-think) that measured how often the associated memory came to mind during the retrieval or suppression attempt. If the memory entered participants' conscious awareness despite their best attempts at stopping its retrieval, a memory intrusion is supposed to have occurred. Importantly this phenomenological rating of intrusiveness can be given on a trial-by-trial basis, providing a real-time measure of memory control ability. It is also important to note that to date, these intrusions have only been measured using direct-suppression instructions (see Anderson & Hulbert, 2021) because intrusions are operationally defined as failed suppression attempts. Although one could argue that difficulty in recalling the substitute is conceptually similar to a memory intrusion, to the best of my knowledge this has not been studied on a trial-by-trial basis yet. The difficulty/ease in thinking about the substitute has however been assessed in a post-experimental questionnaire (see description of Benoit & Anderson's, 2012 study in the previous section). Nevertheless, much is now known about the behavioural effects of memory control on intrusions measured immediately after suppression attempts.

On average, participants report experiencing memory intrusions around 60% of the time at the first attempt at memory suppression for weakly related word pairs, but ratings of intrusions are considerably reduced over repeated attempts to around 30% (van Schie & Anderson, 2017) of the time, indicating that although suppression is difficult at first, repeatedly applying memory control can help reduce intrusions. This reduction in suppression failures is often quantified by the slope of the reduction in intrusion frequency over trials. On an individual basis, a steeper slope (greater reduction in intrusions) was found to predict greater suppression-induced forgetting, indicating that the “online” ability to control memories also relates to the ability to forget as a consequence of memory control. Importantly, these results indicate that the unintended retrieval of memories can be controlled and the ability to control improves over repeated attempts.

The neural correlates of memory intrusions were further studied using fMRI (Benoit et al., 2015; Levy & Anderson, 2012). A reduction in hippocampal activity was found in trials that participants reported experiencing intrusions compared to trials with non-intrusions, and this reduction in activity predicted later forgetting. Interestingly, experiencing an intrusion engaged the rDLPFC more than when retrieval was completely avoided for no-think items. Markedly, this intrusion-related increased prefrontal activity predicted the negative slope in intrusion frequency. Moreover, individuals who showed greater rDLPFC → hippocampal negative coupling during early attempts had fewer intrusions later on in the task. This negative-coupling also got weaker over repeated attempts at suppression (Benoit et al., 2015). These two pieces of evidence indicate that the effort taken to control these memories reduces as suppression gets easier.

Taken together, these results indicate that when a memory is reactivated (intrudes), the rDLPFC is particularly recruited to assert top-down control over the intruding memory. Consequently, the need for this top-down inhibitory control seemed to reduce with the rate of greater control achieved by these participants. Thus, the reactivation of the intruding memory could make it especially vulnerable to disruption (also see memory reconsolidation evidence, Dudai, 2004; Nader et al., 2000). Another important outcome of the research is that although the DLPFC activity was greater for intrusions, it was not absent for non-intrusions compared to think trials, and participants indicated that they were still able to notably suppress their memories even during the first attempt (up to 40% of the time). These results indicate that reactivation is not necessary for inhibition processes to be triggered, but intrusions could potentially amplify the need for memory control.

Fascinatingly therefore, the evidence reviewed here indicates that intrusion of memories into awareness when faced with reminders lends those memories more susceptible to the consequences of memory control. Therefore, we may be more motivated to engage memory control processes for particularly intrusive memories (see Anderson & Hanslmayr, 2014). Furthermore, memory control may thus provide an adaptive benefit for mental wellbeing (see Engen & Anderson, 2018). These ideas are central to the present thesis and are explored in greater detail in section 3 of the literature review. They are also directly investigated in empirical Chapters 3 and 5.

1.2 Electrophysiological correlates of memory control

The evidence presented until this point indicates that inhibitory control mediated by the prefrontal cortex suppresses activation of unwanted memory traces as indexed by reduced activation in the hippocampus and other memory-related medial temporal lobe regions. Furthermore, this top-down control is more engaged when a memory intrusion is experienced than when retrieval is completely avoided, making intrusive memories particularly vulnerable to suppression. To this end, repeated attempts at this top-down memory suppression strategy can reduce the intrusiveness of memories, leading to later forgetting when faced with a reminder after attempts at memory control.

The present thesis used electroencephalography (EEG) to investigate the neural dynamics of memory suppression. Chapter 2 of the present thesis provides a more detailed review on EEG data extraction and analysis. In essence, changes in electrical activity generated by neuronal ensembles in response to a stimulus such as a memory cue is measured on the scalp with EEG. These changes are then compared across different experimental conditions to understand the neural mechanisms underlying a cognitive process such as memory retrieval and suppression. ERP effects that have a specific timing, polarity, and scalp location (sometimes referred to as "components") have been reliably found to correlate with different cognitive processes and can therefore serve as markers of those processes.

The EEG correlates of memory retrieval and control provide a complementary approach to understanding the mechanisms underlying memory control. The fMRI technique has high spatial precision and is therefore excellent at investigating these brain regions and systems that contribute to cognitive processes. However, they lack temporal precision and are poor at identifying "*when*" a process occurs. Therefore, high-temporal EEG data can track the dynamics of real-time processes that occur at a trial-by-trial level and provide additional insights into the mechanisms underlying memory suppression and especially memory intrusions.

1.2.1 ERP correlates of memory retrieval and control

A well-documented marker of memory retrieval is the left-parietal episodic memory effect (see Rugg & Curran, 2007 for a review). Around 500-800ms post presentation of a memory cue, event related electric potentials (ERPs) around the left parietal regions of the scalp are higher in amplitude when a memory is retrieved compared to when there is no retrieval in response to a reminder (see Mecklinger & Jäger, 2012). Therefore, this left-parietal positivity component is thought to index recollection, such that contents of episodic memory is retrieved into conscious awareness (Wilding, 2000; Wilding & Ranganath, 2012). Interestingly, this component is modulated by the strength of the recollective experience; strong recollections have found to correlate with larger positive effects compared to weaker or partial recollections (Vilberg, Moosavi, & Rugg, 2006). This literature therefore predicts that if suppression can inhibit retrieval, then such retrieval avoidance should be reflected as a reduction of the left-parietal positivity.

Indeed, a reduction of the left-parietal positivity for suppressed compared to retrieved memories is arguably the most reliable ERP effect in think/no-think literature. There is a consistent finding that memories in the think condition produce a significantly larger left parietal positivity than memories in the no-think condition when participants use a direct suppression strategy or no specific strategy is encouraged (Bergström et al., 2007, 2009a, 2009b; Hanslmayr et al., 2009; Hellerstedt et al., 2016; Mecklinger et al., 2009; Waldhauser et al., 2012). Suppressed retrieval (no-think) ERPs were also found to be similar to perceptual baseline ERPs, compared to voluntary retrieval (think) ERPs (Depue et al., 2013). Therefore, the ERPs provide more objective evidence that retrieval of a particular memory can be avoided. As Bergström et al., (2009b) explains, this suppression-related attenuation of the positivity simply reflects successful avoidance of retrieval but does not necessarily provide an index for the mechanism responsible for memory suppression. When participants report experiencing a memory intrusion, the positivity is attenuated the same way as when they report that they successfully avoided retrieval (Hellerstedt et al., 2016), similar to how hippocampal activity is reduced for both intrusions and avoided retrieval. Collectively, the lack of left-parietal positivity for suppressed memories indicates that retrieval is inhibited, and can be related to the downregulation of activity in the hippocampus for no-think items (Benoit et al., 2015).

Interestingly, Bergstrom et al., (2009b) also tested a group of participants that used a thought substitution strategy to prevent retrieval and found that the left-parietal positivity was not reduced when this strategy was used, indicating retrieval of memory traces during thought substitution, but this positivity was absent during direct suppression. As mentioned earlier (section 1.1), fMRI results also indicate that the hippocampus activity is upregulated for both think and no-think cues when the thought substitution strategy is used. Importantly, this finding indicates that the attenuation of the left-parietal effect reflects an avoidance in retrieval of the memory associated to the cue. These results further solidifies the idea that thought substitution and direct suppression have opposing mechanisms.

Presumably, if the modulation of the left-parietal positivity reflects successful suppression, then earlier time windows should provide an indication for top-down inhibitory mechanisms involved in suppressing the memories. In some studies, think cues elicit a more positive P2 than no-think cues, around 200ms post-stimulus onset in primarily central-parietal regions (e.g., Bergström et al., 2007; Mecklinger et al., 2009); however other studies do not find a modulation of the P2 (e.g., Waldhauser et al., 2012). The P2 peak is thought to reflect allocation of attentional resources to certain stimuli and provides an early indication of an increased attention towards think compared to no-think cues. Another explanation is that the colour differences in cues potentially elicits different early ERPs. Instead, cognitive control mechanisms seem to be reflected by other EEG effects.

Mecklinger et al., (2009) tested participants on the think/no-think memory suppression task and a stop-signal motor inhibition task. In the stop-signal task, participants press a button (left/right button) depending on the direction of an arrow shown on the screen, and importantly, they are asked to withhold (i.e., inhibit) their response when an auditory sound is played, which measures motor inhibition ability. They found that both no-think items, and trials that required motor inhibition elicited more negative ERPs than think items, and trials that had successful responses in central-parietal electrode sites around 400ms post-stimulus onset during both thought and motor inhibition. This effect was related to the deflection of the N2 peak, which is a highly reliable effect that reflects initiation of cognitive control in motor inhibition (see Donkers & van Boxtel, 2004). Mecklinger et al., (2009) also found that the no-think related negativity was correlated with the N2 in the stop-signal task. They also found that this negativity is greater for memories that were forgotten compared to not-forgotten memories that were suppressed in the TNT task. This N2 deflection for no-think items is not a consistent finding as it is present in some studies that use the think/no-think task (Bergström et al., 2009b; Chen et al., 2012; Streb et al., 2016) but not in other studies (Bergström et al., 2007; Bian et al., 2021; Hanslmayr et al., 2009; Hellerstedt et al., 2016; López-Caneda et al., 2019). The reason for this disparity in results is not entirely clear, but one observation is that most of studies that do find the N2 component seem to have based the timing and location of the effect, both on previous research and post-hoc visual inspection of their data, one exception is Bergström et al., (2009b) who analysed ERP data using an exploratory technique. This N2 therefore may not be picked up by exploratory techniques and may explain why different analyses strategies produce different results with the N2 (see also Chapter 6 for further discussion regarding this issue).

Curiously, new evidence suggests that another ERP component that occurs in the same temporal regions as the N2 may also be important for understanding memory control and the temporal dynamics of memory intrusions. An early component related to memory reactivation, known as the FN400, is found to be maximal in mid-anterior regions and peaks around 400ms post-stimulus-onset. Although there is a debate regarding the psychological processes reflected by this component (Leynes, Bruett, Krizan, & Veloso, 2017), it has been associated with reactivation of memories in both cued recall (Bergström et al., 2012; Hellerstedt & Johansson, 2014) and recognition tasks (Opitz & Cornell, 2006). This component is also thought to reflect familiarity to the cue (Mecklinger & Jäger, 2012; Rugg & Curran, 2007)

Hellerstedt et al., (2016) found that memory intrusions, voluntary retrieval, and successful suppression related ERPs elicited more positive FN400s compared to perceptual baseline items, indicating that the memory was reactivated regardless of intention to retrieve. Interestingly, a greater intrusion-related FN400 magnitude was correlated with more suppression-induced forgetting, whereas there was no correlation between the FN400 for avoided retrieval and suppression-induced forgetting. In a retrieval-induced forgetting study, Hellerstedt and Johansson (2014) found that the FN400 magnitude positively correlated with later forgetting only if competing memories were reactivated (see also Bergström et al., 2012). Therefore, Hellerstedt et al. (2016) argued that the FN400 reflects memory reactivation, and greater FN400 during intrusions indicates a greater need for cognitive control, in line with fMRI findings (Benoit et al., 2015).

Hellerstedt et al. (2016) also found a negative slow wave (NSW) component present for voluntary retrieval compared to the baseline from 550 to 3000ms post-stimulus onset in anterior and central regions but no difference in ERPs between avoided retrieval and the baseline was found. Interestingly, intrusions also had a negative slow wave compared to baseline but this effect was diminished over the time and disappeared around 1300ms. Furthermore, mirroring the behaviour results that frequency of intrusions experienced by participants reduce over repeated attempts, the size of the negative slow wave for intrusions was significantly smaller during the first half compared to the second half of the task. Considering that similar NSWs has been found to index traces entering working memory (Drew, McCollough, & Vogel, 2006), the intrusion-NSW effect seems to indicate that participants were able to purge unwanted memories from working memory (conscious maintenance) and improved their ability to control intrusions over repeated attempts at suppression. It is important to note that the study reported above is the only study investigating the ERP correlates of real-time intrusions, other than the present thesis. Therefore, these results should be considered with some caution, but it does provide valuable evidence that intrusions can be tracked in real time using EEG, and the findings do converge with results from fMRI data. The present thesis directly explores this issue further, especially for autobiographical memories, in Chapter 3.

In contrast to the negative slow wave, some other studies have found a sustained positive slow wave for think compared to no-think items in the same time region as the NSW, from around 800ms till the end of the epoch. This think vs. no-think positivity is thought to reflect cognitive control mechanisms working to keep the memory out of conscious awareness (Depue et al., 2013; Hanslmayr et al., 2009). Hanslmayr et al., (2009) found that this sustained positivity predicted suppression-induced forgetting. Moreoever individuals who showed high suppression-induced forgetting had greater sustained postivity than individuals with low forgetting. Therefore, this component could reflect an ability to maintain sustained cognitive control to limit traces of memories from entering awareness for the duration of exposure to the reminder, in line with the instructions provided to participants.

The ERP evidence reviewed here thus indicates an early P2 effect that potentially reflects selective attention towards cues on the basis of their colour, followed by memory reactivation (FN400) and/or cognitive control processes (N2) that lead to retrieval avoidance, as indexed by a reduction in the left parietal episodic memory effect. Then, sustained control is potentially maintained over the memory whereas memories in the think condition could be maintained in working memory.

1.2.2 Oscillatory correlates of memory retrieval and control

Oscillatory signals are rhythmic changes in electrical activity generated by neuronal ensembles that are reflected in the EEG (Bastiaansen, Mazaheri, & Jensen, 2012). This rhythmic activity can be due to neurons firing together at the same time (synchronisation) or firing at variable times (desynchronisation) within a particular period of time. The oscillatory activity is usually categorised into frequency bands based on the frequency of oscillatory activity: Delta (1-4Hz), theta (4-7Hz), alpha (8-12Hz), beta (13-30Hz), and gamma (30-100Hz). These different bands are also found to correlate with different cognitive functions. In short, theta synchrony and alpha/beta desynchrony are thought to reflect memory retrieval processes (see Hanslmayr et al., 2016 for further discussion). The literature reviewed below will focus on the oscillatory correlates of memory suppression and retrieval from the think/no-think paradigm.

A few studies have investigated oscillatory correlates of direct suppression. In the earliest study, Depue et al., (2013) found that in the theta and alpha frequency bands, think items produced reduced oscillatory power than no-think items around 300-800ms post-stimulus onset in parietal areas. Intriguingly, this think vs. no-think power desynchronisation was present for remembered memories but not forgotten memories, regardless of intention to suppress the memory. Therefore, the authors interpreted this as a retrieval-driven effect, possibly reflecting monitoring processes for successful retrieval of memories. Conversely, Ketz et al., (2014) found that think items had *greater* oscillatory power than no-think items (synchronisation) in the theta frequency bands but *reduced* oscillatory power (desynchronisation) for think compared to no-think items in the alpha and lower-beta frequency bands. Waldhauser et al., (2015) argue that these contradictory findings could be because the studies could not separate cognitive control mechanisms from the consequences in terms of avoided retrieval.

Waldhauser et al., (2015) modified the think/no-think task to measure the anticipatory effects of memory control. They instructed participants to think or not-think of memories *before* the reminder was presented. This way, they were able to delineate the anticipatory neural mechanisms that would be activated to engage intentional cognitive control, and the neural markers of the consequence of this cognitive control on the items where retrieval was successfully avoided. They also used source localisation of the oscillatory effects to understand the brain systems involved in memory suppression.

Waldhauser et al., (2015) found anticipatory markers of engagement of cognitive control processes in the theta band. Interestingly, and importantly to the present thesis, a widespread think vs. no-think power synchronisation was present in the theta, alpha, and beta bands from 500ms to 3000ms post onset of the reminder. This effect was driven by a greater reduction in power for no-think compared to think reminders in all frequency bands. Interestingly, the source analysis indicated that this effect seemed to be originating from different cortical structures for the theta band compared to alpha/beta frequency bands.

In the theta band, this reduction was more pronounced for high forgetters compared to low forgetters in the medial-temporal lobe. Although source localisation of EEG effects should be interpreted with caution, they found that the spatial localisation of the effects were in line with previous findings. Previous research using MEG has also found that theta power increases during retrieval seems to originate from the medial temporal lobe regions (Nyhus & Curran, 2010) and are often associated with successful retrieval of memories (Osipova et al., 2006). As aforementioned in previous sections of the thesis, suppression of retrieval also leads to hemodynamic decreases in these regions (Anderson et al., 2004; Benoit & Anderson, 2012). Therefore, the theta decrease for suppressed items is interpreted as a reflection of success in avoided retrieval. Suppressed items also induced reduced alpha/beta power compared to retrieved items, originating from prefrontal regions of the brain during anticipatory processes. Importantly, the authors argue that the sustained alpha/beta power decrease found post-onset of the reminder, is related to brain regions in the parietal lobe and other MTL regions that are involved in sustenance of maintaining control over retrieval during the time the reminder is shown on the screen, in line with the instructions provided to participants.

Therefore, in line with the evidence provided by other neuroimaging methods, the oscillation data indicates that top-down control exerted by pre-frontal areas on medial-temporal lobe areas are responsible for inhibiting the unwanted retrieval of memories. The oscillation data also indicates that parietal regions are involved in sustaining this control over retrieval throughout the presentation duration of the reminder. Note that results from this study was replicated and extended by Quaedflieg et al., (2020) and was replicated by (Legrand et al., 2020), potentially indicating that the findings were reliable. These two studies are more relevant for section 1.1.3. of this chapter, so they are discussed in more detail there.

A recent study has also investigated the oscillatory correlates of retrieval inhibition measured by intrusion reports after suppression attempts. As aforementioned in a previous section, the upregulation of right dorsolateral prefrontal cortex (DLPFC) is especially increased when participants report experiencing an intrusion. A very reliable effect in motor inhibition research is that beta band activity (around ~16 Hz) in right frontal regions around 200-300ms post-stimulus onset is increased for successful vs. failed attempts at stopping motor retrieval (Wagner et al., 2018). Interestingly, Castiglione et al., (2019) directly compared the oscillation data collected from participants who completed both the stop-signal and think/no-think tasks. Avoided retrieval had greater beta band activity than memory intrusions around 200-300ms post-stimulus, indicating that successful control of memories elicits greater beta activity. This was related to the right frontal beta synchronisation that was found for successful vs failed attempts at stopping motor actions. Legrand et al., (2020) also found that avoided retrieval induced greater beta power than intrusions, but this effect was found in a later time-window around 1000-1500ms post stimulus onset and in more frontal-central regions. Therefore, greater beta band activity could reflect success in controlling intrusive memories.

EEG based investigations of the mechanisms involved in memory intrusions has increased in recent years (e.g., Castiglione et al., 2019; Hellerstedt et al., 2016; Legrand et al., 2020; Streb et al., 2016) and have provided insightful evidence for understanding memory control processes. EEG can be a particularly useful method for this research as it could capture reactivation and cognitive control processes at an item-level at a high temporal resolution (see Benoit et al., 2015; Hellerstedt et al., 2016). Furthermore, to the best of my knowledge, there are no studies yet that use EEG to investigate intentional control over autobiographical memory retrieval in the think/no-think task. Chapter 3 in the present thesis directly investigates the neurocognitive mechanisms of autobiographical memory suppression and memory intrusions using EEG techniques.

1.3 Controlling emotional memories

The evidence presented thus far indicates that people can limit retrieval of unwanted memories by top-down cognitive control mechanisms. The literature discussed until this point considered evidence for memory control from healthy participants attempting to control mostly emotionally neutral memories. Notably, an important facet of what makes a memory unwanted is often the unpleasant emotions experienced while remembering such memories. Therefore, the benefits of memory control is not only concerned with reducing unwanted retrieval of memory contents into consciousness, but also reducing the unintended negative emotions elicited by these memories.

Memory control is now theorised to be a fundamental mechanism of emotion regulation, and it is reasonable to hypothesise that an ability to control unwanted memory retrieval is beneficial for maintaining good mental health (Engen & Anderson, 2018). A recent meta-analysis has revealed that people with clinical or sub-clinical levels of mental health issues perform worse than mentally healthy people in the think/no-think task, indicating that they have poorer memory control (Stramaccia, Meyer, Rischer, Fawcett, & Benoit, 2020). Persons with anxiety (Dieler, Herrmann, & Fallgatter, 2014; Marzi, Regina, & Righi, 2014a), depressive disorders (Sacchet et al., 2017), and PTSD (Mary et al., 2020) do not show the suppression-induced forgetting effect compared to healthy counterparts. Even sub-clinical populations that have indicators of poor mental health have impaired ability to control memories. For instance, participants with ruminative (Fawcett et al., 2015) or worry (Benoit, Davies, & Anderson, 2016) tendencies, intrusive thoughts (Catarino, Küpper, Werner-Seidler, Dalgleish, & Anderson, 2015), and dysphoria (Noreen & Ridout, 2016a, 2016b) all have impaired suppression-induced forgetting.

Strikingly, recent research provides evidence that memory control plays an important role in maintaining resilience after trauma (Hulbert & Anderson, 2018; Mary et al., 2020). In one study, survivors of the Paris terrorist attacks in 2015 were tested on their ability to suppress memories using the standard think/no-think paradigm with emotional neutral words (Mary et al., 2020). These victims were split into two groups depending on whether they developed PTSD post the terrorist attacks or not. The top-down connectivity between the right DLPFC and medial-temporal lobe was intact for survivors who did not develop PTSD, and this connectivity was comparable to participants who were mentally healthy and not exposed to the terrorist attacks. In contrast, this top-down connectivity was considerably diminished for survivors who did develop PTSD after the attacks, and moreover, these individuals also had impaired suppression-induced forgetting compared to the survivors without PTSD and non-exposed healthy participants. These results not only suggest that a lack of memory control is related to poor mental health, but also suggest that memory control can help build resilience towards traumatic events in life indicating its adaptive benefit for mental health and emotion regulation.

In another study, Streb et al., (2016) tested healthy participants on their ability to suppress memories of simple word stimuli. The ability control memories was assessed by suppression-induced forgetting and using ERPs, specifically the N2 component which is thought to index cognitive control processes. After the think/no-think task, disturbing and particularly intrusive memories were induced into participants’ minds, using traumatic films. A week later, they assessed how often memories from disturbing traumatic films (compared to neutral films) intruded into awareness in the past week during everyday activities. Interestingly, participants that showed greater suppression-induced forgetting and had more negative N2 potentials for no-think items, experienced fewer intrusions of the induced traumatic memories. Therefore, this study provided further evidence that having greater ability to control memories helps in reducing its unwanted influences.

New evidence indicates that inducing stress in healthy participants can influence the electrophysiological oscillatory dynamics of suppression (Quaedflieg et al., 2020) and consequently impair memory control. Before the think/no-think task, stress was induced for one group of participants whereas another group completed the task as normal. To induce stress, the researchers used the gold-standard Trier Social Stress Test (TSST; Kirschbaum et al., 1993) that mimics a tense job-interview with harsh interviewers that wear lab-coats and behave coldly with the participant. They found that an increase in cortisol was related to impaired suppression-induced forgetting. Expectedly, the EEG theta frequency power was reduced for no-think compared to think items in parietal areas around 2000-3000ms post onset of the reminder for the control group , in line with previous findings (Waldhauser et al., 2015). However, this theta power reduction for no-think items was not present for participants who faced stress before the task. Additionally, increased cortisol for stressed participants was related to reduced connectivity between the DLPFC and hippocampus, whereas in the control condition, this DLPFC to hippocampus connectivity was intact. Therefore, the results suggest that inducing acute stress before a memory control task can impair the ability to inhibit retrieval of unwanted memories. These findings are echoing the research reviewed earlier on the impaired memory control for people with mental health issues.

There is also direct evidence that attempts at suppressing memories controls both the intrusiveness and emotional responses to unwanted stimuli in parallel. Gagnepain et al. (2017) investigated the neurobiological mechanisms of controlling intrusiveness of negative compared to neutral memories using the think/no-think task (see also Depue et al., 2007). The right DLPFC was upregulated and medial temporal lobe regions were downregulated for intrusions compared to non-intrusions for both neutral and negative memories. Interestingly, the amygdala was also downregulated for intrusions compared to non-intrusions, and this effect was enhanced for negative memories. Using effective connectivity analysis, the reduction in amygdala activity was found to be due to an upregulation of top-down inhibitory control exerted by the DLPFC. This negative coupling was present for both DLPFC → amygdala and DLPFC → hippocampus connections (Benoit et al., 2015). Therefore top-down control mechanisms modulated activity in both emotional and memory retrieval regions, and this inhibitory effect was not restricted to memory retrieval processes but also applied to the emotional domain. Intriguingly, this top-down connectivity for suppressing mnemonic and emotional content was increased when an intrusion was experienced compared to when retrieval was avoided. The intrusiveness of memories may thus play an important role in acutely engaging these top-down inhibitory processes, an idea which is developed further in later paragraphs of this section. Interestingly, participants who had better control over intrusions also indicated (using a self-report questionnaire) that unpleasant emotions associated with viewing the negative images were reduced after the think/no-think task. Therefore, memory control not only downregulates brain activity of emotional areas, but also reduces the phenomenological emotional response to aversive stimuli, indicating a clear role in emotion regulation.

Recently, Legrand et al., (2020) investigated the oscillatory mechanisms of suppressing neutral or negative (disgusting) images. They measured physiological emotion responses using electrocardiogram (ECG) measures during both the learning of these items (before the TNT task) and after repeated attempts of suppression/retrieval (after the TNT task). They found that cardiac activity while viewing disgusting images after the think/no-think task was considerably inhibited for memories that were efficiently suppressed. They also replicated that the theta and alpha/beta EEG frequency bands have significant power desynchronisation for no-think vs. think items (Quaedflieg et al., 2020; Waldhauser et al., 2015), that is usually thought to index avoided retrieval and sustained control over memory respectively. Interestingly, they found that the suppression of theta power was significantly related to the subsequent inhibition of the cardiac activity. Therefore, controlling memories can also directly reduce the unintended physiological emotional responses to negative stimuli, further indicating its crucial role in emotion regulation. Taken together, these two studies (Gagnepain et al., 2017; Quaedflieg et al., 2020) indicate a clear role in memory control reducing the neural activity at the time of suppression, and in reducing physiological and phenomenological responses as a consequence of this successful control over memory, providing direct evidence for the role of memory control in emotion regulation.

There is considerably more research on the role of direct suppression in emotion regulation, so it is also important to consider the role of thought substitution in regulating negative affect. Curiously, some studies indicate that attempting direct suppression can have an ironic effect and lead to increased recall rates when faced with reminders after the think/no-think phase. This effect could arise because when participants fail at suppression attempts doing so can strengthen the memory trace rather than weaken it, leading to better recall when later faced with reminders. Thought substitution could therefore be a more successful strategy than direct suppression for people who find suppression difficult (Hertel & Calcaterra, 2005), and for reminders and/or memories that are so strong that direct suppression would fail (Ehlers, 2010). In reaction to an unwanted reminder, thought substitution is thought to involve selectively retrieving alternate memories rather than re-engaging attention towards to reminder itself. Therefore, not only is an alternative new pathway encoded between the reminder and an alternate memory, but the existing unwanted pathway is also inhibited (Engen & Anderson, 2018).

There is relatively less research on the role of thought substitution in emotion regulation. Interestingly, Stramaccia et al., (2020) point out that the thought substitution instructions could play a role in successful memory control. Substitution seems to be most effective when people are provided specific substitute memories that are semantically and categorically different from the to-be-suppressed memory (Benoit & Anderson, 2012; Hertel & Calcaterra, 2005; Joormann, Hertel, LeMoult, & Gotlib, 2009). Stramaccia et al., (2020) suggest that thought substitution seems to not be as effective when participants are asked to generate their own substitutes (e.g., Bergström et al., 2009b), or if there is considerable overlap between the substitute and unwanted memory which could be counterproductive as it could lead to reactivation of the unwanted memory. Therefore, it is important to consider the actual content of the substitute thought to maximise the efficiency of this strategy.

Recent research indicates that thought substitution may also not be effective at inducing forgetting for dysphoric individuals who have difficulty in inhibiting unwanted thoughts (Noreen & Ridout, 2016a, 2016b). This research also provide evidence for the idea that thought substitution, like direct suppression, works by an inhibitory mechanism and not only influences memory via competitive retrieval/interference. Participants in Noreen & Ridout's (2016a) study were provided emotionally neutral substitutes, whereas in another study (Joormann et al., 2009) emotionally positive substitutes were used and forgetting was observed for depressed participants. Therefore, the authors argue that the ability to successfully control memories using thought substitution potentially depends on the emotional nature of the substitutes (Noreen & Ridout, 2016a). Nevertheless, existing evidence on this issue is still unclear and it is important to compare the effectiveness of thought substitution strategies in memory control and emotion regulation. This question is directly tested in Chapter 5 of the thesis (also see section 1.4.3 for more detail).

Overall, the evidence presented above indicates that mentally healthy people tend to have robust top-down memory control mechanisms that also potentially helps regulate emotions, whereas this mechanism is often absent for people with poorer mental health. Furthermore, parallel mechanisms seem to support the control of mnemonic and emotional recollections of unwanted memories, and this control can reduce the physiological and phenomenological emotional reactions to aversive stimuli. These issues are empirically investigated and reported in chapters 3 and 5 of the thesis. It is also interesting to distinguish the mechanisms used to control memories to understand if one strategy provides an advantage over the other. This is specifically studied in chapter 5 of the thesis.

Negative memories are found to capture our attention more (negative bias), and are thus traditionally thought to be more intrusive and difficult to control (see Compton, 2003). However, there is mixed evidence regarding the intrusiveness and consequent forgetting of negative emotional memories in the think/no-think paradigm. Some studies find that forgetting for emotionally negative stimuli is greater than (Depue et al., 2006; Marzi et al., 2014; Noreen & Macleod, 2013), lesser than (Chen et al., 2012; Nørby et al., 2010; Sakaki et al., 2014) or not significantly different from (Joormann et al., 2005; Murray et al., 2011, 2015; van Schie et al., 2013) neutral stimuli. There have been fewer studies investigating intrusiveness (or the difficulty to suppress) negative compared to neutral or positive stimuli.

Although in all these studies, rate of intrusions decrease over repeated attempts at suppression, the overall difference in intrusiveness between negative and neutral memories is varied. One study did not find any difference in intrusiveness between negative and neutral images (Gagnepain et al., 2017). Another study found that negative words are less intrusive than neutral words (van Schie et al., 2013), but this difference between negative and neutral words was found primarily for items that were later forgotten. A recent study has found negative images are overall more intrusive than neutral images (Davidson et al., 2020). These contrasting results in our ability to control emotional memories could potentially be due to the relatively simple nature of stimuli used in most of these experiments (see also van Schie et al., 2013). These stimuli potentially do not capture the multiple reasons why people would find memories “unwanted”. This issue should therefore be investigated further with more complex stimuli that would particularly motivate people to engage memory control processes.

Motivated forgetting is a dominant theory arguing that memory control provides an adaptive benefit for mental health (Anderson & Hanslmayr, 2014; Benoit & Anderson, 2012; Gagnepain et al., 2017). Experiencing an intrusion has found to increase the *need for control* which particularly engages memory control processes, consequently regulating both unwanted thoughts and negative emotions associated to the memory (Benoit et al., 2015; Gagnepain et al., 2017). Therefore, people are expected to be more motivated to control particularly intrusive memories and consequently reduce retrieval of memory traces and regulate negative emotions associated to these memories. As this in an emerging field of investigation there are relatively fewer studies directly investigating these research questions. There are two specific gaps in the literature that are evident, which the present thesis aims to investigate: 1) Most of the literature uses relatively simple stimuli to investigate these questions. Although these stimuli are useful for understanding the neural mechanisms and behavioural outcomes on the actual task, they do not test the requirements of controlling complex autobiographical memories that people deal with in everyday life. 2) Most of the research contrasts negative and neutral or positive memories, which could be unpleasant, but potentially not particularly good motivators for memory control. Instead, research focussing on specific emotions could provide a significant advantage for teasing out the role of motivation in our ability to control memories.

There is evidence that memories of acting immorally could be strong motivators for engaging memory control processes due to the unpleasant feelings of guilt, shame and threat to self-concept associated with thinking of these events. The social psychology and behavioural economics literature has suggested that memories of one's own unethical actions become obfuscated over time, sometimes referred to as "unethical amnesia" (Kappes & Crockett, 2016; Kouchaki & Gino, 2016). Over several experiments, it has been found that memories of past unethical acts are less vivid and perceived to be more distant (see also Escobedo & Adolphs, 2010) than memories of ethical actions. However, recently, a team of cognitive memory researchers (Stanley, Yang, & de Brigard, 2018), were unable to replicate some of these findings and cast doubt on the concept of unethical amnesia, arguing that one's own immoral actions are actually likely to be more, rather than less memorable than other types of events. Specifically, Stanley et al. (2018) argue that memories for our unethical actions are associated with strong negative emotions that lead to particularly vivid and detailed unwanted memories. There is thus a conflict in the literature regarding whether memories of unethical actions are more or less memorable than other types of memories.

Notably, these apparently contradictory ideas can be reconciled from the viewpoint of theories of motivated forgetting (Anderson & Hanslmayr, 2014; see also Stanley & De Brigard, 2019). As indicated by the motivated forgetting theory, people should be particularly motivated to recruit cognitive control processes to suppress/distort such vivid and potentially intrusive unwanted memories, because intrusive negative memories are most necessary to suppress. Therefore paradoxically, if memories for immoral actions are initially very intrusive as Stanley et al., (2018) argue, then this may indeed make them more susceptible to motivated forgetting in the long term, as Kouchaki and Gino (2016) find. Therefore, in the present thesis, healthy individuals' ability to control autobiographical memories of past immoral actions was investigated in all three empirical chapters (3, 4, and 5).

1.4 Autobiographical memory retrieval and control

Autobiographical memories refer to a combination of events and facts from the personal past that usually defines the self (St Jacques & De Brigard, 2015). Therefore, retrieval of autobiographical memories is complex and usually includes several processes such as retrieving semantic content, personal significance, subjective re-experience, spatiotemporal context and emotion. Although simple mnemonic material (such as stimuli used in typical think/no-think experiments) can intrude into awareness in everyday life such as flashbacks of negative images or words, arguably it might be particularly difficult to control intrusive autobiographical memories because such complex contents of thought intruding into awareness may enhance feelings of reliving the experience and would also elicit unpleasant emotions (Marks, Franklin, & Zoellner, 2018). It is therefore important to investigate the mechanisms underlying autobiographical memory control and if such mechanisms can be successfully recruited for stopping unwanted mnemonic and emotional content from entering awareness.

Studying the neural mechanisms of autobiographical memory retrieval in a laboratory setting can be challenging however and has thus traditionally gained relatively less attention compared to retrieval of other types of memories (Cabeza & St. Jacques, 2007; St. Jacques & Cabeza, 2012). The primary challenge lies in eliciting autobiographical memories in neuroimaging scanners and lab-based experiments. Such experiments require unique reminders for a specific memory because a single cue could elicit multiple autobiographical memories. Specific to EEG research, the cues need to be relatively short (one or two words) to avoid ocular muscle activity generating noise in the EEG. Perhaps this is also the reason for the considerable lack in ERP research of autobiographical memory retrieval. Our understanding of brain systems involved in autobiographical memory retrieval has been largely based on fMRI and EEG oscillation methods. To the best of my knowledge, only one study has directly investigated the ERPs associated with autobiographical memory retrieval (Conway, Pleydell-Pearce, & Whitecross, 2001), and recent research has focused on understanding retrieval of specific contents of autobiographical memory (e.g., Tanguay et al., 2018). A few studies have investigated our ability to control autobiographical memories with the think/no-think task. One study used fMRI (Noreen et al., 2016), and a couple of studies used EEG measures albeit in a different paradigm (Hellerstedt et al., 2021; Hu et al., 2015) to investigate the neural correlates of autobiographical memory suppression. The present thesis therefore is the first investigation of electrophysiological correlates underlying autobiographical memory control using the think/no-think paradigm, specifically investigated in chapters 3 and 4.

1.4.1 Eliciting autobiographical memories in lab experiments

St Jacques and De Brigard, (2015) outline several ways to cue autobiographical memories in the lab, but two techniques are particularly relevant for the present thesis. One method is to conduct a “pre-scan interview”, wherein specific words related to a theme is provided to cue autobiographical memories. These words can be both neutral (e.g., “sport”) or emotional (e.g., “kiss”) and can be manipulated by the experimenter to elicit memories from a certain theme. Then, the participants usually provide a written or oral description of each memory. These generic cues are then used in experimental paradigms as reminders to elicit the associated autobiographical memories. One issue with this technique is that multiple memories can be associated to a certain cue word and could potentially interfere with each other when the cue is used to elicit that specific memory in the experimental paradigm. There are two ways experimenters have tried to overcome this issue: 1) Rather than using the original generic cue word to elicit the memory in the experimental paradigm, experimenters choose important snippets from the actual memory and present that along with the generic cue as reminders in the experimental paradigm (e.g., De Brigard et al., 2017); 2) Another method is to ask participants to come up with one-word “titles” that are unique to the autobiographical memory to ensure that only that memory is elicited by the reminder (e.g, Stephens et al., 2013).

Another method that provides control over the encoding of autobiographical memories unlike any other technique. This is known as the prospective method. In an innovative study, St. Jacques et al., (2008) attached a camera to the participants as they navigated around a museum and encoded autobiographical memories. The camera took pictures of specific scenes as participants navigated the room, and these pictures acted as memory cues in a later memory retrieval paradigm to elicit the specific autobiographical memories. The prospective method for encoding autobiographical memories is used in applied forensic memory research in the concealed information test, which involves instructing participants to encode an autobiographical memory of conducting a mock crime (see e.g. Hu et al., 2015; Hellerstedt et al., 2021). This method will be elaborated in the section 5 of the literature review which refers to the concealed information test.

1.4.2 Mechanisms of controlling autobiographical memories

In the think/no-think paradigm, the pre-scan interview technique has been primarily used to understand if autobiographical memories can be controlled. The first investigations of memory control over autobiographical memories were conducted by Noreen & Macleod, (2013) and Stephens et al., (2013). In the traditional think/no-think paradigm, participants learn associations between two stimuli pairs wherein one stimulus in the pair acts as the reminder to the other stimulus. In the autobiographical version of the paradigm, participants are provided cues to elicit autobiographical memories, and the memory description and phenomenological and emotional characteristics of the memory is collected by the experimenter for each memory. Then, either these cues act as reminders in the think/no-think task (Noreen, Bierman, & MacLeod, 2014; Noreen et al., 2016; Noreen & Macleod, 2013) or participants generate a unique reminder for each memory themselves, and these reminders are used in the think/no-think task (Stephens et al., 2013). This is the main paradigm difference between the traditional and autobiographical think/no-think tasks. Then, as usual, the reminders are presented in either green or red and participants are instructed to retrieve or suppress the autobiographical memories associated to the reminder. After the think/no-think task, participants describe and rate the autobiographical memories in response to the reminders in a surprise memory test.

To the best of my knowledge, only suppression-induced subsequent forgetting of autobiographical memories has been investigated to date, whereas the “online” ability to suppress these memories using real-time ratings of intrusions is yet to be studied. Therefore, the present thesis (see Chapter 3) is one of the first investigations of the intrusiveness of autobiographical memories measured immediately after attempts at memory suppression in the think/no-think paradigm, using a combination of ratings and electrophysiological measures.

In the traditional think/no-think paradigm, suppression-induced forgetting refers to words or pictures that could not be recalled at all in the surprise memory test. However, due to the emotional and personally relevant nature of autobiographical memories, such memories are unlikely to be *completely* forgotten. Therefore, it is important to carefully consider how to quantify forgetting in autobiographical think/no-think tasks. Stephens et al.'s (2013) results indicate an overgeneralisation of memories as an indication of a suppression-induced effect. As Stephens et al., (2013) describe, a memory initially described as "My sister accused me of turning our parents against her because they didn't want her to go to her favourite college" (p. 167) was later described as "My parents didn't like my sister's favourite college." (p. 167) in their data set. Autobiographical details of the memory were omitted from the latter description compared to the first description for memories that were in the suppression condition, and this was taken as evidence for suppression-induced forgetting. Noreen and Macleod, (2013) quantified forgetting by a slightly different method. When they collected autobiographical memories before the think/no-think task, they asked for specific details for each memory in addition to a general description of the memory. A strict criterion was applied to measure forgetting such that only if these specific details of the memory were not remembered in the surprise memory test, then forgetting occurred.

Both studies found that details of the memories were significantly reduced for no-think compared to baseline memories (Noreen & Macleod, 2013; Stephens et al., 2013). Importantly, a suppression-induced reduction in specificity of memories was found by both studies rather than forgetting of the complete episode of autobiographical memories. This effect has also been replicated in later studies with both direct suppression and thought substitution strategies (Noreen et al., 2016; Noreen & MacLeod, 2014). Noreen and Macleod (2013) also measured memory accuracy of a gist-based representation of memories and found that it was at ceiling regardless of suppression or retrieval instructions. Therefore, they argue that it is possible that certain details of a memory can be especially susceptible to memory control and subsequent forgetting based on people's motivations to forget.

The neurobiological mechanisms of autobiographical memory control has recently gained attention from researchers. As discussed earlier, and as Hu et al., (2017) explain, there is growing evidence that memory control has a domain-general effect: Memory control targets inhibition of specific contents of memory by top-down control mechanisms. A reliable effect of suppression is the downregulation of hippocampal activity indicating retrieval inhibition (Anderson & Huddleston, 2012; Benoit & Anderson, 2012), but memory suppression also targets neocortical regions, especially when memories intrude despite best efforts to inhibit retrieval (Gagnepain et al., 2014, 2017; Hu et al., 2017). Suppressing visual objects is coordinated by both DLPFC → hippocampus and DLPFC → fusiform gyrus negative coupling (Gagnepain et al., 2014). Therefore, both mnemonic and visual contents of memories are inhibited. Moreover, this negative coupling predicted greater impairment in implicit memory of these visual objects. Memory suppression can also target inhibition of semantic content (Hertel, Large, Stück, & Levy, 2012), potentially by downregulating areas in the medial temporal lobe responsible for conceptual priming (Anderson & Hanslmayr, 2014; Mayes et al., 2007). As aforementioned in a previous section, inhibiting emotional memories engages top-down control by DLPFC areas on the amygdala, and this activation is increased during unwanted intrusions (Depue et al., 2007; Hu et al., 2017).

Interestingly, autobiographical memories retrieval involves recruitment of most of these areas. Memory reactivation engages the medial temporal lobe network such as the hippocampus (Diana, Yonelinas, & Ranganath, 2007), whilst detailed recall and subjective recollection is carried out by other regions in the memory network (Addis, Moscovitch, Crawley, & McAndrews, 2004; Daselaar et al., 2008). The visual cortex is also especially recruited during autobiographical memory retrieval, potentially due to vivid visual imagery of past episodes (Cabeza & St. Jacques, 2007). The amygdala is also sensitive to autobiographical memory retrieval considering its emotional content (Young et al., 2016). Therefore, controlling autobiographical memories would theoretically involve activity-dependent inhibition of medial-temporal lobe and content specific neocortical areas, potentially based on our motivations to control specific contents.

Noreen et al., (2016) directly tested the brain systems underlying autobiographical memory control and found downregulation of hippocampal and visual areas in response to direct suppression compared to baseline items, whereas thought substitution similarly engaged brain areas as voluntary retrieval, in line with previous findings with simpler stimuli (Benoit & Anderson, 2012). It is however unclear which regions are responsible for top-down control of autobiographical memories. Noreen et al., (2016) did not find increased activity in the DLPFC for memories undergoing direct suppression, unlike think/no-think studies using simpler stimuli as reviewed earlier in the thesis. The authors argue that other frontal regions could be involved in autobiographical memory control but this is speculative. Two changes in the experimental design could explain the lack of DLPFC recruitment in direct suppression: 1) As aforementioned, generic cues were used to remind participants of the memories, and thus could elicit multiple memories during suppression attempts, which could lead to competitive inhibition instead of direct suppression. 2) Similarly, because direct suppression and thought substitution were manipulated within-subjects, participants could have used both strategies while not-thinking of memories. Therefore, there is clear evidence that retrieval of autobiographical memories can be avoided by downregulation of hippocampal and visual cortex areas, but the mechanisms underlying the cognitive control that leads to this downregulation is yet to be understood.

1.4.3. Different strategies for autobiographical memory control

An essential facet of memory control is its benefit for regulating negative emotions. Due to the narrative nature of autobiographical memories, novel strategies for applying memory control should also be investigated with the aim of understanding how they can help regulating negative emotions. In the present thesis, two different strategies that people may use in everyday life to cope with unwanted emotions is studied.

One strategy is to imagine alternate narratives to the autobiographical memory, known as episodic counterfactual thinking. Research on the mechanisms of episodic counterfactual thinking and its effects on emotions has gained traction in recent times. The neural and cognitive mechanisms responsible for counterfactual thinking are similar to retrieving episodes from memory (De Brigard et al., 2013). This is not surprising because the primary difference between mechanisms involved in retrieving the original memory and counterfactual thinking is that the content of memory is changed to come up with counterfactual scenarios, so rather than retrieving all traces from the original memory, different memory traces are reactivated and combined during counterfactual thinking (De Brigard et al., 2017; Parikh, Ruzic, Stewart, Spreng, & De Brigard, 2018; Schacter, Benoit, De Brigard, & Szpunar, 2015). More relevant to the present thesis, increasing evidence indicates counterfactual thinking has consequences on the affective responses when faced with reminders after counterfactual thinking. The alternate narratives that people generate can either be more positive (upward counterfactual), more negative (downward counterfactual), or neutral (changing semantic information) in the valence of the outcome. Kahneman and Miller, (1986) argued that emotions associated to imagining counterfactual scenarios are heightened during counterfactual thinking. Many studies have found that participants that imagine better alternatives to unwanted memories report feeling more regret and greater negative affect compared to participants who generated worse alternatives to the memory, since the latter reported feeling grateful and generally more positive about the original memory (Allen, Greenlees, & Jones, 2014; Epstude & Roese, 2008; Gilovich & Medvec, 1995; Rim & Summerville, 2014; Stanley, Henne, et al., 2017; White & Lehman, 2005).

A recent study compared the emotional valence before and after repeatedly inducing counterfactual thoughts for autobiographical memories (De Brigard et al., 2018). Regardless of the upward or downward counterfactual imagination, participants reported feeling less negative for negative autobiographical memories after compared to before the counterfactual task. This reduction in negative valence for negative autobiographical memories was also found for straightforward retrieval. Interestingly, the effect size in reduction of negative affect was greater when memories were simply retrieved compared to when counterfactuals were imagined, indicating that both simple retrieval may be more beneficial than counterfactual imagination for emotion regulation.. The think/no-think task also provides an avenue to test these questions further. Nevertheless, it is therefore important to consider counterfactual thinking as a strategy for emotion regulation. Therefore, in the present thesis, the effects of counterfactual thinking on the phenomenological emotional response to memories was also studied.

In terms of the neurocognitive mechanisms of memory control described in previous sections of the review, counterfactual thinking could be argued to fall into the category of thought substitution. However, a critical difference between the two strategies is that in thought substitution, the alternative memory trace should ideally be completely different from the original memory to ensure success in memory control (Stramaccia et al., 2020). However, counterfactual thinking involves reactivating some traces from the original memory and combining them with traces from imagination or other previously learned information (De Brigard et al., 2013), which may have complex effects on the original memories.

In the present thesis, counterfactual thinking was compared against direct suppression and thought substitution, which in this case was implemented by asking participants to use positive imagery to substitute negative thoughts of the memory. A recent meta-analysis indicates that imagery can be helpful for mental health patients to regulate their emotions, especially for overcoming negative cognitive biases (Hitchcock, Werner-Seidler, Blackwell, & Dalgleish, 2017). One study outlines the guided imagery scenarios such as mentally placing oneself as in a “safe space” (Adams, 2008; see also van der Hart, 2012). Participants are instructed to vividly imagine and focus on particular aspects of the beach such as the waves hitting the beach side, the smell of the beach, and the sounds of birds and also generally place themselves in that mental space. This emphasis on positive details could be particularly helpful because it could focus attention away from the unwanted memory to retrieving these vivid details about the beach from memory and consequently eliciting positive emotions associated with remembering these details. There is also evidence that memory control is more effective if specific - potentially more positive - substitutes are provided to participants prior to the think/no-think task to enable successful thought substitution (Joormann et al., 2009; Noreen & Ridout, 2016a; Stramaccia et al., 2020). Therefore, participants were provided with positive scenarios prior to the think/no-think task, which they used during the actual task.

The think/no-think task also provides a unique opportunity to investigate the effectiveness of these memory control strategies in more detail by comparing counterfactual thinking, direct suppression, and thought substitution instructions for the no-think condition. Furthermore, immediate phenomenological reports can be collected to understand if repeated attempts at memory control can help regulate negative affect related to unwanted memories.

Then, the effects of these strategies on forgetting, and phenomenological and emotional memory characteristics of these memories can also be investigated. These questions are directly addressed in empirical chapter 5.

1.5 Memory control in the concealed information test

In addition to the clinical applications of memory control discussed in prior sections, memory control can also be used in an applied setting to stop retrieval of guilty knowledge and consequently appear innocent in forensic memory detection tests that rely on recognition of incriminating information. The concealed information test (CIT; originally devised by Lykken, 1959), is used in forensic settings to detect knowledge about a significant piece of evidence from a crime, that is pertinent only to the criminal and the investigators (Rosenfeld et al., 1998). In this paradigm, a series of stimuli (words or images) are presented one at a time to a suspect. One stimulus is a key piece of information from the crime (such as the murder weapon), known as the “probe”, whereas the other stimuli are related to the probe but irrelevant to the crime, known as “irrelevants” (such as other weapons that could have been used in the murder). Importantly, only the criminal has *guilty knowledge* of the murder weapon and can specifically recognise the probe, whereas an innocent person who does not have guilty knowledge will not be able to differentiate between the probe and irrelevant stimuli.

In the computerised version of this task, suspects are asked to respond to stimuli on the screen by pressing a button to indicate “yes I have seen it before” and another button to indicate “no I have not seen it before”. When the person that has guilty knowledge, i.e., the criminal, perceives the probe (the murder weapon in this example), a physiological response is elicited that is clearly distinguishable from the physiological responses elicited by that person when other irrelevant stimuli are perceived. This difference is amplified if the person explicitly states that they do not recognise the probe by pressing the “no I have not seen it before" button, as they would respond for the irrelevant stimuli. However, if the person does not have guilty knowledge, i.e., they are innocent, then the physiological response is the same for both probe and irrelevant stimuli.

The CIT is based on detecting the orienting response, which occurs in many situations when there is a change in stimulus. The orienting response manifests as a physiological change and can be measured by skin conductance response, heart rate variability, and ERP measures (Verschuere et al., 2011). Physiological changes in the CIT are therefore argued to be indicators of the orienting response towards crime-relevant stimuli and markers of guilty knowledge because more personally significant stimuli elicits a larger orienting response than meaningless stimuli. Therefore, researchers argue that this is an accurate method to demonstrate that a suspect is guilty if there is a distinct physiological response to the probe compared to irrelevant items, and also to prove that a suspect is innocent if there is no distinguishable response between probes and irrelevants (see Meijer et al., 2014; Peter Rosenfeld, 2018).

There are several ways to design and conduct this test, ranging from self-report questionnaires to measuring physiological responses to the orienting response (see Meijer et al., 2014). Important to the present thesis, event-related-potential (ERP) differences between stimulus types as measured by an EEG is a popular method to detect the orienting response in guilty persons (see Rosenfeld, 2019 for a review). Research finds that if the suspect has guilty knowledge, the probe elicits a more positive ERP compared to irrelevant stimuli in mid-parietal electrodes (Pz in the 10/20 system) around 300-800ms post-stimulus-onset, but there is no difference between ERPs elicited by probes and irrelevants in the same region of interest if the suspect does not have guilty knowledge (but also see Bergström et al., 2013; Hu et al., 2015, 2017). This test is used in Japan to investigate real-world crimes (Osugi, 2018), but not in most other countries as it is found to be susceptible to countermeasures (see Ben-

Shakhar & Nahari, 2018 for a discussion of different countermeasures). Scientific research hence plays an important role in investigating the effectiveness of this test.

Researchers have adapted this paradigm to a lab-based setting using the prospective method mentioned in section 1.4.2, in order to provide a more controlled and reliable environment for investigating how well the test works. In these studies, participants are asked to simulate a “mock crime” and then the CIT was used to detect knowledge of that crime. In these experiments, participants are first instructed to navigate to a room and steal an object (usually a ring) from a bag that is hidden in that room. They are then asked to inspect the stolen object carefully to ensure strong encoding of the object, before bringing that object to the experimenter. The CIT is then administered as usual; the word “ring” (the probe) is inter-mixed with 6 to 7 irrelevant words referring to objects that could have also been in the bag. These words are shown one by one on a computer screen while EEG data is recorded, and participants indicate whether they recognise the object, via button press. Participants are instructed to press the button corresponding to “yes I have seen it” for one irrelevant word that the experimenter reveals to them before starting the CIT. This is known as the “target” word. This “target” always elicits a larger P300 ERP than other stimuli because the participant recognises the word, and also because the participants use the other hand to press the “yes” button, leading to additional motor-related amplification of the P300 response. Participants are told to press “no I do not recognise it” to all other stimuli, including the probe so as to conceal their knowledge of the mock crime.

The ERP elicited for the target is significantly more positive than the ERP for irrelevant items, in the P300 region of interest described earlier. This ERP difference thus provides a benchmark-P300 effect for that individual (or an average benchmark for a group of individuals), that can then be compared to the probe vs. irrelevant P300 difference. Another group of participants who do not commit a mock crime also complete the CIT with the same stimuli. Importantly, the P300 probe vs. irrelevant difference is not present for this “innocent” group, even though the target vs. irrelevant P300 is present. Therefore, this P300 component is not due to visual or semantic differences between stimulus types, implying that this component is reflecting recognition of the stolen object, and thus, guilty knowledge. This paradigm is traditionally known as the 3-stimulus-protocol (3SP) CIT (see Rosenfeld, 2018 for other types of CIT paradigms).

Interestingly, recent research has found that “guilty” participants can *beat* the test by intentionally suppressing the P300 response to the probe, such that there is a reduced or absent difference between the ERPs elicited by probes and irrelevants (Bergström et al., 2013; Hu et al., 2015, 2017), similar to the absence of P300 response found in the innocent group. Bergström et al., (2013) and Hu et al., (2015) adapted the memory suppression instructions from the TNT paradigm to the CIT procedure to investigate if participants can suppress guilty knowledge detected by the CIT. So, in addition to the CIT instructions, participants were told that the test relies on recognition of the stolen object, and to beat the test, they should avoid any thoughts about the mock crime from coming to mind, especially when the crime-relevant stimulus (i.e., the probe) appears on the screen. They found that participants were able to successfully attenuate the P300 ERP component. Importantly, there still was a significant P300 target vs. irrelevant effect in this group, suggesting that the P300 probe component was being actively suppressed by participants. This finding has important implications, as it suggests that criminals in the real world can beat the CIT and escape conviction, and also provides vital information about the susceptible nature of P300-related paradigms.

However, there is an important difference between a mock crime conducted in a lab setting and a real crime; committing a real crime involves more emotional valence and arousal than the mock crime. Osugi and Ohira (2017) found that higher emotional arousal during encoding (i.e., committing a mock crime) can significantly amplify the P300 difference. Additionally, Ben-Shakhar and Nahari, (2018) argue that these mock-crime procedures are artificial and can lead to inflated detection efficiency estimates in CIT paradigms that use a mock crime manipulation. Therefore, there is a clear reason to test if the P300 response can be suppressed when elicited by guilty autobiographical memories that would be more similar to memories associated with committing a real-life crime. These questions are directly investigated in Chapter 4 because to the best of my knowledge, the effects of memory suppression on emotional autobiographical memories in the concealed-information test has not yet been investigated.

1.6 Overview of the thesis

The present thesis is therefore an investigation of the mechanisms and consequences of controlling autobiographical memories of past immoral actions. This introductory chapter reviewed the different mechanisms of general memory suppression, thought substitution and counterfactual thinking. Furthermore, the consequences of memory control on retrieval of mnemonic and emotional traces of unwanted memories was discussed. Then, the challenges of studying autobiographical memories in the lab, followed by the extant autobiographical memory control literature was reviewed. Finally, the ability to beat an EEG based forensic memory detection test using memory suppression strategies was reviewed. As EEG methods were primarily used in Chapters 3 and 4 to understand the neurocognitive mechanisms of memory control, a more detailed review of EEG data handling and analysis is presented in Chapter 2.

Firstly, it is important to investigate if autobiographical memories of immoral actions are particularly intrusive and if they can be successfully suppressed. The think/no-think task in combination with behavioural, ERP, and oscillation measures of memory was used to answer these questions in Chapter 3. Furthermore, the consequences of suppression on the mnemonic and emotional contents of the memories were analysed using self-report questionnaires. Chapter 4 investigated the ability to beat the P300 based concealed-information test for both mock crimes and autobiographical memories of past morally-relevant acts using memory suppression. Participants were therefore tested on their ability to suppress P300 responses in the CIT for mock crime, morally right, and morally wrong memories. In Chapter 5, three different strategies to control memories were compared in their efficiency at regulating emotions associated to autobiographical memories of past immoral acts. For the first time in memory control research, the emotional responses using self-ratings were collected at the time of attempted memory control in the think/no-think task. Similar to Chapter 3, the consequences of memory control on the mnemonic and emotional contents of memory was also investigated. Finally, the main findings and their implications are discussed in Chapter 6, together with limitations and suggestions for future research.

Chapter 2 - Review of methodology for analysing EEG data

EEG was used in the experiments reported in both chapters 3 and 4. Similar EEG data processing and analysis techniques were used for both experiments. ERP and oscillatory data was extracted and analysed in both exploratory and focal analyses. These two methods provide a complementary approach to EEG data analysis. However, as experiment one (Chapter 3) was relatively exploratory and experiment two (Chapter 4) had strong predictions regarding the EEG results, slightly different parameters and/or techniques were used for analysis of the EEG data. Furthermore, in chapter 4, ERP data was used to individually classify participants as either guilty or innocent.

2.1 EEG data extraction

Event-related potentials and time-frequency oscillation data were extracted and analysed for both chapter 3 and 4 in the thesis.

2.1.1 Event-related potentials

Event-related potentials (ERPs) are small changes in electrophysiological voltages produced by neural activity in response to an event (Luck, 2005). ERPs are extracted by averaging together multiple EEG data segments that are time-locked to a certain event such as the presentation of a stimulus (e.g., memory cue), or a participant response (e.g., a button press).

ERPs were analysed in both chapters 3 and 4 because they can be used to measure different component processes associated with episodic memory retrieval (see Wilding & Ranganath, 2012), and ERPs have been successfully used in past research to understand the neurocognitive processes associated with memory suppression and intrusions (Bergström et al., 2007; Hellerstedt et al., 2016; Mecklinger et al., 2009), Furthermore, this method is critical to the concealed-information test (chapter 4) and a large body of research exists on the ERP-based CIT (see Rosenfeld, 2019). The P300 ERP component is useful as a memory marker because it is a very large amplitude effect (effect sizes around .8 or greater) and can therefore be detected at a trial-by-trial level. Importantly, this allows for individual guilt classification and is a clear advantage of the ERP technique.

2.1.2 Oscillations

Neural oscillations are rhythmic patterns of activity generated by populations of neural ensembles that can be measured on the scalp using EEG (Bastiaansen et al., 2012). A group of neurons can either synchronise (fire simultaneously) or desynchronise (either reduce in firing rates or not fire simultaneously). As neuronal synchrony increases, it enhances the amplitude of oscillations as measured in the scalp EEG. Increases or decreases in EEG oscillation amplitudes in response to events have been successfully used to investigate different cognitive processes associated with memory retrieval and suppression (Hanslmayr et al., 2016; Waldhauser et al., 2015), as a complementary measure to ERPs.

EEG oscillations are categorised based on the frequency of oscillatory cycles per second (1Hz would refer to 1 cycle in one second). Patterns of oscillatory activity tend to correlate within certain frequency bands (e.g., 4-7Hz) and vary across different frequency bands (4-7Hz vs. 8-12Hz), and therefore these different bands are often thought to reflect different cognitive processes (but see Cohen, 2019). The most common bands measured in scalp EEG are from 1-3Hz (delta), 4-7Hz (theta), 8-11Hz (alpha), 12-29Hz (beta), and 30-90Hz (gamma). The delta band strongly contributes to ERP effects so does not provide as much complementary information as higher frequency oscillations (Bastiaansen et al., 2012), and requires relatively long duration of EEG recordings for accurate analysis. Gamma frequencies above around 40Hz are difficult to measure because of interference from mains and muscle activity in high frequency bands, and high frequency oscillations are also typically very low amplitude, making them difficult to distinguish from noise. Therefore, EEG bands between around 4-30Hz are most useful and feasible to measure with scalp electrodes.

To estimate the amplitude (power) of EEG oscillations, the raw EEG data needs to be transformed into the frequency domain. In order to retain information about timing of oscillations, time-frequency decomposition can be conducted on EEG epochs using complex Morlet wavelet convolution (Roach & Mathalon, 2008). Complex morlet wavelets are complex sine waves (containing both real and imaginary values) tapered by a real-valued gaussian distribution (Cohen, 2019; Tallon-Baudry & Bertrand, 1999). This complex Morlet wavelet is then convolved with the EEG signal, resulting in a complex-valued signal from which both power and phase can be extracted at each time-point (Cohen, 2019). In the present thesis, only the power (in dB) of oscillations was extracted and analysed.

The width of the Gaussian distribution used during time-frequency decomposition determines the trade-off between the temporal vs. frequency precision. Traditionally, this parameter is measured by the number of cycles of the morlet wavelet, wherein lower number of cycles provides high temporal resolution but low frequency resolution, and higher number of cycles provides low temporal resolution but high frequency resolution (see Cohen, 2019 for a coherent explanation of this trade-off). In the present thesis, a 3-cycle wavelet width was used to improve temporal precision albeit at the cost of frequency precision. This decision was partly determined by the duration of epochs that were possible to extract given the experimental paradigm, as the same number of cycles were used to decompose all frequencies. To estimate lower frequencies (such as 4Hz) with a greater number of cycles (such as 5-6 cycles), the length of the morlet wavelet needs to be greater than 1300ms (which is the length of the epoch in the CIT). Therefore, although previous think/no-think studies such as Waldhauser et al., (2015) use a 5 cycle wavelet width, to maintain consistency across both EEG studies in the thesis, a 3 cycle width was used for both studies (similar to Hellerstedt et al., 2021).

2.1.3 ERPs and oscillation data comparison

ERPs are extracted by averaging EEG from many trials of a condition together, which cancels out time-variable noise over several trials and maintains the time-locked signal, thus improving the signal-to-noise ratio (Burgess, 2012; Cohen, 2019; David, Kilner, & Friston, 2006). However, such time-locking could also be disadvantageous in instances where the event-related processes are *jittered* (varying across time) across trials: For example, the parietal positivity ERP component related to recollection can vary in the latency across different trials for older adults (Murray et al., 2019). Similarly, one could reasonably expect that retrieval of autobiographical memories (as investigated here) could vary in time both within and across trials because different aspects of the memory could come to mind at different times post-stimulus onset. This could lead to an increase in false negative rates as true effects could be "averaged" out of the ERPs.

In contrast, oscillation-based analyses can be applied so that they are less susceptible to jitter than ERPs and could provide an advantage for analysing these types of processes. As Burgess, (2012) explains, event-related EEG data contains both phase-locked (evoked) and non-phase-locked (induced) brain responses. In short, non-phase-locked responses are power changes in ongoing EEG oscillations without the peaks and troughs being perfectly aligned across trials, whereas phase-locked responses are thought to originate from temporary synchronisation of neuronal networks evoked by an event, and independently of the ongoing EEG (but also see Burgess, 2012; Cohen & Gulbinaite, 2014 for alternate theories). ERPs are primarily reflective of phase-locked responses because those are maintained after averaging over trials, while the non-phase-locked responses cancel out after grand-averaging. However, induced activity can be retained in estimates of oscillatory power by applying the wavelet transform to single trial EEG data and then averaging the output of the transform across trials. Therefore, oscillation analysis output typically contains both phase-locked and non-phasedlocked activity. The non-phase-locked oscillatory activity could be especially useful as it can capture jittered cognitive processes (such as autobiographical memory retrieval) that could be averaged out when analysing ERPs.

2.2 EEG statistical analysis

EEG recordings are rich multi-dimensional data sets. Neural activity is measured across several electrodes across the scalp (1st dimension, e.g., 28 electrodes), several time-points (2nd dimension, 1750 time-points for 3500ms epochs at 500Hz sampling rate), and a 3rd frequency dimension (27 frequency points, with steps of 1Hz, from 4Hz to 30Hz). Therefore, a key step is to “reduce” the data-points to a manageable level before conducting significance tests.

In the present thesis, three techniques were employed to analyse the EEG data. Having *a-priori* knowledge of the expected effects provides a relatively straightforward analysis procedure. For instance, the concealed-information-test is based on the P300 component which is defined by particular time-points and electrodes. So, one can zoom in on this region for analysis (see Focal Analysis section). However, for more exploratory questions, a more whole-head approach is required. Both data-driven and region-of-interest based multifactorial ANOVA techniques were used in the thesis to investigate exploratory research aims.

2.2.1 Focal analysis

The most traditional method to analyse ERPs is to identify a certain region-of-interest (ROI) across the two dimensions (Hoormann, Falkenstein, Schwarzenau, & Hohnsbein, 1998). For instance, the P300 peak is generally found around 300-800ms in parietal sites (primarily in the Pz electrode). This technique has been used extensively in the P300-based CIT (Rosenfeld, 2019) and was thus also used in chapter 3. Within this region, the mean amplitude of the ERP is measured and is used as a dependent variable in parametric tests such as t-tests and ANOVAs.

The primary advantage of this method is that the region of interest can be consistent across all experiments that investigate a particular ERP component. Furthermore, this technique leads to a significant reduction in data-points to be tested and thus multiple comparisons is not an issue unlike exploratory whole-head analyses. However, there are some disadvantages to this technique: Importantly, it ignores all activity in electrodes and time-points lying outside the ROI, and thus misses out on possible effects in other regions. Furthermore, if the research is relatively exploratory, then choosing precise ROIs can be challenging and prone to experimenter bias. For instance, autobiographical memories include several details that could be retrieved at different times post-stimulus onset that may not be captured by a particular ROI. Therefore, in this thesis focal ROIs were used only for analysing ERPs in the P300-based CIT (chapter 4), because a rich body of literature exists on the P300 and it was therefore possible to decide the exact focal analysis details *a-priori* (as shown in the associated pre-registered analysis plan, see https://osf.io/c5uym/[1]).

2.2.2 Whole-head analysis

For the more exploratory research questions in both EEG studies, whole-head approaches were used.

Multifactorial ANOVAs. This method considers multiple electrodes and time-points in the data set. Several different ROIs across electrodes and time-windows are identified and are used as levels of a factor in an ANOVA (Hoormann et al., 1998). The electrodes are often grouped based on location on the scalp (e.g., left anterior, mid-posterior), and the time-points are categorized using adjacent time-windows starting from 0ms till the end of the epoch (e.g.,

---

[1] Note that the pre-registered is embargoed until Feb 10, 2022 and can be viewed only after that point of time.

0-100ms, 100-200ms… and so on). The ERP amplitudes are then averaged over each ROI (e.g., average amplitudes over 200-300ms in left-central electrode sites) and the values are used as dependent variables in ANOVAs.

Importantly, because electrode regions are analysed as factors along with the experimental condition factors in the ANOVA, the topographical distribution of the effect can be interpreted using this technique. For example, if a condition effect interacts with left vs. right hemisphere as a factor, this would mean that the condition effect is significantly larger across one hemisphere than the other. However, a critical issue with ANOVA analysis using multiple time-windows and electrode locations is the *multiple comparisons problem*, which considerably increases the family-wise error rate, consequently increasing the likelihood of false positives (inflated type I error; see Maris & Oostenveld, 2007). ERP differences due to noise could be interpreted as meaningful simply because they are statistically significant, even though they could be occurring due to chance. If Bonferroni corrections are conducted the alpha threshold would be extremely conservative due to the large number of tests, risking a failure to detect true effects (inflated type II error rate).

An additional problem with using multiple ROIs in ANOVAs is that it is difficult to determine the optimal ROI boundaries without selecting ROIs on a post-hoc basis, which is inappropriate due to being statistically biased. Determining ROIs on an a-priori basis however risks different effects being grouped together into one ROI, or split across multiple ROIs, which impedes clarity in the results. For instance, if a true effect is maximal across 150-250ms, but the data is grouped into two separate ROIs such as 100-200ms and 200-300ms, then there may not be significant results in either time window, but if the 150-250ms time-window is chosen, then the effect may turn out to be significant.

In the current thesis, the multifactorial ANOVA method was used to analyse ERPs in chapter 3 in order to facilitate direct comparisons of my results against prior ERP literature on memory suppression and intrusions that had used similar analyses (Hellerstedt, et al., 2016).

Cluster-based permutation tests. To address the problems with multifactorial ANOVAs, Maris and Oostenveld, (2007) proposed a non-parametric cluster-based technique for analysing ERP and oscillation data without splitting the data into separate ROIs, but that can still correct for multiple comparisons. This technique was used to analyse both ERP and oscillation data in both experiments. Unlike the analysis methods mentioned above that require the experimenter to identify specific regions of interest which could lead to false-positive results, this analysis identifies such regions using a data-driven procedure albeit at relatively lower temporal precision (see Sassenhagen & Draschkow, 2019 for an explanation). This analysis is run in two stages. In the first stage, adjacent data-points that show differences between conditions are identified and grouped into clusters. In the second stage, a permutation test is applied to determine which of those identified clusters survive a threshold for statistical significance based on the likelihood of finding clusters by chance in the data.

Identifying clusters. In step 1, t-tests comparing two conditions of interest are conducted at every time point (and frequency point for oscillation data) at each electrode site. Neighbouring data points in the 2-dimensional spatial-temporal (for ERPs) or 3-dimensional spatial-temporal-frequency (for oscillation data) space that individually result in p values < .05 form a cluster. To be considered a cluster, a certain number of neighbouring electrodes need to have significant t-test results in the same time (and frequency) points. This parameter is manually decided by the experimenters; the least number (two) of electrodes required was used in the present thesis as the spatial resolution was relatively low in both experiments.

This process of identifying clusters is similar to choosing a region of interests across the 2- or 3-dimensional space, but critically, free from any arbitrary divisions that experimenters may create since it is completely data driven.

Next, a cluster-statistic is calculated for each cluster, which is a vector containing two values: Magnitude, representing the total sum of the t-values in the cluster; and a direction that represents if the values of condition one are greater than (positive cluster) or lesser than (negative cluster) condition two.

Permutation tests. In step 2, non-parametric permutation tests using the Monte-Carlo resampling method is conducted. The Monte-Carlo method follows these steps: Step 1) Collate the trials from condition one and two into a large data set; Step 2) *Random partition*: Consider “x” number of trials were observed in condition one and “y” number of trials in condition two, then *randomly* sample “x” trials and sample “y” trials from the collated data set. This creates permutated-condition one (with “x” number of trials), and condition two (with “y” number of trials) data sets. Step 3) Calculate cluster-statistics for the permutated condition one vs. condition two comparison. Step 4) Repeat steps 2 and 3 several times (1000 permutations were conducted in the present thesis) and obtain a distribution of cluster-statistics. Step 5) Calculate the proportion of permutated cluster-statistics that are larger than the observed cluster-statistic, which yields the *p*-value for that cluster. If the *p*-values for the observed clusters are smaller than the alpha level ($p < .025$ for two-tailed tests), then there is a significant difference between the two experimental conditions. This process is first conducted for the largest observed cluster compared to the distribution of the largest permutated cluster, then the second largest, third largest… and so on.

Intuitively, one can interpret that a cluster that is defined by boundaries in the 2- or 3-dimensional space is statistically significant, implying a localisation of the effect in that particular region. However, researchers argue that the exact boundary of these clusters cannot be conclusively interpreted because the method does not control for error across time-points: The precise timing of the cluster (e.g., 245ms-502ms) is not necessarily accurate, but one can say it falls approximately between 250-500ms for instance (Maris & Oostenveld, 2007; Sassenhagen & Draschkow, 2019). This is a primary disadvantage of the technique: A cluster that seems to be present in left- but not in the right-hemisphere electrode sites cannot be firmly interpreted as a left-lateralised effect because the analysis has not determined whether the condition difference is significantly larger in the left than right hemisphere (i.e. a condition x hemisphere interaction effect). However as aforementioned, the multifactorial ANOVAs are able to address such questions regarding the topographic distribution of the effects more firmly.

Nevertheless, the primary advantage of the cluster-based permutation analyses is that it overcomes the multiple comparison problem and also avoids the issues of violating assumptions of parametric tests. An important assumption of parametric tests is that the observations are independently sampled, which is potentially violated by electrophysiological data. Sassenhagen and Draschkow, (2019) suggest that data from neighbouring time-points and electrodes may be highly correlated as the same neural generators in the brain can propagate signals to multiple neighbouring electrodes. Therefore, clustering data overcomes this assumption. Other assumptions of parametric tests such normal distribution of the data, which EEG data tends to violate is also avoided by non-parametric tests. Permutation tests could thus be useful in reducing the type I and II error rates for EEG data (Pernet, Latinus, Nichols, & Rousselet, 2015).

In the current thesis, the focal analysis of ERPs was used in chapter 4, and the multifactorial ANOVA analysis of ERPs was used in Chapter 3. The cluster-based analysis was conducted on both ERP and oscillation data in both Chapters 3 and 4.

2.3 Individual guilt classification

The primary function of the concealed-information-test used in Chapter 4 is its application in forensic settings. Therefore, individually classifying participants (suspects of a crime) as guilty or innocent is a key component for this research area. In the present thesis, two methods of classification were used: The traditional bootstrap method (see Rosenfeld, 2019; Farwell & Donchin, 1986; Farwell & Donchin, 1991), and a permutation-based classification method that is argued to be more sensitive than bootstrap tests (see Zoumpoulaki et al., 2015; but also see Rosenfeld & Donchin, 2015). Essentially, both these techniques work by re-sampling the observed EEG data at the trial level and creating a distribution of values, which is used to test if the observed ERP differences between conditions are reliable/significant within each person's data.

2.3.1 Bootstrap based classification

The bootstrap tests uses resampling with replacement to determine if an individual has a reliably different ERPs between conditions. If 50 observed probe EEG trials were collected for each participant, then the average of all these trials forms an "observed-probe-ERP", similarly an "observed-irrelevant-ERP" is computed from the irrelevant trials (which are typically more frequent, e.g. 250). During bootstrapping, a separate set of 50 trials is randomly chosen from this observed set of probe trials with replacement, so that some trials can appear more than once and other trials do not appear. This set can be termed as a bootstrapped sample, and the average of all these trials forms a "bootstrapped-probe-ERP". Similarly, a bootstrapped-irrelevant-ERP is generated by computing the average of a bootstrap sample of 250 trials that have been drawn with replacement from the irrelevant EEG trials.

This process results in two different ERPs: One bootstrapped-probe-ERP, and the other bootstrapped-irrelevant-ERP. The ERP amplitude at a region of interest (such as the

P300 peak) is computed for the bootstrapped-irrelevant-ERP and then subtracted from the equivalent value of the bootstrapped-probe-ERP. This equates to the critical measure, a bootstrapped-probe vs. bootstrapped-irrelevant difference value.

This process of generating bootstrapped-ERPs and computing the difference value is repeated many times (e.g., 1000 iterations as used here; but see Rosenfeld et al., 2016). This provides a sampling distribution of bootstrap-difference values which can be used to estimate the intervals of the difference. The final step is to compute the mean of this bootstrapped distribution and determine whether the 90% confidence intervals of this mean includes 0. If this value is above 0, then researchers can conclude with 90% (depending on the chosen cut-off) confidence that the observed-probe-ERP is truly more positive than the observed-irrelevant-ERP, suggesting that the participant is guilty. This process was thus conducted for every participant in each group, to classify them as either guilty or innocent.

2.3.2 Permutation based classification

The permutation-based classification method uses random permutations of condition labels assigned to EEG trials at the within subjects level to determine if an individual has a statistically significantly different ERPs between conditions The Monte-Carlo resampling method is used, similar to cluster-permutation tests. Trials from both probes and irrelevant conditions are combined to create a large dataset. Then, a set of randomly selected 50 trials (assuming 50 probe trials) are chosen to create a “permutated-probe-ERP”, and the remaining 250 trials are averaged to create “permutated-irrelevant-ERP”. This process involving randomisation of trials and computing permutated-probe vs. permutated-irrelevant difference is performed 1000 times. Thus, a distribution of the 1000 probe vs. irrelevant ERP differences is derived. If the observed probe vs. irrelevant difference is larger than 90% (arbitrary cut-off) of the values in the distribution of permutated probe vs. irrelevant differences, then the observed difference is interpreted as statistically significant, indicating guilt.

2.3.3 Bootstrap vs. Permutation tests

As Zoumpoulaki et al., (2015) explain, the bootstrap and permutation based tests are conceptually different because the bootstrap test approximates the test statistic distribution whereas the permutation test approximates the null hypothesis. By simulating data, they illustrate that the bootstrap distribution underestimates the maximum value. They also show that permutation tests do not suffer from this bias. This is especially important for the CIT as the test-statistic is usually the maximum value of the P300 component, or the difference between maximum and minimum values of the ERP (peak-to-peak measure, see methods section in chapter 4). However, in response to this article, Rosenfeld and Donchin (2015) argued that the bootstrap method that CIT researchers used was different from Zoumpoulaki et al., (2015) method and thus is not prone to this bias. Therefore, in the present thesis, both the permutation and bootstrap method (as used more predominantly in the literature) was conducted to individually classify participants as guilty.

2.3.4 ROC analysis

Both bootstrap- and permutation-based techniques require an arbitrary threshold for guilty classification, and the optimal cut-off may vary across samples. Receiver-operator-characteristics provide a threshold-independent technique to assess classification performance. As (Florkowski, 2008) explains, two groups of samples are submitted to the analysis, wherein one group is “positive” (guilty in this case) and the other is “negative” (innocent in this case). The proportion of participants that were in the guilty condition and tested guilty (“Sensitivity” or true positive rate) is plotted against the proportion of participants that were in the innocent condition and tested innocent (“1 – Specificity” or true negative rate). The area under this curve (AUC) is then calculated, which provides an estimate of how likely it is that a random participant will be classified as guilty or innocent. The AUC ranges from 0 = 100% wrong classification, .5 = chance classification, to 1 = 100% correct classification. This analysis was used to test the performance of both bootstrap- and permutation-based classification techniques.

Chapter 3 – EEG and behavioural evidence for direct suppression and intrusions of morally relevant autobiographical memories

Most of the extant evidence for memory control is derived from research on memories encoded in a lab-based environment, and they are relatively simple stimuli such as words or pictures (Anderson & Hanslmayr, 2014). There is limited research on the mechanisms involved in, and outcomes of, memory control of real-world complex autobiographical memories. The first experiment in the thesis investigated the neurocognitive mechanisms of suppressing autobiographical memories, and whether this can be measured using the think/no-think paradigm (see also Noreen et al., 2016; Stephens et al., 2013). Specifically, EEG data was considered in the present thesis because it provides real-time reflections of cognitive processes and has not yet been studied in the context of autobiographical memory control or retrieval. These seemingly theoretical questions can also have considerable implications to everyday life. It is unclear how the emotional nature of memories can affect our ability to control memories (see Davidson et al., 2020; van Schie et al., 2013). Morally relevant memories are usually associated with negative emotions of guilt and shame, and also tend to be threatening to the concept of the self, which are all ingredients that lead us to distance ourselves from such memories (Stanley & De Brigard, 2019). Therefore, memories of our past morally wrong actions could be particularly difficult to suppress, and consequently more vulnerable to change as a result. This is another research question that the present experiment aims to address.

The think/no-think paradigm used in the present experiment was adopted to study autobiographical memories, and was inspired by two previous studies (Noreen & Macleod, 2013; Stephens et al., 2013). Participants first remembered and described several autobiographical memories involving morally wrong and morally right actions that they committed in their past. They then reported phenomenological characteristics of each memory and rated the emotions associated with the memories using questionnaires. Importantly, they also created a title for each memory, which were used as cues in a subsequent think/no-think task to elicit the associated autobiographical memory.

The think/no-think task was conducted 24 hours later, identical to the traditional think/no-think paradigm for simpler memories. The critical difference being that the reminders elicited the autobiographical memories that participants provided in the previous session. As explained in Chapter 1, in the think/no-think paradigm, participants voluntarily retrieve (think) or suppress (no-think) memories when faced with reminders (Anderson & Green, 2001; Anderson et al., 2004). EEG was measured during the think/no-think task, at the time of retrieval/suppression attempts. Two behavioural measures were derived from the think/no-think paradigm in the present study, in line with previous research. Firstly, memory intrusions were measured in real-time on a trial-by-trial basis, similar to previous research (Benoit et al., 2015; Hellerstedt et al., 2016; Levy & Anderson, 2012). Participants indicated how often the memory came to mind while the reminder was on the screen. Importantly, an intrusion was thought to have occurred if the memory comes to mind *despite* attempts at memory suppression. Secondly, the after-effects of memory suppression on mnemonic and emotional content of memories was measured by a surprise memory test, which was completed after the think/no-think paradigm. Reminders of the memories were shown and participants were required to recall and describe the associated autobiographical memory, regardless of whether they were retrieved or suppressed in the think/no-think paradigm. To assess any consequence of suppression/retrieval on memory characteristics and/or on emotions associated to the memory, the phenomenological ratings of these measures were collected, similar to when they first generated these autobiographical memories.

Behaviourally, a consistent finding in previous research is that memories that were directly suppressed in the think/no-think task have poorer recall in the surprise memory test compared to memories that were retrieved or not shown in the think/no-think task (baseline), and this effect is known as suppression-induced forgetting (Anderson & Levy, 2009). In the present study, suppression-induced forgetting was operationalised using two measures of "specificity" (consistent with Noreen & Macleod, 2013) and "similarity" (consistent with Stephens et al., 2013). Effectively, specificity reflects a concrete change in how much specific details participants provide in their memory descriptions collected before and after the think/no-think task, whereas similarity refers to a relative change in any aspect of the two descriptions before vs. after the think/no-think task.

Secondly, previous research found that the frequency of intrusions that participants experience during early attempts of memory suppression is relatively high, but repeated attempts at memory suppression lowers intrusion frequency (Levy & Anderson, 2012). The present experiment provides the first investigation of this intrusion measure for autobiographical memories, and thus aims to replicate and extend on these prior results. In terms of the moral nature of memories, morally wrong memories are theorised to be more difficult to suppress (Stanley & De Brigard, 2019), and this would reflect in the intrusion ratings in the present study, such that morally wrong memories are rated as more intrusive than morally right memories by participants self-reports. Furthermore, if morally wrong memories are more susceptible to change, we should find greater suppression-induced forgetting for such memories than morally right memories, as predicted by motivated forgetting theory (Anderson & Hanslmayr, 2014).

Finally, the role of memory suppression in emotion regulation was also tested in the present study, such that specific emotions of guilt and shame, and general arousal were measured for each memory before and after the think/no-think task. If suppression does help regulate emotions (Engen & Anderson, 2018), then there should be a suppression-induced decrease in ratings of guilt, shame, and arousal for morally wrong memories but not morally right memories.

As this was the first investigation of EEG correlates of autobiographical memory suppression, an exploratory approach was used, and thus there were no concrete hypotheses regarding specific differences in brain activity in terms of retrieval/suppression or moral memory type. The research aims were thus to delineate the neurocognitive processes involved in retrieval, suppression, and consequent intrusions of autobiographical memories. Importantly, the research also concerns how these suppression-related components in brain activity are modulated by the moral nature of memories. The present study aims to find similar EEG effects for autobiographical memories as found in previous studies investigating EEG correlates of suppressing simpler memories.

The ERP data from previous think/no-think research indicates that selective attention to reminders based on its colour occurs first, around 200ms post-stimulus onset, followed by a combination of memory reactivation processes and consequent cognitive control processes (N2 or FN400; Hellerstedt et al., 2016; Mecklinger et al., 2009) from around 200-500ms that causes retrieval avoidance, which is evidenced by positivity for retrieved vs. supressed items from 500-800ms (Bergström et al., 2009a, 2009b, 2007; Depue et al., 2013). From that time point onwards, sustained control over memory is maintained to ensure that the memory does not enter into awareness for as long as the person is exposed to the reminder. The oscillatory data from previous studies indicate a widespread decrease in power for suppressed compared to retrieved items across theta, alpha, and beta bands, which potentially reflect retrieval avoidance and maintenance of cognitive control to ensure that the memory does not enter awareness (Legrand et al., 2020; Quaedflieg et al., 2020; Waldhauser et al., 2015). If morally wrong memories are indeed difficult to control compared to morally right memories, these components that delineate retrieval and suppression should be less defined for morally wrong than morally right memories.

3.1 Method

3.1.1 Participants

Forty-one students at the University of Kent completed the study in exchange for a combination of course credits and £5. Data from seven participants were invalid due to a programming error and were not analysed. Therefore, only data from 34 participants between the ages 18 to 21 years (24 females; $M_{\text{age}} = 19.15$; $SD_{age} = 0.78$) were included in further analyses. The sample size was based on related think/no-think studies that report around 30-36 participants (the closest being Hellerstedt et al., 2016; see also Depue et al., 2013). All participants had normal or corrected-to-normal vision, were psychologically and neurologically healthy and were not taking any psycho-active medication. They were also advised to not take part if they were generally feeling low or depressed due to potentially upsetting nature of the study. Each participant provided informed consent before taking part. The study was approved by the University of Kent's School of Psychology ethics committee.

3.1.1 Design Materials and Procedure

This study was run in two sessions. In the first session, participants generated autobiographical memories and associated memory cues for each memory. These cues were used as stimuli in the TNT task in the second session, conducted 24 hours later (see Figure 3.1). Session one took 1 to 1.5 hours to complete whereas session two took 2 to 2.5 hours to complete.

Session one. This session was conducted in a lab using the online Qualtrics survey software. Participants were instructed to think of 22 different autobiographical memories one at a time (10 morally wrong, 10 morally right, one birthday, and one holiday), and type a description in a text box provided to them on a computer screen using a keyboard. To aid recollection of morally wrong actions, they were instructed to think of memories where they lied, cheated, physically or emotionally harmed someone, or any other act they considered was morally wrong. Similar examples were also provided for remembering past morally right actions, such as memories where they were truthful, helped someone physically or emotionally, or other morally right actions. Participants had three minutes to think of and write about a specific memory in two to three sentences by describing their actions, the persons involved, the location, and how they felt. They were instructed to avoid writing about events that easily blend with other memories and to describe each event in as much detail as possible.

After describing each memory, participants were instructed to think of a unique and specific personal title (henceforth referred to as the “cue” in this thesis) that would help them recollect the same exact memory in the next session when cued with the title. They were also instructed to avoid titles that could evoke multiple memories. The cues were split into a set of cues for morally wrong memories, one set of cues for morally right memories, and one set of filler cues.

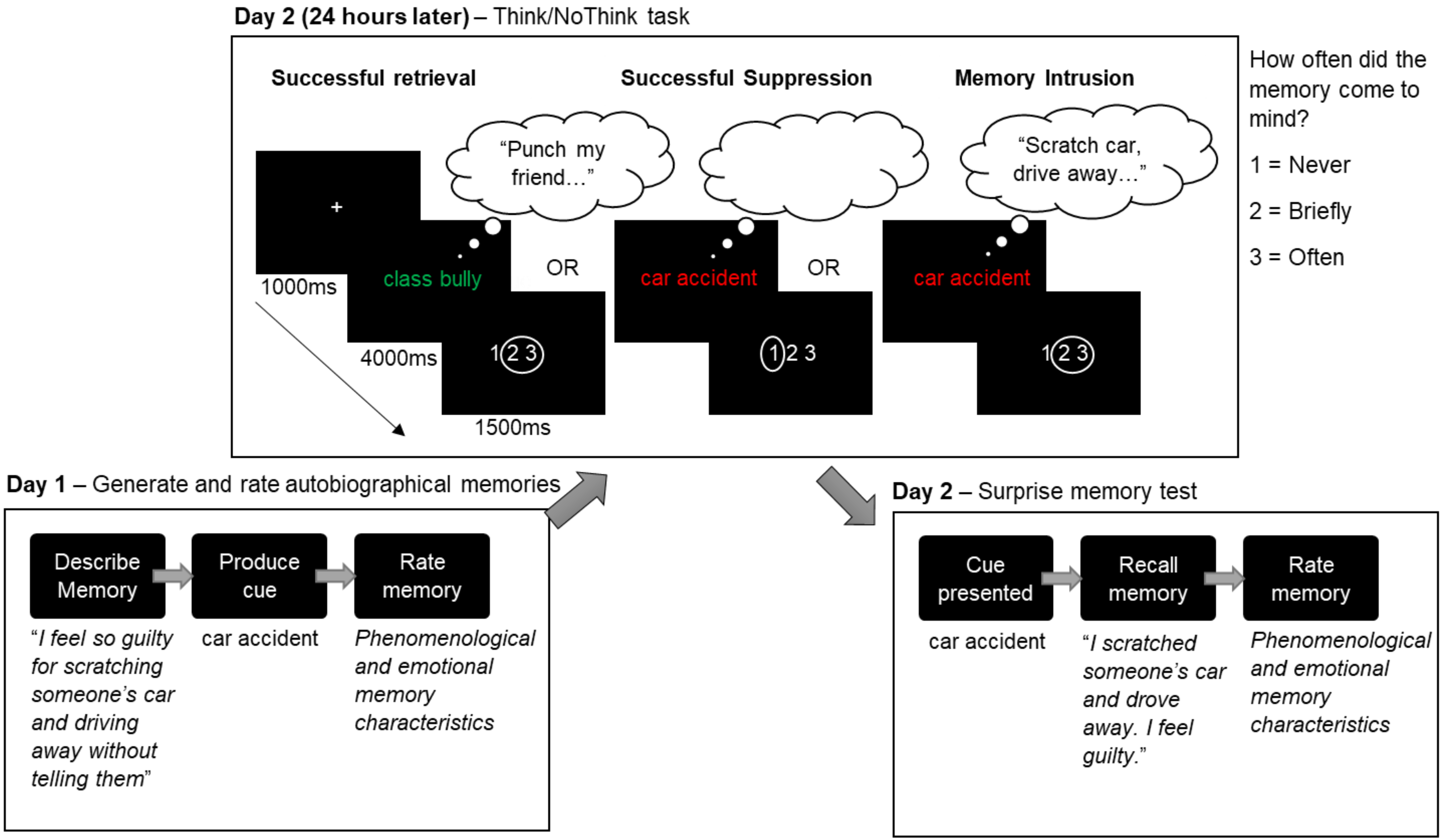


Figure 3.1. Illustration of one trial in the critical procedureal phases in the experiment. Participants attempted to suppress (no-think) memories associated to red cues, but consciously recollect the associated memories green cues. Participants indicated how often the associated memory came to mind using the “1 2 3” rating scale.

After participants generated the cue, they rated the memory for age (in years), vividness (How well do you remember the event? *1 = not well at all, 5 = extremely well),* intentionality (How intentional were your actions? *1 = not intentional at all, 5 = extremely intentional*) and morality (How morally right or morally wrong were the actions you performed during the event? *1 = very morally wrong, 7 = very morally right).* These phenomenological ratings have been used previously in moral autobiographical memory research (see Stanley et al., 2017). The participants were then asked to report emotional affect and valence using two scales: a modified version of the I-PANAS-SF: "Indicate the extent to which you feel this way at the present moment of the event" (e.g., Guilty, *1 = very slightly or not at all, 5 = extremely*); and SAM, which is a pictorial scale that measures emotional pleasure and arousal. Refer to Appendix A.1 for an example of the questionnaire.

Session two. Participants were fitted with electrodes attached in an EEG cap, and impedances were reduced. As aforementioned, the titles that the participants generated in session one were used as stimuli in this session to cue the autobiographical memories. To that end, an initial cue-test phase was conducted to ensure that participants could remember the autobiographical memory associated to the cue. Next, they extensively practiced the think/no-think task and intrusion ratings over three separate stages until they understood and were following all instructions. See Appendix A.2 for details of these two phases of the experiment.

Think/no-think phase. In the next phase, the 10 morally wrong and 10 morally right cues were pseudo-randomly assigned to the think and no-think conditions in equal proportions by the software, resulting in five cues in each of the four conditions (morally wrong think, morally wrong no-think, morally right think, and morally right no-think).

In each trial, a white fixation cross was first presented on a black background for 1000ms, followed by a cue presented in either green or red for 4000ms. If the cue was displayed in green (think), participants were instructed to constantly think about the associated memory for as long as the cue was on the screen, whereas if the cue was displayed in red (no-think) they were instructed to prevent any thoughts of the associated memory from coming to mind by pushing it out of awareness, while paying full attention to the cue the whole time it was on the screen. They were instructed to use a direct suppression strategy and not replace thinking of the memory with any other memory, image, or word (Bergström et al., 2009b).

After each cue, a black screen was shown for 200ms, followed by the intrusion rating scale for 1500ms, which showed “1 2 3”, wherein participants responded using keyboard button press. Participants were asked to withhold responses until the rating scale was displayed, and then indicate their response by pressing buttons on the keyboard. They were instructed that if the cue was shown in green and they managed to think of the associated memory for as long as the cue was displayed on the screen, they should press 3. However, if they failed to think of the memory, they should press 1, and they should press 2 if they thought of the memory only briefly. They were then instructed that if the cue was shown in red and they managed to prevent any thoughts of the associated memory from coming to mind then they should press 1. However, they should press 3 if they struggled to keep the memory out of mind and thought of it repeatedly, or if it came to mind briefly, they should press 2. After the rating scale screen ended, the next trial began.

The software was programmed to first present all 20 cues pseudo-randomly, one at a time, then after all cues were presented, the order of presentation was pseudo-randomised again and presented one at a time again, and this process was repeated 16 times. So, each cue was repeated 16 times, leading to a total of 320 trials. Participants were given a short break after each set of 40 trials. The software was also programmed to present cues in one condition (either think or no-think) not more than three times in a row, in line with previous research (Hellerstedt et al., 2016).

Surprise memory test phase. Participants completed a surprise test of the autobiographical memories associated with all cues during the final phase of the experiment, in the Qualtrics software. They were instructed to disregard all previous tasks and recall the memory for each cue regardless of the colour it had been displayed in during the previous phase. Like in session one, participants typed descriptions of the memories in a text box on the computer using a keyboard. In this phase, a cue was displayed on the screen for 10s and participants pressed a button as soon as the associated memory came to mind, which was followed by a text box. They had two minutes to describe the associated memory in as much detail as possible. They then reported same memory characteristics as in session one: memory age, vividness, intentionality, morality rating, I-PANAS-SF, and SAM measures. See Figure 3.1 for an illustration of the procedure.

At the end of this phase, participants completed a compliance questionnaire (adapted from Hu et al., 2015) to ensure they had completed the task as instructed. All participants were compliant with the instructions and were therefore included in further analyses.

3.1.3 Autobiographical memory description analysis strategy

Participants provided a written description of each memory both before and after the think/no-think task, which allowed us to investigate if the descriptions changed as a result of thinking or not thinking about the memories, and to test if this change was modulated by the moral content of the memory.

Each memory description was manually coded on the following measures: i) Specificity, which was quantified by identifying four core aspects of the memory event – *who* was involved, *what* happened, *where* did the event take place, and *when* did it take place. As these were moral memories of personal actions towards others, the "who" and "what" aspects were prioritised. This specificity measure was rated on a Likert scale ranging from *1 = Not at all specific* to *5 = Extremely specific*; ii) Emotional valence was quantified by identifying specific positive or negative words that conveyed the emotion felt by the participant. This was rated on a scale of *1 = Very negative, 4 = Not negative or positive, 7 = Very positive*. Finally, for each memory, the iii) similarity of both descriptions were rated on a general intuitive sense of similarity between the descriptions. Importantly, this measure was used to test if the content of the memory was changed or forgotten between the first and second report. A 5-point scale was used to measure similarity, *1 = Not at all similar, 5 = Extremely similar*.

The experimenter rated all memories in the study and five different independent coders second rated 50% of the memories (10% each coder). Although all coders had prior knowledge of the experimental design, they were blind to whether the memories had been assigned to think or no-think instructions and did not know if the descriptions were from before or after the think/no-think task. Due to the nature of the memory descriptions, it was not possible to blind the coders to whether the memories were morally wrong or right. The correlation coefficient between the experimenter and independent coders was .75 (ranging from .66 to .89), suggesting good agreeability between coders. Please see Appendix A.3 for full instructions that both the experimenter and independent coders used to rate the autobiographical memories.

3.1.4 EEG recording and pre-processing

The EEG data were recorded with a bandpass of 0.05-70Hz using 30 Ag/AgCl electrodes fitted in an EasyCAP with a sampling rate of 500Hz, amplified with a BrainAmp DC amplifier. Data was recorded with an average reference, and AFz was used as the ground electrode. Electrodes were placed below and above the right eye to measure vertical eye movements, and on both left and right outer canthi to measure horizontal eye movements. Impedances were reduced below 5 kΩ before starting the experiment by gentle abrasion of the scalp using cotton buds and saline gel.

EEG data was pre-processed and analysed using the EEGLAB toolbox (Delorme & Makeig, 2004) for Matlab version R2018b and self-written code. For each participant, data from all electrodes were first re-referenced offline to the average readings of the left and right mastoid electrodes. A 0.1Hz high-pass filter was applied to the re-referenced data. Then, the continuous EEG data was divided into epochs beginning 1000ms pre-stimulus-onset and ending 4000ms post-stimulus onset, baseline corrected on the -200-0ms pre-stimulus period. This data was visually inspected to delete epochs with extremely noisy data or channels. Independent component analysis (ICA) was then conducted to identify and remove ocular and muscle related components. Then, a 30Hz low-pass filter was applied to the data. Finally, any remaining epochs with large noise was removed based on visual inspection, and epochs were baseline corrected again against the -200-0ms pre-stimulus period.

3.1.5 EEG data extraction

Two complementary measures were derived from the epochs for analysing EEG data. First, ERPs were computed by averaging over all trials in each condition for each participant, and the amplitude of these ERPs (measured in μV) was one measure. Second, the single trial epochs were submitted to a time-frequency decomposition in FieldTrip (Oostenveld, Fries, Maris, & Schoffelen, 2011), using a complex Morlet wavelet transform to decompose the EEG into estimates of oscillation power across different frequencies and timepoints. Twenty-seven wavelets with centre frequencies ranging between 4-30Hz (in steps of 1 Hz), were convolved with the EEG data to produce power estimates across time-steps of 5ms. Each wavelet had a width of three cycles to prioritise temporal resolution over frequency resolution. To remove edge artefacts, epochs were truncated to -625 to 3500ms. A pre-stimulus baseline period of -625 to -325ms was used to normalise the oscillatory power to decibels (dB), since using a baseline period closer to stimulus presentation may lead to artificial “bleeding” of post-stimulus activity into the pre-stimulus period.

3.1.6 EEG statistical analysis

For the primary analysis of ERPs and EEG oscillations, epochs were divided into separate conditions based on a 2 (Instruction Type) x 2 (Memory Type) design in each half of the think/no-think task. Therefore, there were four different conditions in each block. On average, 36 trials remained in each condition after pre-processing the data (with individual ERPs ranging from 17-40 trials).

A complementary analysis of ERPs and EEG oscillations was also conducted to investigate the neural correlates of memory intrusions. Epochs were categorised into three conditions based on think vs. no-think condition and participants’ introspective reports of retrieval using the intrusion rating scale. The conditions were: Successful Retrieval, if participants indicated that memories came to mind briefly or often in the think condition; Intrusions, if participants indicated the memories came to mind briefly or often in the No-think condition; and Successful Suppression, if participants reported that the memories did not come to mind in the no-think condition. This analysis was collapsed across morally right and wrong memory types and block to ensure sufficient trial numbers for adequate signal-to-noise ratio. There were on average 130 (range 79-158) trials in the successful retrieval condition, 101 trials on average (range 69-131) in the successful suppression condition, and 31 trials on average in the intrusions condition (range 13-64) after pre-processing the data.

Because of the relatively exploratory research questions in this experiment, statistical analyses investigated possible condition differences across locations and time without focusing on pre-determined time-points or electrodes of interest. Two different statistical methods were used to conduct these exploratory analyses.

Multifactorial ANOVA. In an initial analysis, I followed the most relevant prior ERP research by Hellerstedt et al., (2016), which involved applying multifactorial ANOVAs to ERP amplitudes that was averaged across different scalp regions and sliding time-windows. This analysis was conducted to enable a direct comparison between my ERP results and those prior findings reported in the literature. Multifactorial ANOVAs with electrode region as a factor has often been used in ERP research to investigate the topographical patterns of ERP effects within sliding time-windows. Such analyses involve testing a very large number of main effects and complex interactions and are therefore not typically corrected for multiple comparisons, because doing so would render the analysis overly conservative and prone to Type-II errors. For this reason, multifactorial ANOVA analyses (without alpha-level correction) of ERPs are sensitive but prone to false positive results.

Since time-frequency decomposed EEG data contains frequency as an additional dimension, dividing all three dimensions (time, electrodes, and frequency) into different variables for multifactorial ANOVAs would be even more prone to false positive results than two-dimensional ERPs, so EEG oscillations are typically not analysed with factorial ANOVAs. Hence in this experiment, only ERP amplitudes were analysed using this method.

For the ERP ANOVAs, electrodes were separated into nine different regions (see Figure 3.2) and ERP amplitudes were averaged within each region and across 11 different time-windows: 100-200ms, 200-300ms, 300-400ms, 400-500ms, 500-750ms, 750-1000ms, 1000-1500ms, 1500-2000ms, 2000-2500ms, 2500-3000ms, 3000-3500ms.

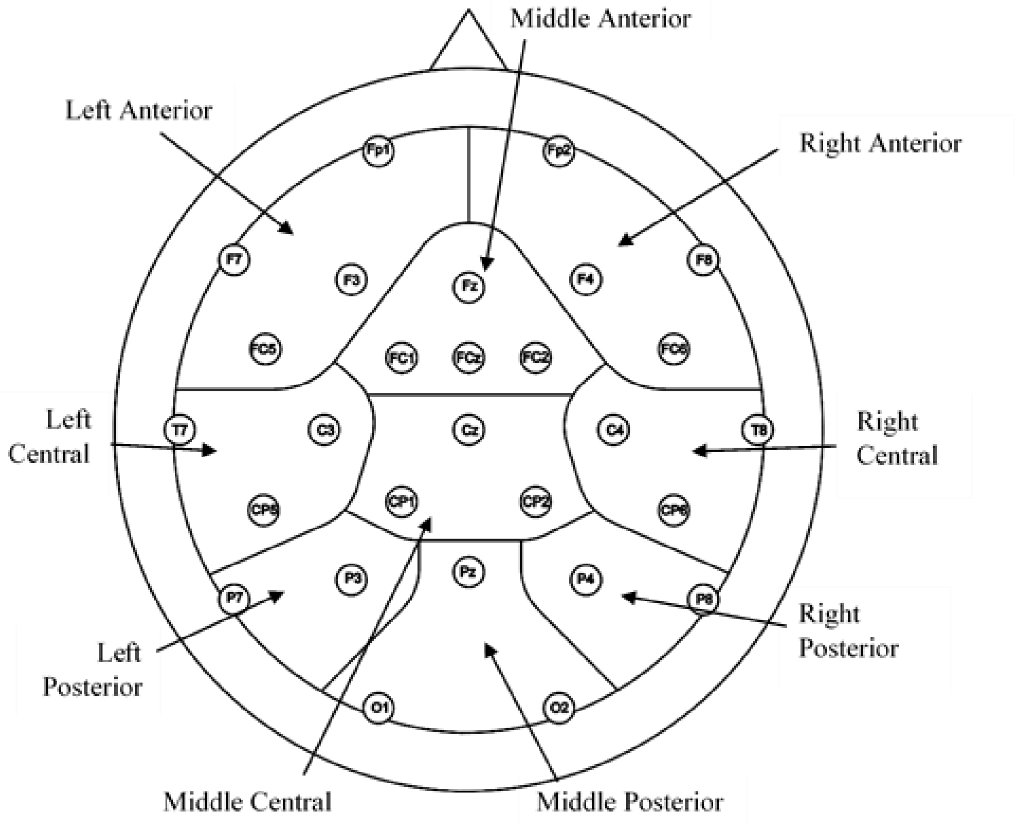


Figure 3.2. This is an illustration of the 28 electrodes that were chosen for analysis, divided into nine different regions of interest. Left Anterior (LA; consisting of electrodes Fp1, F7, F3, FC5); Middle Anterior (MA; consisting of electrodes Fz, FC1, FCz, FC2); Right Anterior (RA; consisting of electrodes Fp2, F8, F5, FC6); Left Central (LC; consisting of electrodes CP5, C3, T7); Middle Central (MC; consisting of electrodes CP1, Cz, CP2); Right Central (RC; consisting of electrodes CP6, C4, T8); Left Posterior (LP; consisting of electrodes P7, P3); Middle Posterior (MP; consisting of electrodes Pz, O1, O2); and Right Posterior (RP; consisting of electrodes P8, P4).

Multifactorial repeated measures omnibus ANOVAs were conducted separately at each time-window, using factors based on the experimental variables plus two scalp location factors to include all nine regions of interest and investigate topographical patterns of any experimental variable effects. One scalp variable was Anterior/Posterior (AP), with three b levels: Anterior, central, and posterior, and the other variable was Hemisphere (H), also with three levels: Left, central, and right. Mean amplitudes of ERPs in these regions of interests were used in the ANOVAs. Follow-up analyses was conducted only for omnibus interactions between one or more experimental factors and scalp factors in each time window. Effects only involving scalp variables are not theoretically meaningful to interpret, so were not followed up and are not reported. Greenhouse-Geisser correction was applied if the sphericity assumption was violated (based on Mauchly's test of sphericity). This analysis was conducted using SPSS.

Cluster-based permutation tests. In addition, a more data-driven approach was used to analyse the data, which considered all time-points and scalp electrodes without collapsing across regions and time-windows (which can induce some bias). Therefore, this approach considerably reduced the number of comparisons needed to test for condition effects while also applying thresholding to control for false positives (refer to Chapter 2 for further information). Data from all 28 scalp electrodes and all time-points split in two separate time-windows (0ms to 1750ms and 1750ms to 3500ms) were included and submitted to nonparametric cluster-based permutation tests. This approach was used for both ERP amplitudes and oscillatory power derived from time-frequency decomposition, wherein frequency (ranging from 4-30Hz) was added as a third dimension along with electrodes and time-points. This analysis was conducted using FieldTrip toolbox for Matlab (Oostenveld et al., 2011).

Planned comparisons. For both ANOVA and cluster-based permutation test analyses of both ERPs and oscillations, the main analysis was conducted to test for potential differences in neural activity related to thinking versus not-thinking of autobiographical memories (main effects of the think/no-think manipulation) and whether the think/no-think neural effect differed for morally wrong versus right autobiographical memories (interactions between think/no-think and the memory type manipulations). We applied this main 2x2 factorial analysis separately in the first half (Block 1) and second half of think/no-think trials (Block 2) to investigate if these effects were present or absent in each half of the task. If the interaction between these factors was significant, we tested the effects of the think/no-think manipulation separately for each memory type. For ANOVAs, follow up tests were conducted only for the specific time-windows and scalp factors where significant interactions were found. For cluster-based permutation tests, follow up tests included all scalp locations and timepoints within the whole longer time-window (0ms to 1750ms or 1750ms to 3500ms).

For the intrusions analysis, the ANOVA analysis included a condition factor with three levels (Successful Retrieval, Successful Suppression, Intrusions) which was analysed with the scalp location factors, in line with Hellerstedt et al.'s (2016) analysis strategy. If a significant effect of condition was found in a time-window/location, then three pairwise comparisons were conducted: i) Successful Retrieval vs. Successful Suppression; ii) Intrusions vs. Successful Retrieval, and iii) Intrusions vs. Successful Suppression. For the cluster-based permutation analyses of ERP and EEG oscillation power, the three pairwise comparisons were conducted directly.

3.2 Results

3.2.1 Behavioural results

First, the effect of moral memory type on retrieval and intrusions in the think/no-think task, compared across the first and second halves of the task is reported. Then, the change in autobiographical memory descriptions as an effect of the think/no-think task as quantified by experimenter ratings is reported. Finally, the change in participants' ratings of autobiographical memory characteristics as an effect of think/no-think task is described.

Memory retrieval and intrusions. Trial-by-trial intrusion scale responses during the think/no-think task were analysed for think/no-think and morally right/wrong conditions for the first versus second block of trials, to test if intrusion rates reduced over repeated attempts to suppress memories, as is typically found in the memory suppression literature. Based on previous research, the percentage of trials where retrieval occurred were computed by combining trials where participants reported thinking of the associated memory either "briefly" or "often" (Hellerstedt et al., 2016; Levy & Anderson, 2012).

The percentage of retrieval in the think condition relates to voluntary retrieval, whereas retrieval in the no-think condition is interpreted as memory intrusions. Therefore, the measures in these two conditions are conceptually different and it is more logical to conduct analyses separately for these two conditions. Therefore, a 2 (Memory Type) x 2 (Block) repeated measures ANOVA was conducted separately for the ratings in the think vs. no-think conditions.

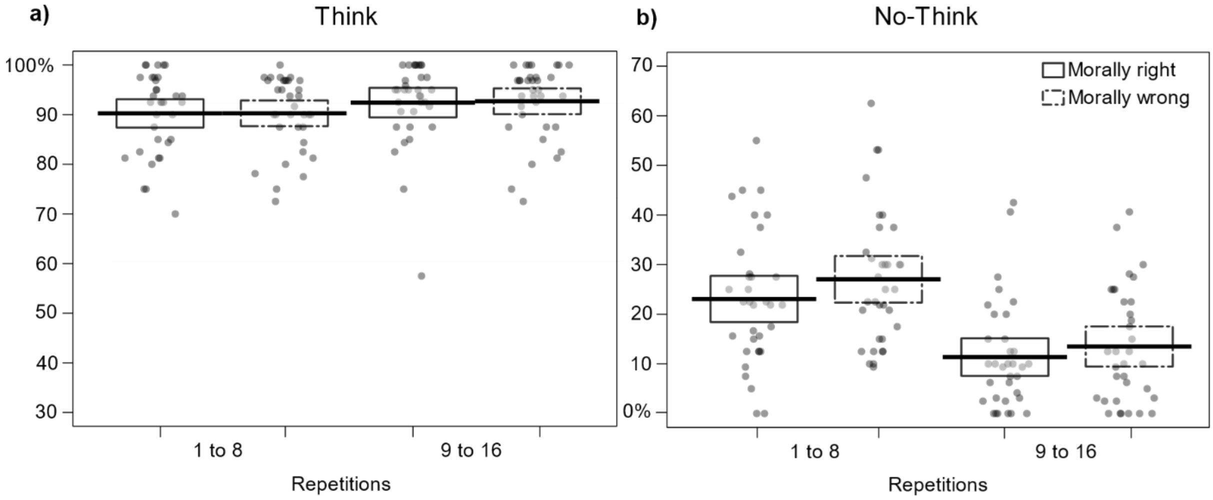


Figure 3.3. Percentage of trials with memory retrieval compared across memory type and block (repetitions) in the a) think condition, and b) no-think condition. The scatter dots show percentage of trials with retrieval of each individual. The thick lines show the group means and the boxes depict the 95% confidence interval of the group means.

Successful retrieval increases over repeated attempts at thinking. See Figure 3.3.a for an illustration of these results. The main effect of memory type was not significant, *F*(1,33) = .03, *p* = .87, partial η^2 = .001); and neither was the interaction between memory type and block, *F*(1,33) = .02, *p* = .89, partial η^2 = .001). However, there was a significant main effect of block: Percentage of trials with voluntary recall success was lower in the first half (*M* = 90.26%, *SD* = 6.86%) than in the second half (*M* = 92.56%, *SD* = 7.20%), *F*(1,33) = 4.52, *p* = .041, partial η^2 = .12, as expected.

Intrusions reduce over repeated attempts of suppression and morally wrong memories were overall more intrusive than morally right memories. See Figure 3.3.b for an illustration of these results**.** There were significant main effects of both memory type and block: Percentage of trials with memory intrusions were significantly higher for morally wrong (*M* = 20.28%, *SD* = 11.43%) than for morally right memories (*M* = 17.22%, *SD* = 11.47%), *F*(1,33) = 4.62, *p* = .037, partial η^2 = .123; memory intrusions were also significantly higher during the first (*M* = 25.06%, *SD* = 12.48%) than the second half (*M* = 12.44%, *SD* = 10.34%) of the task, *F*(1,33) = 78.51, *p*<.001, partial η^2 = .704. However, there was no interaction between the two factors *F*(1,33) = 1.56, *p* = .22, partial η^2 = .045.

Autobiographical memory description analysis. A 2 (Instruction Type) x 2 (Memory Type) repeated measures ANOVA was first conducted on the ratings of how similar the memory descriptions were that participants provided before and after the think/no-think task (see Table 3.1). See tables A.1. in Appendix A.3 for average ratings of similarity provided by the independent coders.

Table 3.1. Average similarity ratings across instruction and memory types.

	Experimenter ratings	
	Right	Wrong
Think	4.07(.43)	3.84(.54)
No-Think	4.05(.52)	3.65(.51)

Note. Similarity was rated on a 5-point scale (1 = Not at all similar, 5 = Extremely similar). Standard deviation is reported in brackets.

Although memory descriptions in the no-think condition were rated as overall less similar than descriptions in the think condition, this difference was not significant¸ *F*(1,33) = 2.17, *p* = .15, partial η^2 = .065. There was however a significant main effect of memory type, as morally wrong memories (*M* =3.75, *SD* = .48) were rated as overall less similar across the two times than morally right memories (*M* = 4.06, *SD* = .37); *F*(1,33) = 18.27, *p* < .001, partial η^2 = .356. The interaction between instruction type and memory type was not significant; *F*(1,33) = 2.14, *p* = .15, partial η^2 = .06.

Next, a 2 (Instruction Type) x 2 (Memory Type) x 2 (Administration time: Before TNT vs After TNT) repeated-measures ANOVA was conducted on the rating of specificity and valence of the memory descriptions (Table 3.2). See table A.2. in the Appendix A.3 for average ratings of similarity provided by the independent coders.

Table 3.2. Average ratings of specificity and valence across instruction and memory types. Ratings were provided by the experimenter.

	Right		Wrong	
	Pre-TNT	Post-TNT	Pre-TNT	Post-TNT
		Specificity		
Think	2.93(.64)	2.93(.67)	2.95(.63)	2.94(.65)
No-Think	2.96(.59)	2.91(.69)	2.93(.72)	2.95(.68)
		Valence		
Think	4.21(.35)	4.20(.46)	3.35(.51)	3.84(.46)
No-Think	4.15(.43)	4.26(.48)	3.35(.59)	3.35(.57)

Note. Specificity was rated on a 5-point scale (1=Not at all specific to 5 = Extremely specific), emotional valence was measured on a 7-point scale (1=Very Negative, 4 = Not negative or positive, 7 = Very Positive). Standard deviation is reported in brackets.

Unsurprisingly, descriptions of morally right memories were rated as containing more positively valanced words than morally wrong memories $F(1,33) = 42.01$, $p < .001$, partial $\eta^2 = .56$. For both measures, neither the main effects of administration time, nor its interaction with instruction type and memory type were significant (all $Fs < 2.34$, $ps > .14$).

Thus, both self-reports and experimenter's ratings showed that morally wrong memories changed between the first and second day in terms of how they were experienced and described by participants, but no such changes occurred for the morally right memories, and the changes were not influenced by the think/no-think manipulation.

Phenomenological characteristics of autobiographical memories. The following measures were used for further analyses: a) memory age (measured in years), vividness (ranging from 1 to 5), intentionality (ranging from 1 to 5), and morality (ranging from 1, *morally wrong* to 7, *morally right*); b) from the I-PANAS-SF, two morally relevant emotions were chosen (guilt, shame; ranging from 1 to 5); and c) both pleasure and arousal were chosen from the SAM measures (ranging from 1 to 5). These measures were chosen from the later battery of self-report ratings because of a priori predictions that these characteristics may change as a function of memory suppression, if suppression inhibits unwanted memories.

In an initial supplementary analysis (not reported here, see Appendix A.4), differences in characteristics between morally wrong and right memories before the think/no-think task were analysed. These results verified that memories were very similar across the think/no-think conditions at the time they were assigned to these conditions (i.e., that the random assignment had succeeded). The results also replicated previous findings that morally wrong memories are rated to have occurred in the more distant past (Escobedo & Adolphs, 2010), and are less vivid (Kouchaki & Gino, 2016) than morally right memories. Moreover, to extend on these findings, participants reported their actions as less intentional in morally wrong than morally right memories.

For the key analyses, a difference score was computed between ratings reported in the first session and the second session, and this was used as a dependent variable to test the effect of the TNT manipulation on phenomenological characteristics of memories (see Figure 3.4).

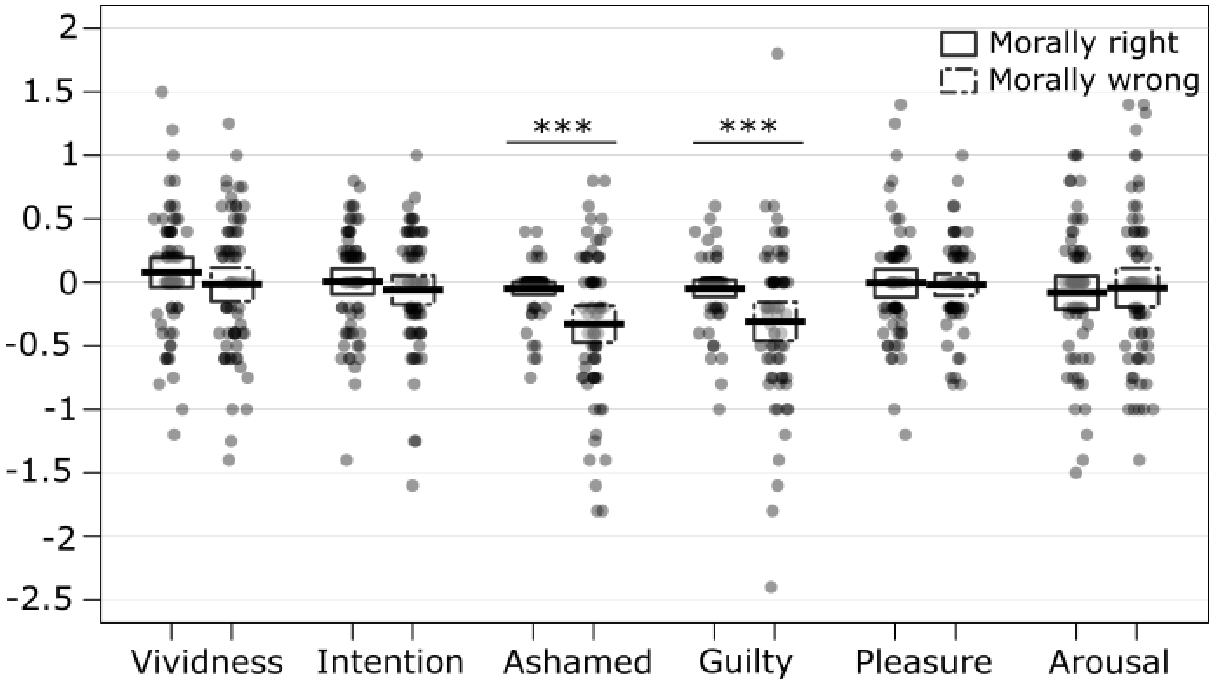


Figure 3.4. Average of difference score (before TNT – after TNT) compared across memory type. The scatter dots show difference scores of each individual for each. The thick lines show the group means and the boxes depict the 95% confidence interval of the group means. ****p*<.001.

A 2 (Instruction Type: Think, no-think) x 2 (Memory Type: Morally right, morally wrong) repeated measures ANOVA was conducted to test the effect of the TNT task on the difference score of the chosen self-report measures. See Table 3.3 for the results from the ANOVAs. Neither the main effects of instruction type, nor instruction type x memory type interactions were significant for any measure. However, there were significant main effects of memory type for some measures: Participants reported feeling less ashamed and guilty for morally wrong memories after the TNT task, but there were no such changes in guilt and shame for morally right memories.

Table 3.3. ANOVA results from omnibus test with the difference score (After TNT – Before TNT).

	Phenomenological memory characteristics											
	Memory Age			Vividness			Intention			Morality		
	F	p	η^2_p	F	p	η^2_p	F	p	η^2_p	F	p	η^2_p
IT (T vs. NT)	0.01	0.91	0.00	0.00	0.95	0.00	0.05	0.82	0.00	0.47	0.50	0.01
MT (MW vs. MR)	0.44	0.51	0.01	2.07	0.16	0.06	0.51	0.48	0.02	0.01	0.91	0.00
ITxMT	0.61	0.44	0.02	0.68	0.42	0.02	0.49	0.49	0.02	1.51	0.23	0.04

	I-PANAS-SF						*SAM*					
	Ashamed			Guilty			Pleasure			Arousal		
	F	p	η^2_p	F	p	η^2_p	F	p	η^2_p	F	p	η^2_p
IT (T vs. NT)	0.31	0.58	0.01	0.19	0.66	0.01	0.15	0.70	0.01	2.35	0.14	0.07
MT (MW vs. MR)	**12.42**	**0.001**	**0.27**	**8.93**	**0.005**	**0.21**	0.03	0.88	0.00	0.30	0.59	0.01
ITxMT	0.69	0.41	0.02	2.23	0.14	0.06	0.00	0.96	0.00	0.36	0.55	0.01

Note. IT = Instruction Type, MT = Memory Type, T = Think, NT = No-Think, MW = Morally Wrong, MR = Morally Right. η^2_p = partial eta sq. (effect size). I-PANAS-SF = International – Positive and Negative Affect Scale – Short Form, SAM = Self-Assesssment Manikin. I-PANAS-SF = International – Positive and Negative Affect Scale – Short Form. Significant effects ($p < .05$) are shown in bold. *N = 34*.

3.2.2 ERP results

The results from the data driven cluster-based analyses are presented in detail here because they control for false-positives and are therefore more reliable than uncorrected multifactorial ANOVAs (see Chapter 2 for more information). Detailed results from the multifactorial ANOVAs are presented in the Appendix (see Appendix A.5.). Overall, the cluster results were significant across longer time periods and were more broadly distributed across the scalp than the ANOVA results, because of the laxer statistical threshold. The main divergence between cluster-based tests and ANOVAs were found in the intrusion analysis, where ANOVAs showed additional intrusion-related ERP effects that were in line with previous research (see Appendix A.5.). The results from these two analyses are compared in Tables 3.4, 3.5, 3.6. The analyses were conducted separately for the first and second blocks and tested for main effects and interactions between the experimental factors in the 2 (Instruction Type; Think, no-think) x 2 (Memory Type; Morally right, morally wrong) design across all scalp electrodes and timepoints with a cluster-based threshold to control for multiple comparisons (see methods section). Results from the first half (block 1) are reported first, followed by the second half (block 2). Then, results from the intrusion analysis are reported.

Block 1. Grand-average ERPs from six different electrode sites (F3, F4, P3, P4, P7, P8), showing all four conditions are presented in Figure 3.5. Topographical maps that illustrate the raw difference in amplitudes between conditions, significant ANOVA effects and associated effect sizes for effects that involve Instruction Type and/or Memory Type as factors, and t-values from significant clusters are presented in Figure 3.6.

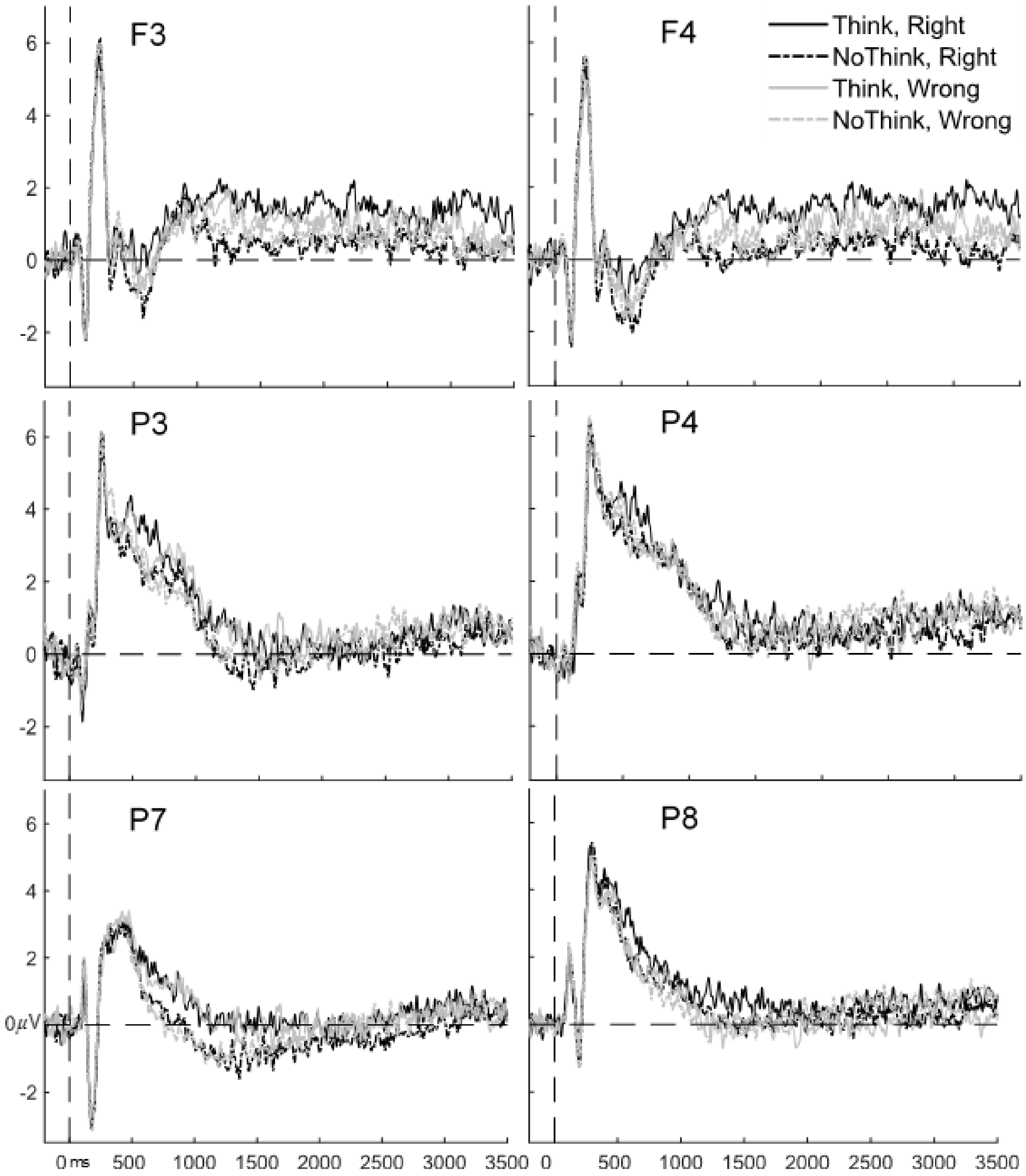


Figure 3.5. Grand-average ERPs for the four experimental conditions in the first half of the TNT task from six electrode positions (F3, F4, P3, P4, P7, and P8).

Overall, no significant clusters were found for right vs. wrong memory type comparison, but importantly, there were significant clusters showing differences in ERPs between think vs. no-think conditions and the interaction between instruction type and memory type revealed significant clusters, as described below.

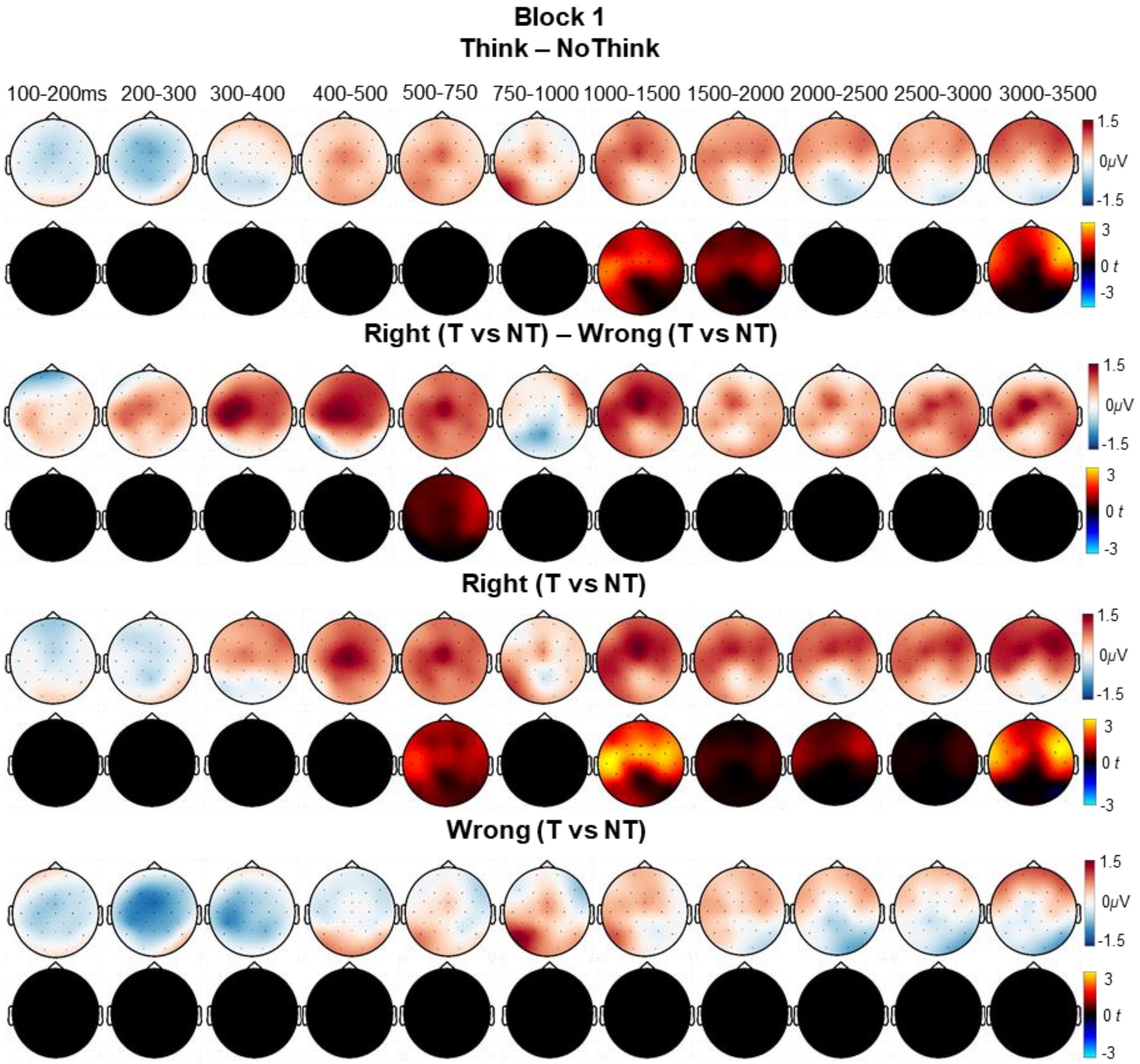


Figure 3.6. Results from follow-up pairwise comparisons in the first block of the TNT task cluster-based permutation testing (bottom row). For each comparison, images on top are topographical maps of ERP amplitude differences between conditions (blue/white/red colourmap, in μV) and images in the bottom row illustrates t-values (blue/black/red colourmap, in t) for the differences. The colour scale represents magnitude and direction of the effect.

General think vs. no-think effects. Four significant positive clusters from around 1080 to 1360ms (*p* <.001), 1500 to 1670ms (*p* = .005), 2990 to 3160ms (*p* = .005), and 3210 to 3390ms (*p* = .015) were found, indicating more positive sustained ERP amplitudes during retrieval compared to suppression. These effects were present across left parietal (in the 1080-1360ms cluster), frontal, and central regions (in all clusters) as can be seen from the *t*-value distributions in Figure 3. Thus, consistent with our predictions, suppression attempts reduced parietal ERP positivities and such effects were relatively late and prolonged and had a broad scalp topography.

Interaction between moral memory type and instruction type. The think/no-think x memory type interaction analysis showed a significant positive cluster early on between 550-700ms (p = .023), but no significant clusters were found in later time points. Follow up analyses showed that this interaction was driven by a significant positive cluster for morally right memories around the same time as the interaction cluster (550-700ms, p = .009), whereas no significant clusters were found for morally wrong memories. Positive think vs. no- think clusters were also present for morally right memories from around 1080 to 1580ms (p < .001), 2150 to 2340ms (p = .017), and 2940ms to 3400 (p < .001) in frontal and central regions, whereas there were no such significant effects for morally wrong memories (see Fig. 3)**.**

Table 3.4. Comparison of two analyses strategies for different ERP components from block 1

	Block 1	
	Cluster-permutation tests	Multifactorial ANOVA
P2 (~100-300ms)	×	✓
Early Positivity (~300-750ms)	✓	✓
Left Parietal Positivity (~500-1500ms)	✓	✓
Positive slow wave (~1000-3000ms)	✓	✓

Note. A tick mark indicates that the component was statistically significant in that analysis, whereas a cross indicates that the component was not statistically significant. Please note that the results are summarised as components here only for illustrative purposes, see text and discussion for more detail.

Block 2. Topographical maps that illustrate the raw difference in amplitudes between conditions, significant effects and effect sizes that involve instruction type and/or memory type as factors, and t-values from significant clusters are illustrated in Figure 3.8.

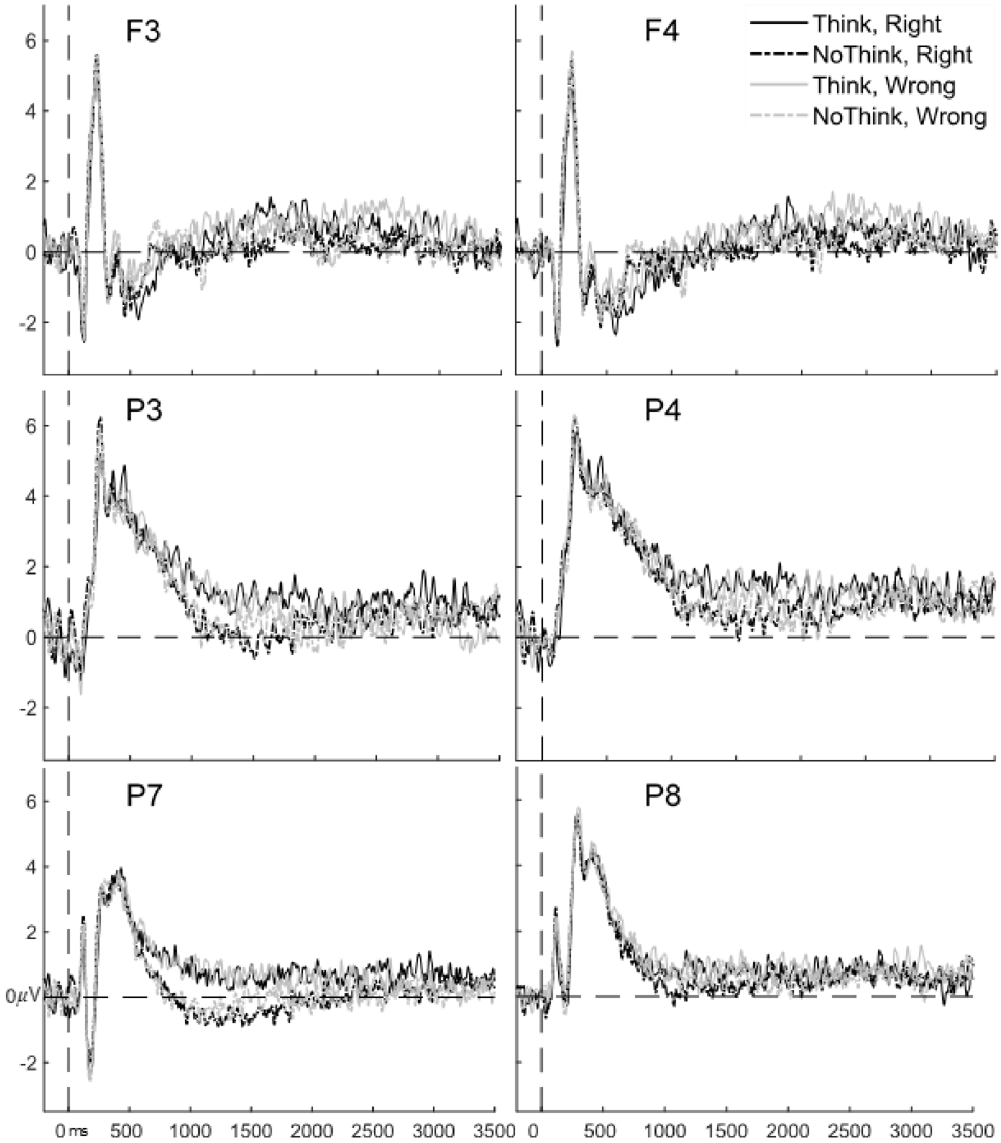


Figure 3.7. Grand-average ERPs for the four experimental conditions during the second half of the TNT task ERPs from six electrode positions (F3, F4, P3, P4, P7, and P8).

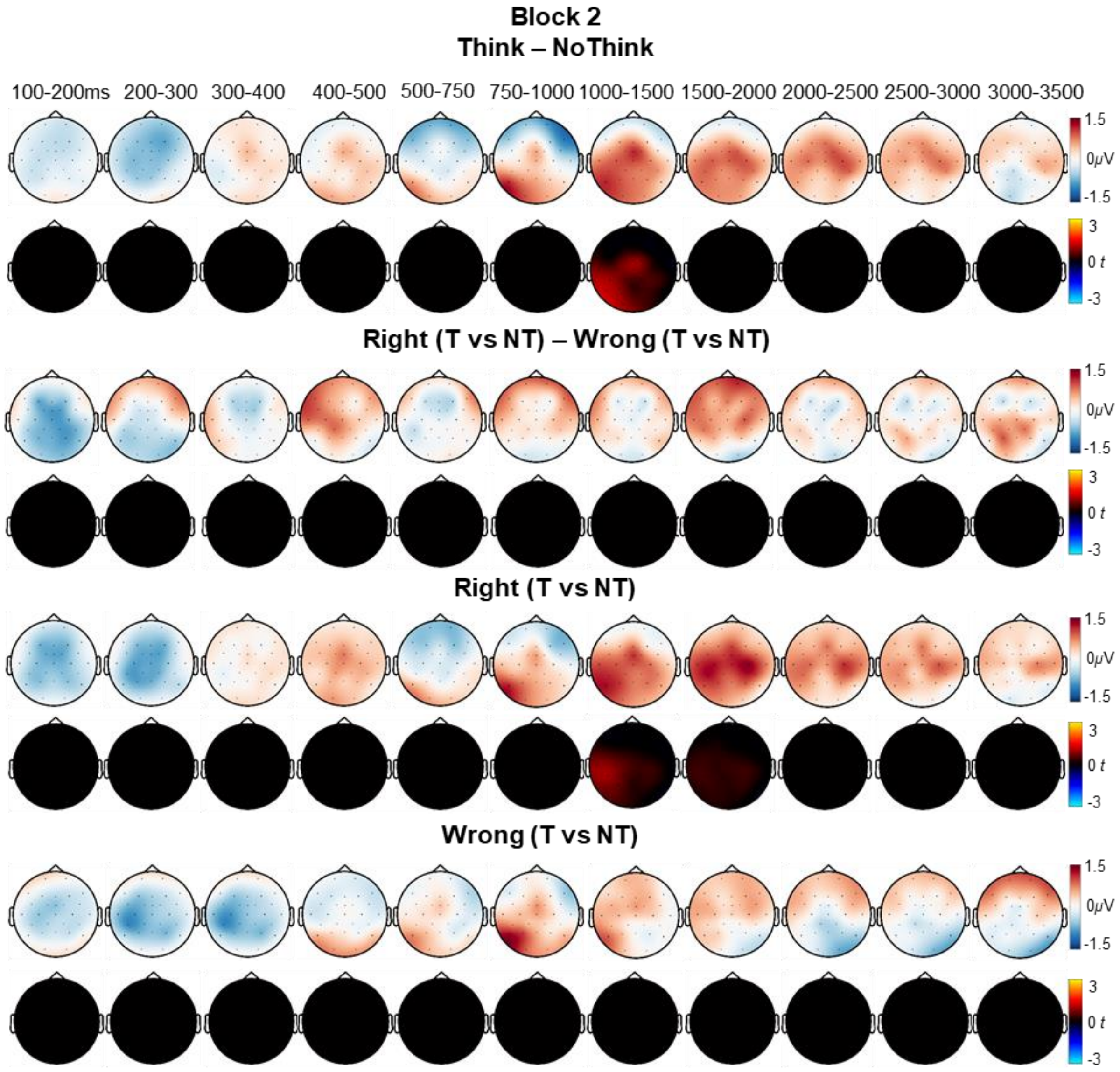


Figure 3.8. Results from follow-up pairwise comparisons in the second block of the TNT task cluster-based permutation testing (bottom row). For each comparison, images on top are topographical maps of ERP amplitude differences between conditions (blue/white/red colourmap, in μV) and images in the bottom row illustrates t-values (blue/black/red colourmap, in t) for the differences. The colour scale represents magnitude and direction of the effect.

In the second half of trials, only one significant think vs. no-think positive cluster was found from 1000-1230ms (p = .013), across left-posterior and central regions (Fig. 3). This analysis therefore showed that suppression of autobiographical memories reduced late left parietal positivities also in the second half of trials, in line with our predictions. There were however no significant ERP differences between right vs. wrong memory types, nor any significant clusters for the memory type x think/no-think instruction type interaction in the second half, suggesting that the neurocognitive processes engaged during retrieval and suppression were similar for morally right and wrong memories in the second half. See Table 3.5 for a summary and comparison of results from cluster-based permutation tests and multifactorial ANOVAs.

Table 3.5. Comparison of two analyses strategies for different ERP components for block 2

Block 2		
	Cluster-permutation tests	Multifactorial ANOVA
P2 (~100-300ms)	×	✓
Left Parietal Positivity (~500-1500ms)	✓	✓
Positive slow wave (~1000-3000ms)	×	✓

Note. A tick mark indicates that the component was statistically significant in that analysis, whereas a cross indicates that the component was not statistically significant. Please note that the results are summarised as components here only for illustrative purposes, see discussion for more detail.

Intrusion analysis. Participants with low trial numbers in the intrusions condition were excluded. Thus, 26 participants qualified the trial number cut-off (more than 13 intrusion trials) in this analysis. Because splitting trials into two blocks was not possible due to low trial numbers, ERPs were collapsed across blocks. Grand-average ERPs from six different electrode sites (F3, F4, P3, P4, P7, and P8) are presented in Figure 3.9. Topographical maps that illustrate the raw difference in amplitudes between conditions, significant effects and effect sizes that involve the factor condition, and t-values from significant clusters are illustrated in Figure 3.10.

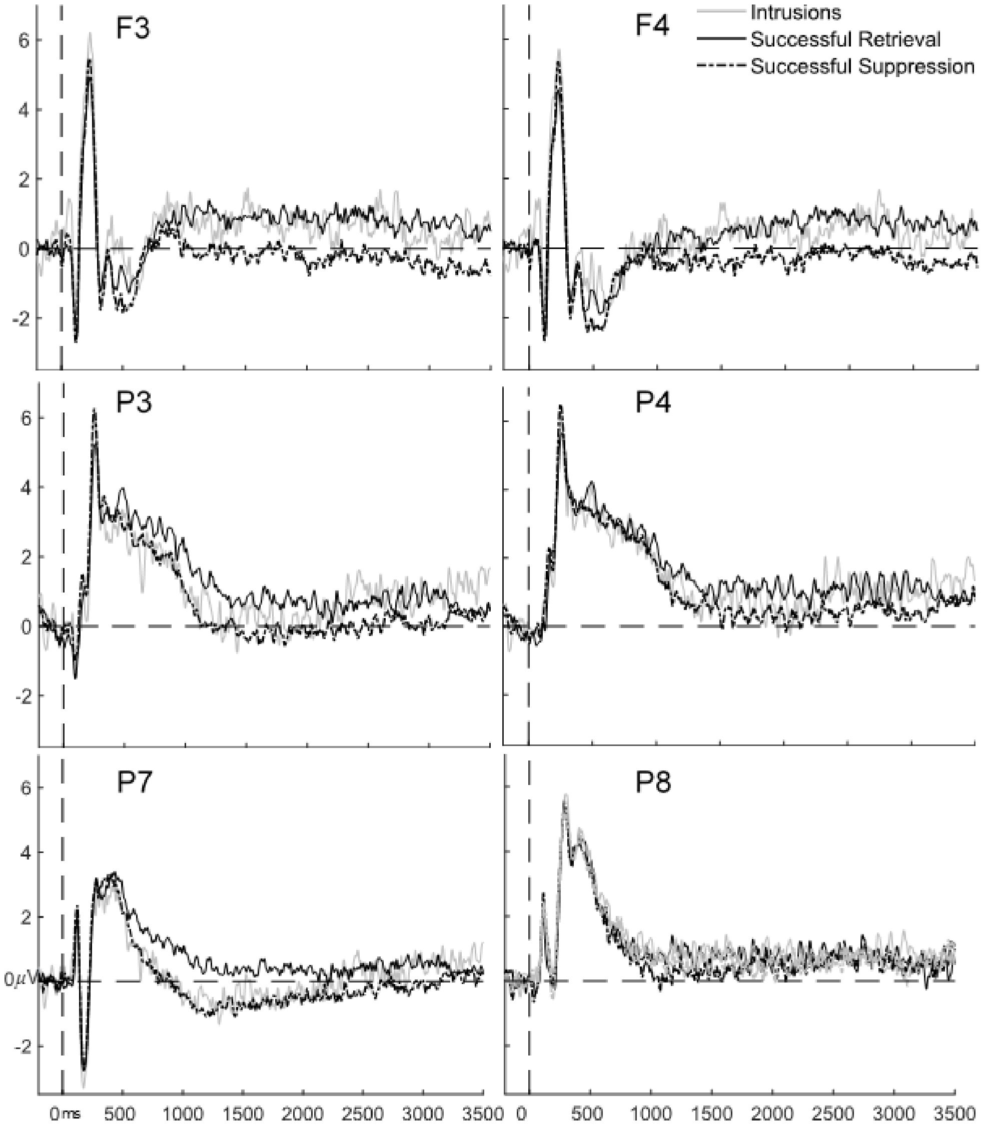


Figure 3.9. Grand-average ERPs for the three experimental conditions in the intrusion analysis of the TNT task ERPs from six electrode positions (F3, F4, P3, P4, P7, and P8).

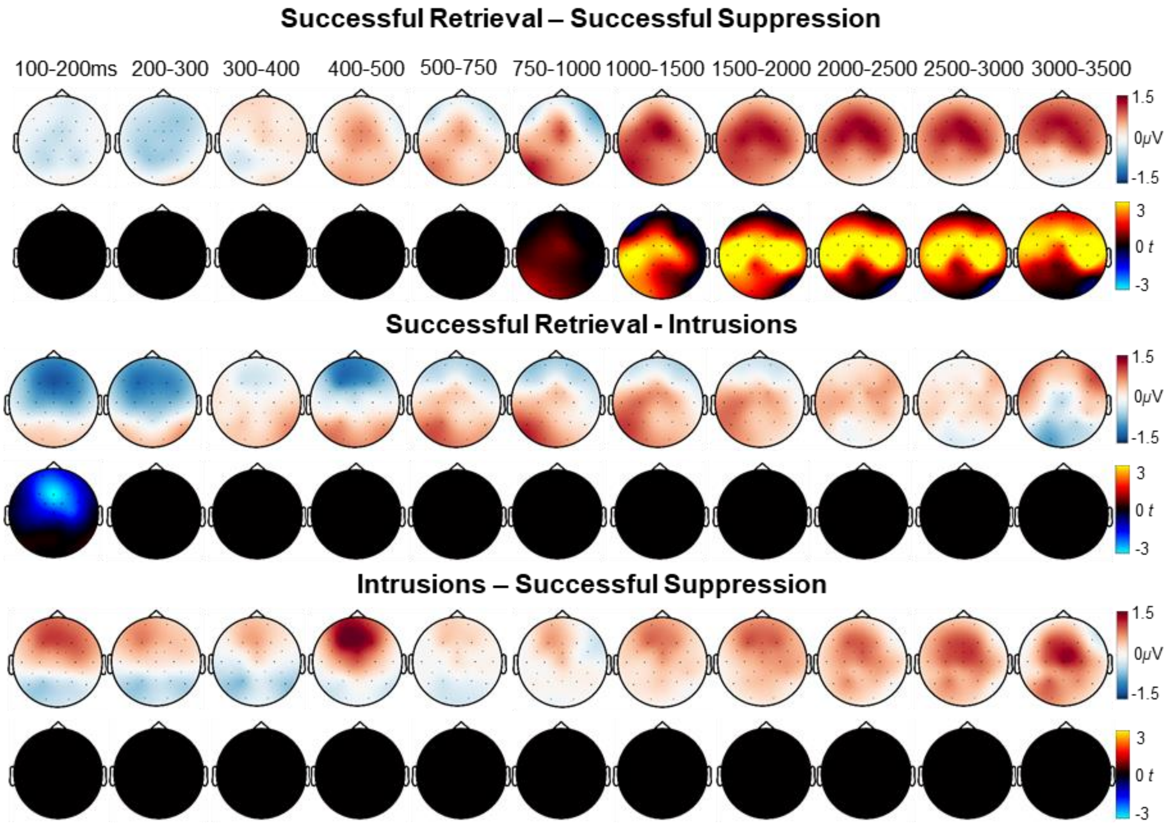


Figure 3.10. Results from pairwise comparisons in the complementary intrusions analysis using cluster-based permutation testing. For each comparison, images on top are topographical maps of ERP amplitude differences between conditions (blue/white/red colourmap, in μV) and images in the bottom row illustrates t-values (blue/black/red colourmap, in t) for the differences.

A sustained ERP positivity was found for successful retrieval vs. successful suppression, beginning around 950ms and lasting until the end of the epoch (3500ms, see Fig. 4). This was reflected by significant positive clusters from 950 to 1340ms (p = .001), 1350 to 1580ms (p = .001), and 1750 to 3500ms (p < .001). This effect was spread across left posterior regions in the early clusters, and in central and anterior regions across all clusters. When comparing successful retrieval vs. intrusions, one significant negative cluster was found from 50 to 200ms (p = .01), caused by more positive ERPs for intrusions than retrieval trials across anterior regions. There were no significant clusters when comparing intrusions and successful suppression. There were significant effects for intrusion analyses in the multifactorial ANOVA that did not survive cluster permutation tests, please see Table 3.6 for a comparison, and refer to Appendix A.5. for more detail.

Table 3.6. Comparison of two analyses strategies for different ERP components from the intrusions analysis

Intrusions Analysis		
	Cluster-permutation tests	Multifactorial ANOVA
	Successful Retrieval vs. Successful Suppression	
P2 (~100-300ms)	×	✓
FN400 (~400-500ms)	×	×
Left Parietal Positivity (~500-1500ms)	✓	✓
Positive slow wave (~1000-3000ms)	✓	✓
	Successful Retrieval vs. Intrusions	
P2 (~100-300ms)	✓	✓
FN400 (~400-500ms)	×	×
Left Parietal Positivity (~500-1500ms)	×	✓
Positive slow wave (~1000-3000ms)	×	×
	Intrusions vs. Successful Suppression	
P2 (~100-300ms)	×	✓
FN400 (~400-500ms)	×	✓
Left Parietal Positivity (~500-1500ms)	×	×
Positive slow wave (~1000-3000ms)	×	✓

Note. A tick mark indicates that the component was statistically significant in that analysis, whereas a cross indicates that the component was not statistically significant. Please note that the results are summarised as components here only for illustrative purposes, see discussion for more detail.

3.2.3 EEG oscillation results

Like in the ERP analysis, cluster analysis of EEG oscillation power differences between conditions was conducted separately for the first and second blocks and tested for main effects and interactions between the experimental factors in the 2 (Instruction Type; Think, no-think) x 2 (Memory Type; Morally right, morally wrong) design across all scalp electrodes and timepoints with a cluster-based threshold to control for multiple comparisons (see methods section).

Block 1. The time-frequency plots representing raw dB differences from four electrode positions (F3, F4, P3, P4) and t-values from significant clusters are illustrated in Figure 3.11. The topographical maps of power differences and significant clusters are presented in Figure 3.12. There were no significant clusters when comparing morally right and wrong memories, either collapsed or when separated between the think and no-think conditions. The interaction contrast comparing think vs. no-think oscillation differences between right and wrong memories also did not reveal any significant clusters[2].

There was however a general think vs. no-think synchronisation effect across the whole frequency band (4 – 30Hz) starting from around 330ms and ending at 3500ms ($p < .001$), maximal in parietal regions. When investigating this difference for morally right memories only, a large think vs. no-think synchronisation was present across the whole frequency band (4 – 30 Hz) from 300ms to 3500ms (both *ps* < .001), maximal in parietal regions. This synchronisation effect was also present for morally wrong memories, but the significant cluster started later in time, from around 950 to 1750ms and ranging across 6 to 30Hz frequencies ($p < .001$), whereas from 1750 to 3500ms, the synchronisation effect was across the whole frequency range (4 – 30Hz). This effect was also maximal in parietal regions.

---

[2] When only the theta band (4-7Hz) is analysed, three significant positive clusters were found for the interaction between think vs. no-think of morally right and think vs. no-think of morally wrong memories. Significant positive clusters in posterior regions from around 500 to 1000ms ($p = .01$), 1000 to 1300ms ($p = .01$), and 1960 to 2500ms ($p = .023$). Morally right think vs. no-think comparison revealed these positive theta clusters, but there were no significant theta clusters for morally wrong memories in any time region.

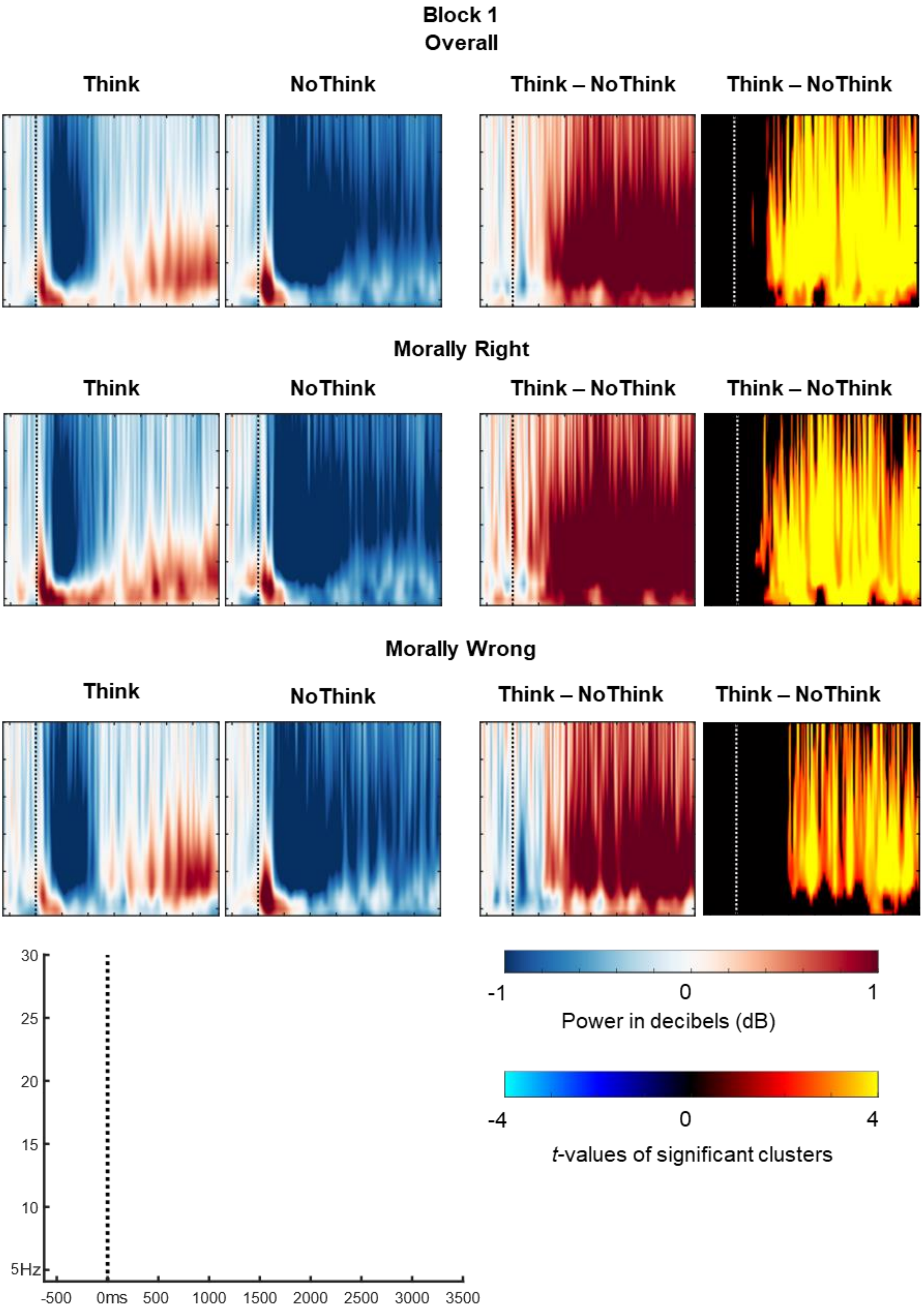


Figure 3.11. Time-Frequency plots of think vs. no-think and its interaction with memory type from P4 electrode. Raw power differences are illustrated in the blue/white/red colourmap) and t-values for the differences are represented in the cold/black/hot colour map. Only significant clusters (*p* < .025 at the cluster level) are shown in the t-statistic.

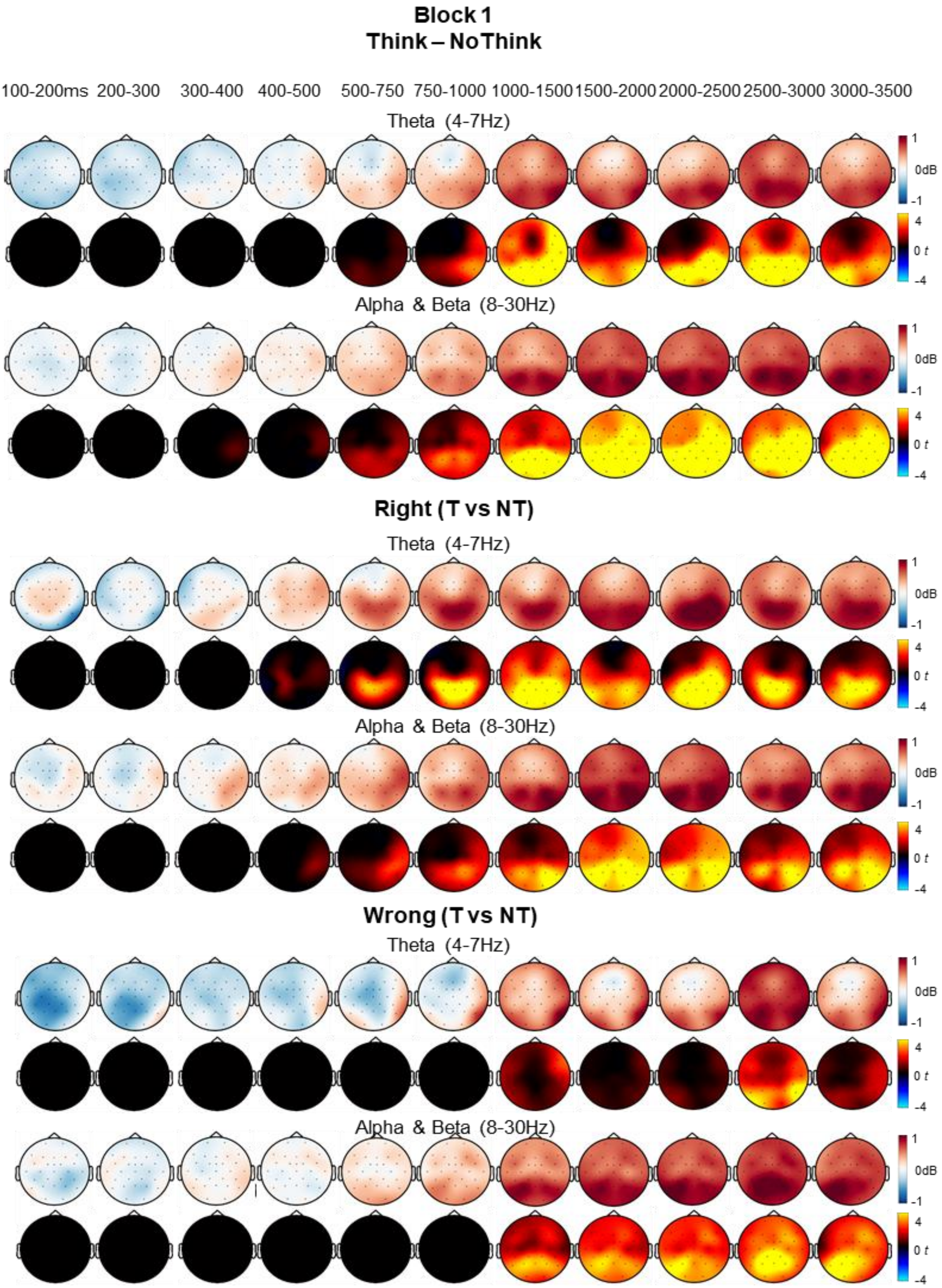


Figure 3.12. Data driven analysis of time-frequency decomposed data in the first block of the think/no-think task. Topographical maps of amplitude differences (top rows, blue/white/red colourmap) and dB-values for the differences (bottom rows, cold/black/hot colour map) for think – no-think conditions for each memory type and the interaction contrast. Note that the theta effects in the interaction contrast are significant when analysing only the theta band oscillations in the cluster permutation test, but these clusters are not statistically significant when all frequency bands are added in the analysis.

Block 2. The time-frequency plots from four electrode positions in block 2 is illustrated in Figure 3.13. The topographical maps of power differences and significant clusters are illustrated in Figure 3.14. The think vs no-think synchronisation found in the first block starts earlier in block 2: The general think - no-think cluster was significant across the whole frequency range (4 – 30Hz) and from around 440 to 3500ms ($p < .001$). For morally right memories, it is present across the whole frequency band (4 – 30Hz) and is significant from around 90ms to 3500ms (both *p*s < .001). Similarly, for wrong memories, it ranges across the whole frequency band (4 – 30Hz) but starts around 600ms and ends at 3500ms (both *ps* < .001). Similar to block 1, the interaction contrast comparing think vs. no-think differences between moral memories was not significant[3], and there were no differences between morally right vs. wrong memories in the main effects either.

---

[3] There were no significant clusters in the instruction type x memory type interaction contrast analysis when only the theta band was analysed in this block.

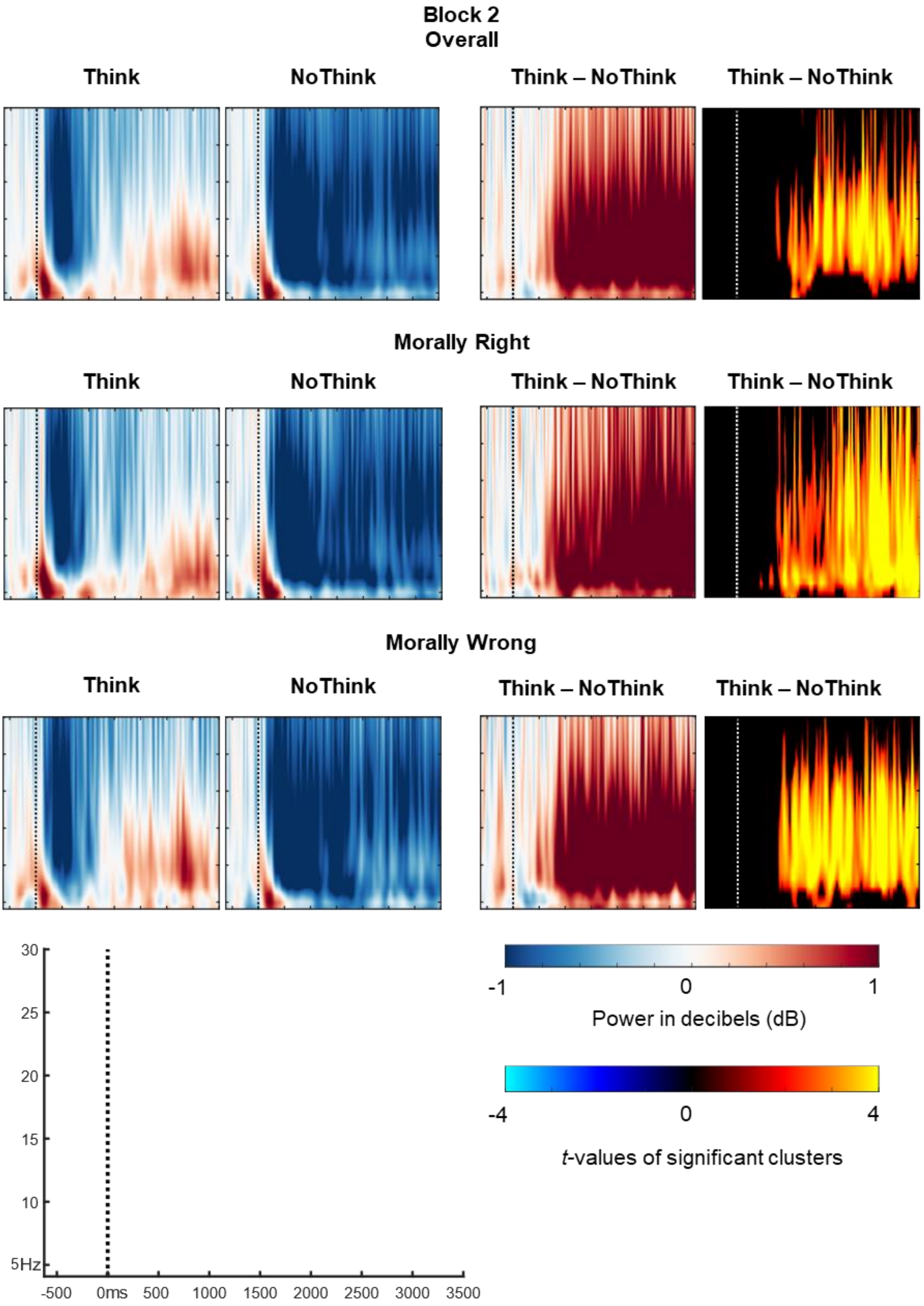


Figure 3.13. Time-Frequency plots of think – no-think differences in power (left image, blue/white/red colourmap) and t-values for the differences (right image, cold/black/hot colour map) from P4 electrode. Only significant clusters ($p < .025$ at the cluster level) are shown in the t-statistic time-frequency plots. Note that the pre-stimulus interval was not analysed statistically.

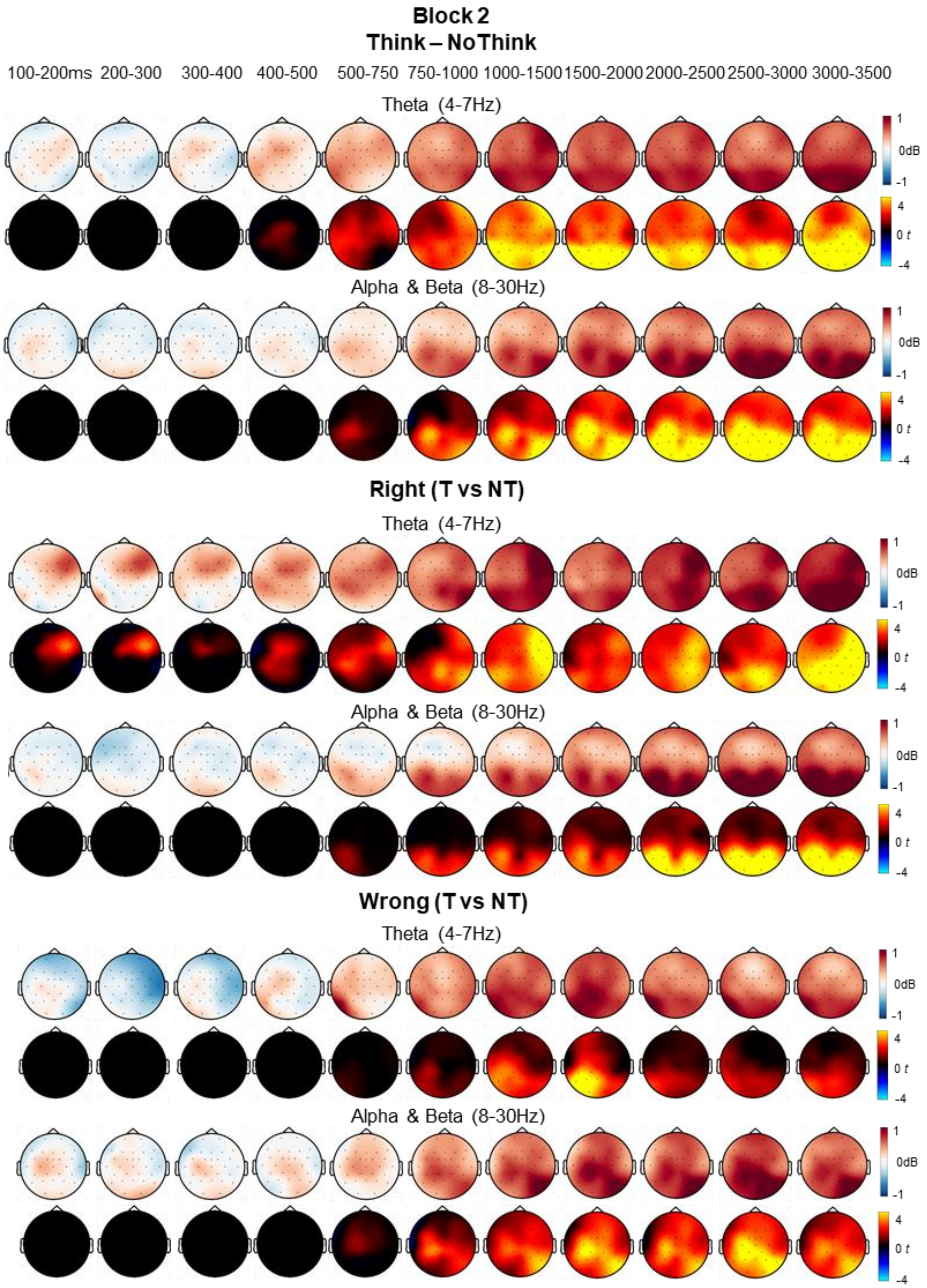


Figure 3.14. Data driven analysis of time-frequency decomposed data in the second block of the think/no-think task. Topographical maps of amplitude differences (top rows, blue/white/red colourmap) and dB-values for the differences (bottom rows, cold/black/hot colour map) for think – no-think conditions for each memory type and the interaction contrast.

Intrusion Analysis. The time-frequency plots from four electrode positions in intrusion analysis is illustrated in Figure 3.15. The topographical maps of power differences and significant clusters are illustrated in Figure 3.16. A synchronisation effect across the whole frequency range (4 – 30Hz) starting around 400ms to 3500ms (both *ps* <.001) was found when comparing successful retrieval with successful suppression, maximal in parietal regions. Similarly, there was a synchronisation effect across the whole frequency range (4-30Hz) and starting around 650ms to 3500ms (both *ps* < .001) for successful retrieval - intrusions. When comparing Intrusions and Successful Suppression, a significant synchronisation was found primarily in the alpha and beta bands (7-30Hz), late in the epoch, around 2800ms to 3500ms (*p* = .008). There were no additional significant clusters when only the theta band was analysed.

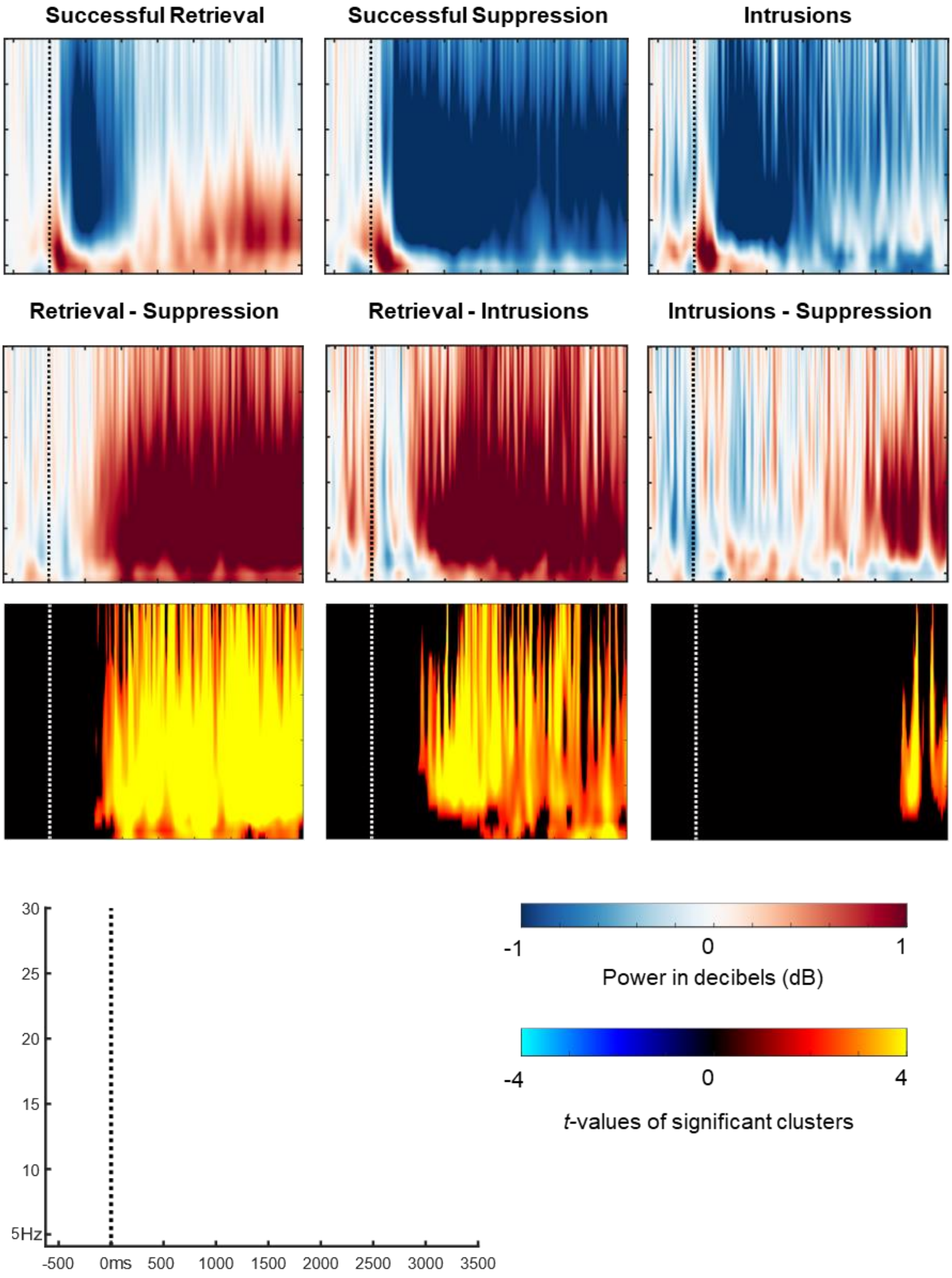


Figure 3.15. Time-Frequency plots of think – no-think differences in power (left image, blue/white/red colourmap) and t-values for the differences (right image, cold/black/hot colour map) from P4 electrode. Only significant clusters (*p* < .025 at the cluster level) are shown in the t-statistic time-frequency plots. Note that the pre-stimulus interval was not analysed statistically.

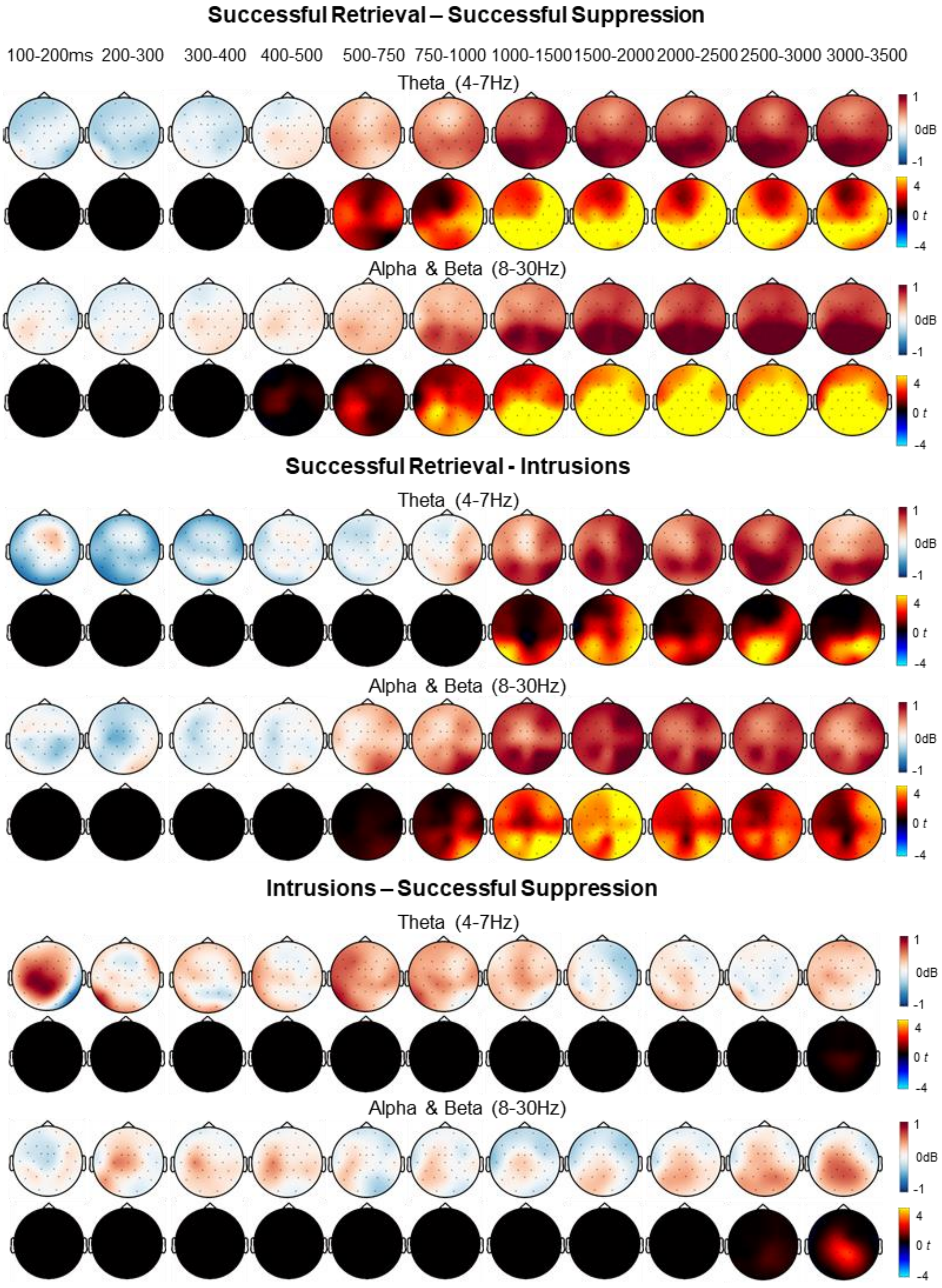


Figure 3.16. Data driven analysis of time-frequency decomposed data of the intrusion analysis. Topographical maps of amplitude differences (top rows, blue/white/red colourmap) and dB-values for the differences (bottom rows, cold/black/hot colour map) for think – no-think conditions for each memory type and the interaction contrast.

3.3 Discussion

The primary aims of this study were to 1) elucidate the neurocognitive mechanisms of autobiographical memory retrieval, suppression, and intrusions; and 2) investigate how the moral nature of memories affects our ability to control autobiographical memories, measured by how the moral nature of the memory affects the neurocognitive and behavioural processes of suppression. A rich body of literature has provided an understanding of the cognitive mechanisms and brain systems involved in suppressing relatively simple memories (see Anderson & Hanslmayr, 2014; Anderson & Hulbert, 2021). Research on autobiographical memory suppression and intrusions is scarce however, with a few studies investigating suppression (Noreen & Macleod, 2013; Noreen & MacLeod, 2014; Noreen et al., 2016; Stephens et al., 2013), but no studies on autobiographical intrusions. The present study thus aims to directly fill this gap in memory control research. I specifically investigated memories of past immoral/moral actions, because some evidence from social psychology indicates that such memories may be more difficult to control/suppress (Stanley, Henne, et al., 2017), and thus more susceptible to change (Kouchaki & Gino, 2016), in line with theories of motivated forgetting (Anderson & Hanslmayr, 2014; Stanley & De Brigard, 2019). The results provide behavioural and neural evidence that autobiographical memories can be suppressed, and furthermore, that repeated attempts at suppression reduces unwanted memory intrusions frequency. Furthermore, morally wrong memories were found to be particularly intrusive compared to morally right memories, indicating that our ability to control such memories may be relatively more challenging. The study does not provide clear evidence for the consequences of suppression on the characteristics and emotional responses to morally wrong memories, but it does indicate that morally wrong memories in general are more susceptible to change than morally right memories, regardless of control or retrieval.

This is the first study to investigate the phenomenological ratings of intrusions of autobiographical memories in the think/no-think task. Therefore, it is vital to understand how the present results compares to intrusion rates from suppressing simpler memories. The extent of intrusions experienced by participants in this study were overall lower than in previous TNT paradigms with simpler memories (e.g., Levy & Anderson, 2012). This may be due to differences in the strength of association between the cues and memories. In traditional TNT paradigms, participants first learn associations between word pairs, and only overly learnt word (or picture) pairs are used in the TNT task, which is also conducted immediately after the learning phase. This process creates a strong association between the cue and the memory. However, the cues in this study are generated by participants 24 hours before the TNT task, and although we instruct them to generate cues that are unique to the memory, the same cue may evoke memories of multiple events. This potentially leads to relatively weaker associations between cues and the autobiographical memories, leading to lower intrusions rates in this study. Nevertheless, we were able to successfully measure intrusions using the present modified think/no-think task. Critically, the frequency of memory intrusions reduced over repeated attempts at memory suppression, and this is a replication of a reliable effect in previous TNT research (Benoit et al., 2015; Hellerstedt et al., 2016; Levy & Anderson, 2012). Therefore, repeated attempts at memory suppression reduces the unintended intrusions of autobiographical memories in response to reminders, in line with expectations.

Morally wrong memories were rated as more intrusive than morally right memories across the whole task. The existing research investigating how the negative nature of memories affects the phenomenological measure of intrusiveness in the think/no-think task is scarce. Of the three studies in this area, negative memories are found to be less intrusive (van Schie et al., 2013), more intrusive (Davidson et al., 2020a), or not differently intrusive (Gagnepain et al., 2017) than neutral memories. The present finding provides new insight into this issue, by using complex real-world memories that could be more conducive to intrusiveness and suppression. These results are also in line with traditional theories that negative memories are more salient and capture more attention and should thus be more difficult to control (see Compton, 2003).

Next, the consequences of memory suppression/retrieval on the mnemonic and emotional aspects of memory was analysed. There was no clear indication of suppression-induced forgetting of mnemonic content for either morally right or morally wrong memories. This was measured by detecting changes in specific contents in the memory descriptions provided by participants, due to suppression or retrieval, and was termed as “specificity” in this study, in line with previous studies measuring forgetting of autobiographical memories (Noreen & Macleod, 2013; Noreen et al., 2016). Unexpectedly, there was no significant change in this specificity measure after, compared to before, the think/no-think task, regardless of retrieval or suppression instructions. A less strict criteria was also used to quantify this effect, and any general differences in memory descriptions were noted and rated in a “similarity” rating. Curiously, morally wrong memories were rated as less similar than morally right memories, regardless of if they were retrieved or suppressed in the think/no-think task. In the same vein, guilt and shame were also reduced regardless of retrieval or suppression for morally wrong memories, whereas there was no change in these emotions for morally right memories. Therefore, the present results indicate that mnemonic contents and emotional responses to reminders of especially morally wrong memories are potentially more susceptible to change over repeated exposure to reminders. This interpretation is not completely conclusive however, because a baseline was not used in the study, and morally right memories had floor effects for guilt and shame measures, so it is not clear why the change in guilt and shame for morally wrong memories is present in this study (see Chapter 6 for further discussion).

In this study, participants were first asked to generate autobiographical memories of past morally wrong and morally right acts, so it was important to also understand the characteristics of such memories. As expected, participants indicated that their actions in morally wrong memories were more morally wrong than their actions in morally right memories. Additionally, they reported feeling more guilty, ashamed, upset, and felt less pleasure while thinking of morally wrong compared to morally right memories. Importantly, participants indicated that they do indeed distance themselves from their past immoral acts. Participants reported that memories of their past morally wrong acts were less vivid (as Kouchaki & Gino, 2016 find) and more remote (as Escobedo & Adolphs, 2010 find) than morally right memories. In line with these findings, participants also reported that their actions in morally wrong memories were less intentional than their actions in morally right memories, which is a novel finding. Therefore, the study successfully replicates and extends on extant evidence from phenomenological ratings that memories of past immoral acts are perceived to be more distant form the self than morally right memories. It is also interesting to note that although morally wrong memories were rated as more distant than morally right memories, they also tended to be more difficult to suppress when faced with reminders of the memories. This seemingly paradoxical pattern is explored further in the general discussion (Chapter 6)

3.3.1 ERP correlates of autobiographical memory suppression and intrusions

The most reliable ERP effect was a large positivity that began around 750ms to 3500ms for successful retrieval compared to successful suppression attempts regardless of moral memory type (see Figure 3.10). This positivity was initially strongest in left parietal regions but later on was stronger in more fronto-central regions. The cognitive mechanisms reflected by this effect is better understood by splitting it into two different components that reflect two different cognitive processes. First, the early parts of this component potentially reflect the left-parietal positivity that is thought to index memory recognition/retrieval (Rugg & Curran, 2007; Wilding, 2000). This reduction in the left-parietal positivity for no-think items is arguably the most reliable effect in the think/no-think literature (e.g., Bergström et al., 2007, 2009a, 2009b; Chen et al., 2012; Depue et al., 2013; Hanslmayr et al., 2009; Hellerstedt et al., 2016), the present study therefore successfully replicates this effect with autobiographical memories. The timing of this effect was relatively later than usually found, as it was maximal around 750-1000ms post-stimulus onset, (but still found early around 500ms) whereas most previous research find that it is maximal around 500-800ms post-stimulus onset. The second segment of this effect is a constant positive slow wave component, potentially beginning around 1500 to 3500ms and maximal in fronto-central regions. This positive slow wave is found in some previous think/no-think research with simpler stimuli, and is thought to reflect cognitive control processes that work to keep the memory from entering awareness, after retrieval has been successfully avoided (see Depue et al., 2013; Hanslmayr et al., 2009). This result also further substantiates new evidence that the parietal positivity increase is related to episodic retrieval of autobiographical memories (e.g., Renoult et al., 2015; Tanguay et al., 2018). Crucially, the results discussed until this point provide clear neurophysiological and behavioural evidence that retrieval of autobiographical memories can be successfully avoided.

An earlier think vs. no-think positivity was found for morally right memories but not morally wrong memories during initial attempts (first block of the TNT task) at memory suppression (see Figure 3.6). This think/no-think positive difference for morally right memories was most reliable around 500-600ms post-stimulus onset but present from around 300-750ms as indicated by the ANOVA analysis. This effect was also maximal in central and frontal regions but present across the whole scalp, indicating that it is potentially different from the left-parietal positivity. Behaviourally, morally right memories were rated as easier to suppress than morally wrong memoires, and therefore the difference due to think/no-think instructions for morally right but not wrong memories could reflect successful cognitive control only for the former memory type. On the other hand, the left-parietal positivity indexing recollection is interestingly not always left-lateralised, for instance, a recent study that used the same cluster-based permutation based method on a recognition test found that the recollection-related positivity was spread across the whole scalp (Hellerstedt et al., 2021), or atleast more parietal but not lateralised to the left hemisphere (see Chapter 4 of the present thesis). Therefore, the greater early positivity when thinking of morally right memories than when suppressing such memories could be indexing suppression success for these memories at an earlier timepoint, whereas such successful suppression was not present for morally wrong memories till a later point (around 750-1000ms). Although it is not entirely clear which cognitive process is reflected by this ERP effect, it does suggest that morally wrong memories were more difficult to suppress than morally right memories, because the brain activity for morally wrong memories was similar regardless of think/no-think instruction, whereas the brain activity for morally right memories was more different. This conclusion is in line with the behavioural results and in line with expectations.

There were other ERP differences due to memory suppression in the present study that were significant in multifactorial ANOVA analyses (see Appendix A.5. for more detail) but did not survive multiple comparison correction in the cluster-based permutation tests (please see Tables 3.4, 3.5, 3.6). This casts doubt on the reliability of these effects as they could be false positives and should be interpreted with caution. However, they are noteworthy as the ERP effects appear similar to ERP components found in previous research. Firstly, the P2 peak was greater for think compared to no-think items in the present study. This peak is suggested to indicate allocation of attention towards think/no-think cues, especially due to the colour differences in stimuli, and has been found in previous think/no-think studies (Bergström et al., 2007; Mecklinger et al., 2009).

The N2 is another component sometimes found in think/no-think research (e.g., Chen et al., 2012; Mecklinger et al., 2009), and in fact, some research asserts that it is an index of cognitive control over memories (Streb et al., 2016). The N2 is a negative deflection that is enhanced for no-think compared to think ERPs (or equivalent, a positivity for think compared to no-think items). This N2 effect is usually found around 350-450ms in fronto-central regions. Looking closer at this time window and scalp location, when think vs. no-think for morally right memories was compared to think vs. no-think for morally wrong memories, then morally right memories seemed to elicit a typical N2, whereas the polarity was reversed for morally wrong memories. Streb et al., (2016) found that participants who experienced fewer intrusions of a traumatic memory had greater N2 amplitudes, whereas individuals with high intrusion frequencies did not have this N2 effect. Therefore, the lack of N2 for morally wrong memories could indicate that they are more difficult to control than morally right memories, which could be easier to control.

Most previous research that found the suppression-related N2 component have focussed on specific ROIs (e.g., Chen et al., 2012; Mecklinger et al., 2009; Streb et al., 2016) as such focal and brief effects are difficult to detect in exploratory whole-head analyses with correction for multiple comparisons (but see also Bergstrom et al., 2009b). The cluster-based permutation test used here is primarily sensitive to large effects that span across time and electrode space. This could be one reason for the absence of the P2 and N2 effects in the cluster-based analyses, because the large effect described above could overshadow these effects. Therefore, future research is needed that directly addresses the role of these components in emotional autobiographical memory suppression.

Another aim of the present study was to understand the neural correlates of autobiographical memory intrusions. Intrusion-related ERP differences were found in the multifactorial ANOVA but did not survive cluster-based permutation tests, potentially because there were fewer trials for intrusions than other conditions in the study, leading to poorer signal-to-noise ratio. The intrusion-related components could of course thus be false positives, but they are consistent with Hellerstedt et al.'s (2016) ERP findings. The left-parietal positivity was also reduced for intrusions, like successful suppression, compared to successful retrieval, and this finding is consistent with Hellerstedt et al.'s (2016) results. Their tentative interpretation is that successful suppression could be occurring even during intrusion trials, but there was no conclusive interpretation. One potential reason for this however is that retrieval may have been successfully avoided early on even during intrusion trials, but contents of memory could have intruded at a later point in time, because control could not be sustained (see oscillation results). Nevertheless, it is important for further research to consider this answering this particular question.

The ANOVA analysis also found that intrusions elicited more positive ERPs than successful suppression trials in frontal regions around 400-500ms post stimulus onset. This effect is overlaps with the timing and topography of the FN400 found for intrusions compared to avoided retrieval in a previous study (Hellerstedt et al., 2016). The FN400 in this instance could therefore reflect reactivation of the memory, and this reactivation has found to be involved in initiating cognitive control processes that are involved in inhibiting traces of the memory from entering awareness (Hellerstedt & Johansson, 2014). Future research is therefore needed to substantiate these hypotheses, potentially experimental designs that could have greater intrusion rates and power to detect these effects.

3.3.2 Oscillatory correlates of autobiographical memory suppression and intrusions

The EEG oscillation data further substantiated our predictions that retrieval of autobiographical memories can be successfully avoided by direct suppression. Overall, a strong decrease in oscillatory power during successful suppression compared to successful retrieval was found across all tested frequencies (4-30Hz), maximal in parietal regions. This effect began around 500 to 1000ms post-stimulus onset and was consistently present for as long as the reminder was presented on the screen (3500ms). This broadband, consistent desynchronisation for no-think compared to think items is consistent with the oscillatory correlates of successful memory suppression found in previous research (Ketz et al., 2014; Legrand et al., 2020; Quaedflieg et al., 2020; Waldhauser et al., 2015).

Overall, the research indicates that the desynchronisation in the theta band (4-7Hz) and alpha/beta bands (7-30Hz) seem to reflect different cognitive processes. Firstly, theta band increases are often related to success in retrieval (Osipova et al., 2006). The raw power data from the think/successful retrieval conditions support this idea, as there is greater power increase for these conditions in the theta band throughout the epoch in the present study. Whereas for no-think compared to think items, there is a significant theta power desynchronisation reflecting successful avoidance of retrieval. The alpha/beta power decrease for no-think vs. think items on the other hand is argued to reflect maintenance of sustained control over the memory for as long as the participant is exposed to the reminder of the to-be-suppressed memory (see Waldhauser et al., 2015). These effects were found regardless of memory type or initial/latter attempts at suppression/retrieval.

Intriguingly, when only the theta band was analysed in the cluster-based test (in a supplementary analysis), the think vs. no-think effect was present for morally right memories, but not morally wrong memories. Note that this analysis was conducted post-hoc in an exploratory manner and should not be taken as conclusive evidence and any interpretations are exploratory and should be confirmed by future research. However, it does match with previous research indicating that mental stress can affect ability to suppress memories (Quaedflieg et al., 2020), and this is evidenced by a lack of no-think vs. think desynchronisation in the theta band. The lack of theta effects for morally wrong compared to morally right memories could potentially indicate that morally wrong memories are harder to control, in line with the behavioural and ERP findings of this study. Further research testing this hypothesis is needed to provide conclusive evidence using oscillatory data. Nevertheless, there is clear evidence from oscillatory data that autobiographical memories can be successfully suppressed.

Intrusions induced a significant increase in oscillatory power in the alpha/beta (8-30Hz) bands around 2500-3500ms post-stimulus onset in parietal regions compared to successful suppression (see Figures 3.15 and 3.16). Contrary to these results, previous research on the oscillatory correlates of intrusions found that intrusions *decreased* rather than *increased* oscillatory power compared to successful suppression (Castiglione et al., 2019; Legrand et al., 2020). However, those effects were generally found earlier in time and tended to have more frontal topographies than the effect found in the present study. The researchers argue that such effects are directly related to top-down inhibitory processes that are engaged to control the memories (Castiglione et al., 2019), similar to motor-action stopping. Therefore, the present results could reflect a functionally different mechanism, even though they are associated with the same behavioural outcome. When considering suppression success vs. retrieval success, a larger decrease in oscillatory power in the alpha/beta band has been suggested to index maintenance of control over retrieval (e.g., Waldhauser et al., 2015; this result is also found in the present study). The power estimates for intrusions (see Figure 3.15) indicates that similar to successful suppression, there is an initial decrease in alpha/beta power, but towards the latter stages of this epoch, this decrease is not present anymore.

However, when suppression is successful, the figures indicate a clear sustained decrease in alpha/beta oscillatory power. Hence, the late intrusion-related increase in alpha/beta oscillatory power could reflect that control could not be maintained throughout the time the reminder was presented to the participant, consequently resulting in an intrusion. This claim is also substantiated by findings that higher frequency of intrusions are reported if the no-think reminder is shown for longer to participants (van Schie & Anderson, 2017). Therefore, the present results potentially provide oscillatory evidence that intrusions could also occur due to a lapse in sustained control.

One issue that is not clear from the present study is *why* morally wrong autobiographical memories were more difficult to suppress than morally right memories. One argument is that the more intrusive nature of morally wrong memories in this study could simply be due to the general negative valence of the memories without being specifically related to morality. However, strong theoretical arguments from social psychology posit that the guilt, shame, and self-threatening aspects of these memories would make them particularly intrusive (Stanley & de Brigard, 2019). However, further research is required to determine which specific features of morally wrong memories render them more intrusive than morally right memories. That is, it could be the self-relevant nature of the memories that makes them particularly memorable (Stanley and De Brigard, 2019), or it could be specific emotional characteristics such as guilt or shame, or general negative vs. positive/neutral emotions. Therefore, further research that separates contributions to intrusiveness from emotions, morality, and how strongly a memory is related to the sense of self will be imperative to understand why these memories could be more difficult to control in everyday life. It may be also important to develop paradigms that have larger power to measure autobiographical memory intrusions, especially to clearly elucidate the neurocognitive mechanisms underpinning such intrusions. As a result, more advanced analysis can be conducted to clearly tease out the contributors for making a particular memory more or less intrusive. See general discussion (Chapter 6) for further discussion.

3.3.3 Summary

The present study provides behavioural and electrophysiological evidence that a direct memory suppression strategy can be successfully used to avoid retrieval of autobiographical memories as soon as 500ms post-onset of the reminder. Behaviourally, participants reported on a trial-by-trial basis that they were successful at stopping the memory from coming to mind in the majority of no-think trials. The left parietal positivity, thought to be an ERP index of memory retrieval, was reduced during suppression attempts. A strong neural oscillatory power decrease was apparent for suppression compared to retrieval potentially indicating successful avoidance of retrieval (theta band), and maintenance of control over memory (alpha/beta bands).

When suppression did fail, operationalised by phenomenological reports of memory intrusions on a trial-by-trial basis, the frequency of intrusions were greater during early attempts at suppression and intrusion frequency noticeably reduced over repeated attempts at memory suppression, indicating better control over memory over time. This pattern of results is also in line with previous findings, but it is novel because it is the first demonstration of this effect for autobiographical memory intrusions. The neural correlates of intrusions of these memories were less conclusive but could be associated with reactivation of memory traces indicating a need for control as indexed by the FN400 (see also Hellerstedt et al.'s, 2016). Intrusions also seemed to be associated with a lapse in sustained control over the memory for the full duration of exposure to the reminder, as evidenced by reduced alpha/beta activity, in line with assertions by another study (van Schie & Anderson, 2017).

Interestingly, morally wrong memories were also found to be more difficult to suppress than morally right memories, a conclusion that seemed quite clearly supported by behavioural and ERP data, but less reliably supported by the oscillation data. Participants indicated that they experienced memory intrusions more frequently during attempts at suppressing morally wrong compared to morally right memories. An early ERP positivity was found maximally around 500-600ms (which could be detected earlier from around 300ms) while retrieving compared to suppressing morally right memories, whereas this positivity was absent for morally wrong memories. The exact cognitive processes reflected by this ERP effect is unclear from the present study, but it could be an indicator of better cognitive control for morally right than wrong memories, or more successful and quicker suppression of morally right compared to wrong memories.

In conclusion, this experiment has contributed novel evidence regarding the avoidance of autobiographical memories in real-time, showing that this avoidance may be more difficult for morally wrong memories. The consequences of suppression on the mnemonic and emotional contents of memory are however less clear. The present study therefore also provides a foundation for future research to investigate the real-time effects of memory control on autobiographical memory intrusions and to further understand the neural correlates associated with suppression and intrusions. Importantly, the results indicate that intrusions of such personal and emotional memories can be controlled, especially over repeated attempts at suppression, and this may alleviate negative emotions associated to such memories.

Chapter 4 – Assessing the effectiveness of direct memory suppression of emotional autobiographical memories in the concealed information test

The first experiment of the thesis reported in the previous chapter explored relatively more theoretical questions regarding the neurocognitive mechanisms underlying retrieval, suppression, and unwanted intrusions of autobiographical memories. In this chapter, I report an experiment that investigated a more applied research question, regarding whether direct memory suppression can be used to beat the P300-based forensic memory detection test, such that a person who is guilty of a crime can appear innocent by suppressing retrieval of crime related memories.

The concealed information test that is used in forensic settings relies on the suspect recognising a unique piece of information that only a guilty person would know – the probe – from a list of irrelevant stimuli (Lykken, 1959). There is reliable evidence that if the suspect identifies the probe word, a greater P300 is elicited for the probe vs. irrelevant word, providing an indicator of guilty knowledge (Rosenfeld, 2018). See Chapter 1, section 1.5 for a detailed review. Recent research has found that direct suppression reduces the P300 ERP response to probes in the concealed information test (Bergström et al., 2013; Hu et al., 2015). Therefore, they argue that suppression can be used by suspects to beat the concealed information tests and appear innocent, questioning the validity of such tests in the everyday life.

The most realistic version of these studies was conducted by Hu et al. (2015), who investigated suppression effects on detection of memories of a mock crime performed in a lab-based controlled setting. Although these memories were rich autobiographical memories, there are potential differences between a mock crime conducted in a lab setting and a real crime, which may involve heightened emotional valence and arousal than a mock crime. For instance, Osugi and Ohira (2017) found that higher emotional arousal during encoding (i.e., committing a mock crime) can significantly amplify the P300. Therefore, based on previous research it is not clear whether guilty suspects would be able to suppress the P300 response when elicited by more personal and emotional autobiographical memories that were encoded outside of a lab environment and may be associated with stronger emotional salience.

Results from experiment 1 (Chapter 3) in this thesis revealed that autobiographical memories of past immoral acts that are associated with feelings of guilt and shame are more difficult to suppress than autobiographical memories of past moral acts. This finding suggests that autobiographical memories of ones' own immoral acts may be particularly difficult to conceal by retrieval suppression, but so far, no studies have directly investigated this issue, to the best of my knowledge. The present study incorporated morally relevant autobiographical memories of participants as stimuli in the CIT to investigate how the nature of memories influence ERP-based detection and suppression of guilty knowledge. In this study, participants simulated a mock crime, but importantly, they also reported autobiographical memories of morally wrong and morally right actions that they committed in their past (using a similar procedure as in Experiment 1). Then, they completed the CIT with three blocks: One block of stimuli referring to the mock crime, one block referring to one of their morally right autobiographical memories, and another block referring to one of their morally wrong autobiographical memories. The stimuli for the mock crime were the same as previous studies (Hellerstedt et al., 2021; Hu et al., 2015), but for the personal memories, the researcher generated a probe that consisted of two key words that best described the memory.

Three different groups of participants completed the study, similar to Hu et al.'s (2015) design. One group (*standard- guilty*) of participants received the standard CIT instructions; another group *(suppression-guilty)* received memory suppression instructions in addition to the standard CIT instructions; and the other group *(innocent),* consisted of participants that did not commit the mock crime, nor provide autobiographical memories. Importantly though, they did complete the CIT, and the stimuli were matched with participants in the two guilty groups. The innocent group provides a comparison measure, so as to show that any difference in the P300 amplitude is not due to differences in stimuli-type, and that the "guilty" participants truly recognise the probes. This design therefore enabled me to directly test whether morally-relevant autobiographical memories and mock crime memories differed in terms of how well they could be detected and how much they could be suppressed.

For this study, experimental methods, data analysis, and hypotheses were pre-registered before data collection (see https://osf.io/c5uym/[4]). The first hypothesis addressed the issue that all groups and memories would be expected to show a recognition-related P300 response to targets: H1) Over all memory types and groups of participants, the target will elicit a significantly more positive P300 ERP than irrelevants. The next two hypotheses related to the issue that all three types of memories should be detected for guilty participants who were not trying to evade detection and should not be detected for innocent participants: H2) In the standard-guilty group, the probe-P300 will be significantly more positive than the irrelevant-P300 for all three memory types. H3) There will be no difference in P300 amplitudes between probes and irrelevants in the innocent group for any memory type. The three hypotheses stated above were based on consistent findings in the P300-based CIT literature (Hu et al., 2015; see Rosenfeld, 2019).

The most novel research questions of this study concerns whether suppression (vs. standard or innocent) instructions can attenuate the probe vs. irrelevant P300 difference such that guilty participants would appear innocent, and importantly, how this P300 component

---

[4] Note that the pre-registered is embargoed until Feb 10, 2022 and can be viewed only after that point of time.

and its suppression-induced modulation differs across different types of memories. Based on previous research of the effects of suppression on CIT of mock crime memories, the P300 effect should be statistically significant in the standard-guilty group, but not significant in the suppression-guilty group and innocent groups (Hu et al., 2015). The same pattern of results was also expected for morally right memories as they should not be particularly difficult to suppress. The fourth hypothesis was hence: H4) We expect a significantly reduced P300 probe-irrelevant difference for mock crimes and morally right memories in the suppression-guilty group when compared to the standard-guilty group.

Suppression effects on P300 amplitudes may be more complicated for morally wrong memories. One potential outcome is that that the P300 effect is statistically significant for both guilty-standard and guilty-suppression groups, suggesting that it is more difficult to suppress these types of memories, as found in experiment 1. On the other hand, the pattern of results could be consistent across all three memory types, suggesting that participants can attenuate the P300 component regardless of the emotional nature of the memory. In support of this idea, previous research has found that the P300 component overlaps with the LPP in some scenarios (see Bergström et al., 2013; Hellerstedt et al., 2021), and in experiment 1 of the present thesis, the left-parietal positivity was reduced for both types of memories. Therefore, for morally wrong memories, the pre-registered hypothesis was: H5) We also expect to find a significant suppression-induced reduction in the P300 probe-irrelevant difference for morally wrong memories, when comparing the suppression-guilty group with the standard-guilty group. However, because our unpublished data from another study (experiment 1) suggests that morally wrong memories are more intrusive than morally right memories, we expect that the magnitude of reduction will be smaller for morally wrong, compared to both morally right and the mock crime memories because morally wrong memories are more difficult to suppress.

Hu et al., (2015) also found that the probe-irrelevant ERP difference was associated with a negative deflection after the P300 component in the same parietal regions. This component is known as the late-posterior-negativity (LPN). Interestingly, they found that suppression did not affect this component. So, we expected to replicate these results with all three memory types: H6) The probe-irrelevant LPN difference will be less affected by suppression than the P300, so will be similar across both the standard-guilty and suppression-guilty groups but will not be present in the innocent group.

An important aspect of the concealed-information test is that these group level effects should also be applicable at an individual level such that suspects can be correctly identified as guilty or innocent of a crime. Therefore, all individuals in the experiment were classified as either guilty or innocent based on their neural activity using bootstrap and permutation based tests (see methods section for further details). The predictions are similar to the group level hypotheses: H7) Individual classification rates based on bootstrap and permutation testing will mirror the group level patterns described above.

Complementary analyses were also conducted to further understand the neural mechanisms involved in the concealed-information test. As mentioned in Chapter 2, specifically focussing on particular ROIs (such as the P300 or LPN peaks) could hide any other effects that could occur in other temporal or spatial sites. Even if there are no other effects, these global analyses could provide more information regarding the scalp topography and timings of the P300 and LPN effects. Therefore, global analyses using cluster-based permutation tests (as used in experiment 1) were used to test for any ERP effects across all scalp electrodes and time points.

The EEG oscillatory correlates of the concealed-information tests were also investigated in the present study. Based on the general memory recognition/retrieval-related

EEG oscillation effects (see Chapter 1, section 1.2.2), two hypotheses were proposed in the pre-registration document: H8) We expect recognition to be associated with increases in parietal theta power, so parietal theta power will mirror the P300 patterns hypothesised above. H9) Memory suppression attempts will be associated with right frontal beta power increases.

However, a very recent study that was published after the pre-registered document was uploaded on OSF, has directly tested the oscillatory correlates involved in the CIT. Overall, they find an early theta power synchronisation for probes/targets vs. irrelevant items, followed by a later alpha/beta power desynchronisation (Hellerstedt et al., 2021), indicating recognition of a stimulus (reviewed in Chapter 1, section 1.2.2). Therefore, an additional prediction in this study that was not pre-registered is that not only the early theta synchronisation (H8) but also a later alpha/beta power desynchronisation effects reflecting recognition will be found in the present study and will have similar patterns as the P300 component.

I also had confirmatory predictions about the phenomenological characteristics of the autobiographical memories. Overall, we expect both morally wrong memories and the mock crime to be rated as more morally wrong, guilty, shameful, and more arousing than the morally right memories, whereas morally right memories will be rated as more pleasurable than the other two types of memories. Based on our predictions about the emotions associated with the mock crime, we expect participants to feel that their actions were less morally wrong, feel less guilty and ashamed, and less aroused, while thinking of the mock crime, than remembering morally wrong memories.

4.1 Method

The methods and analysis plans reported below are largely in line with the pre-registered plan. Any deviations from the original plan are explained in text.

4.1.1 Participants

Fifty-eight participants took part in this study: 18 in the standard-guilty group, 20 in the suppression-guilty group, and 20 in the innocent group. Their ages ranged from 18-20 years (standard-guilty: *M* = 18.78, *SD*= .73; 12 females), 18-21 years (suppression-guilty: *M* = 19.30, *SD* = .92; 15 females), and 18-21 years (innocent: *M* = 19.10, *SD* = 1.02; 14 females). A power analysis conducted by Hu et al., (2015) indicated that 26 participants in each group was required to detect a large suppression effect (Cohen's d = 0.8) with a power of 0.8 at an alpha of .05. Although I planned to collect 26 participants per group (see https://osf.io/c5uym for more details[5]), data collection was stopped due to the COVID-19 pandemic. All participants scored below the criteria for exclusion outlined in the preregistration document, so none were excluded. All participants were right-handed, native English speakers, had no psychiatric or neurological diagnosis, and were not taking any medicine that affects cognition. The study was approved by the Research Ethics Committee of the School of Psychology at the University of Kent.

4.1.2 Design, Procedure, and Materials

The study was conducted in two sessions for the guilty groups, but only one session for the innocent group. In the first session, participants completed a mock crime, and reported morally relevant autobiographical memories. Twenty-four hours later, the CIT was conducted. Three different groups of participants completed the study: Participants in the *standard-guilty* group committed the mock crime, provided autobiographical memory

---

[5] Note that the pre-registered is embargoed until Feb 10, 2022 and can be viewed only after that point of time.

descriptions, and completed the CIT with standard instructions. The *suppression-guilty* group undertook the same tasks as the standard-guilty group but also received memory suppression instructions (based on Hu et al., 2015) in addition to the standard CIT instructions. The *innocent* group neither committed the mock crime nor provided descriptions of autobiographical memories, but they did complete the CIT to provide a baseline comparison for CIT results with an absence of memories. Participants were assigned to groups in a fully randomised order.

Session 1 – both “guilty” groups only.

Mock crime. Participants were instructed to navigate through the School of Psychology building and find a kitchen. They searched through the cupboards in this kitchen, and found a blue bag, and their task was to steal the ring in this bag and bring it to the experimenter, who was in the EEG lab. They were also asked to closely inspect the ring to ensure encoding of the stolen item, but were not told that their memory for the object would later be tested, in line with previous mock crime procedures (Hu et al., 2015).

Autobiographical memory generation. Participants were instructed to think of six different autobiographical events (three memories of past morally wrong actions and three memories of morally right actions), although only two of those memories would later be tested (see next section). The memories were collected using the same procedure as in experiment 1: First, they described the memory in a text box and then self-reported the phenomenological and emotional characteristics of that memory, similar to experiment 1. However, in this study, participants reported emotions they experienced both when they were describing the memory at that time of the experiment, and emotions experienced when the event originally occurred. Participants were also asked to rate the morality of their actions in the mock crime, and report emotions associated to committing the mock crime using the I-PANAS-SF and SAM measures (similar to experiment 1, see Appendix A.1. for an example. The emotions measured by the I-PANAS-SF and SAM scales were also measured for when the event in the memory actually happened, in addition to the emotions felt at the time of describing the memories during the experiment. This is not reported in A.1., but the scales are identical and only the wording was changed).

Session 2 – all "guilty" and "innocent" groups. Participants were told that an object was stolen from a kitchen nearby and that the experimenter knew some personal information about the criminal, and they must complete this test to prove that they are not involved in the crime. EEG electrodes were attached to participants, impedances were reduced, and then the CIT was conducted in three blocks, with one block testing memories for the mock crime, one block testing memory for a morally wrong event reported by guilty participants, and one block testing memory for a morally right event reported by guilty participants. The order of presentation of the three blocks was counterbalanced across all participants in each group. Instructions were delivered such that from the participant's perspective, all three blocks served the purpose of detecting them as guilty of the crime and identifying them as the criminal.

Participants were instructed that, words would appear on the screen one at a time and that they should respond by pressing a button on the keyboard corresponding to "yes" if they recognised the word, or "no" if they did not recognise the word. There were three different stimulus types: a probe (that should remind them of the mock crime or autobiographical event), irrelevant words (control items that should not remind participants of any specific events), and target words (items that were unrelated to the crime/autobiographical event but required a special response). The specific words that were used depended on the block that the participant was completing, but participants were always told which word was the target and therefore required a "yes" response. They were also instructed to try to appear innocent by saying that they did not recognise any of the words other than the target (even if they did recognise those words as related to the mock crime or autobiographical events). They were thus instructed to press “yes I recognise this word” only for the target and “no I do not recognise this word” for the irrelevant and probe stimuli.

If the participant was completing the mock-crime block, the probe was the word “ring”, and the target was the word “watch”, and the irrelevants were the words “wallet”, “necklace”, “keys”, “phone”, and “bracelet”. One of the other two blocks contained phrases referring to morally right memories, and the other block contained phrases referring to morally wrong memories. The probe for these two blocks was a phrase (made up of maximum two words) that relates to one morally right or one morally wrong (depending on block) memory that the participant described in session one. Two memories were selected out of the six that participants reported, with the aim to match these as closely as possible on self-reported memory vividness. The target and irrelevant phrases were chosen by the experimenter from a list of probes used for other participants in this experiment. They were similar to the probe in terms of the memory type (morally right or wrong). This was to ensure that the participant completing the CIT recognised the autobiographical memory probes based on their memory, and not on the basis of these probes being distinctive for other semantic or perceptual reasons.

In all three blocks, participants were told to pay full attention to every word that appeared on the screen. To ensure they did so, seven attention checks were conducted in each block, leading to 21 checks for each participant in total. The number of failed attention checks in the standard-guilty (*M* = 1.61, *SD* = 1.61; Range: 0-5), suppression-guilty (*M* = 1.35, *SD* = 1.03; Range: 0-4), and innocent (*M* = .95, *SD* = .94; Range: 0-3) groups were not significantly different, *F*(2,57) = 1.24, *p* = .23.

In addition to the above instructions that were provided to all three groups of participants, the suppression-guilty group also received additional instructions on how to suppress retrieval of memories of the mock crime and autobiographical memories. They were told that the brainwave test that they were about to complete relies on the recognition of phrases related to their personal memories, and recognition of the object that was stolen in the mock crime. So, to beat the test, they must actively prevent any thoughts about the mock crime and their autobiographical memories from coming to mind. These instructions were adapted from previous research (Bergström et al., 2013; Hu et al., 2015).

Before starting each block, participants practiced pressing the “yes” button on the keyboard when the target appeared and pressing the “no” button when any other stimuli appeared. Each trial in all blocks (including practice phases) of the CIT followed the same structure. A blank screen first appeared on the screen for a jittered duration (500-1000ms), followed by a fixation cross for 500ms before the word that was be displayed for 300ms, followed by an inter-trial-interval for 1000ms. Responses to the word were collected until 1750ms post-stimulus-onset. See Figure 4.1 for an illustration of the key experimental phases.

After the CIT, participants completed a post-experimental questionnaire that assessed their perceived performance in the CIT. Participants were planned to be excluded from further analysis if they reported non-compliance to suppression instructions (Hu et al., 2015). However, no participants were excluded as none crossed the threshold for exclusion. Please see supplemental file for the complete post-experimental questionnaire.

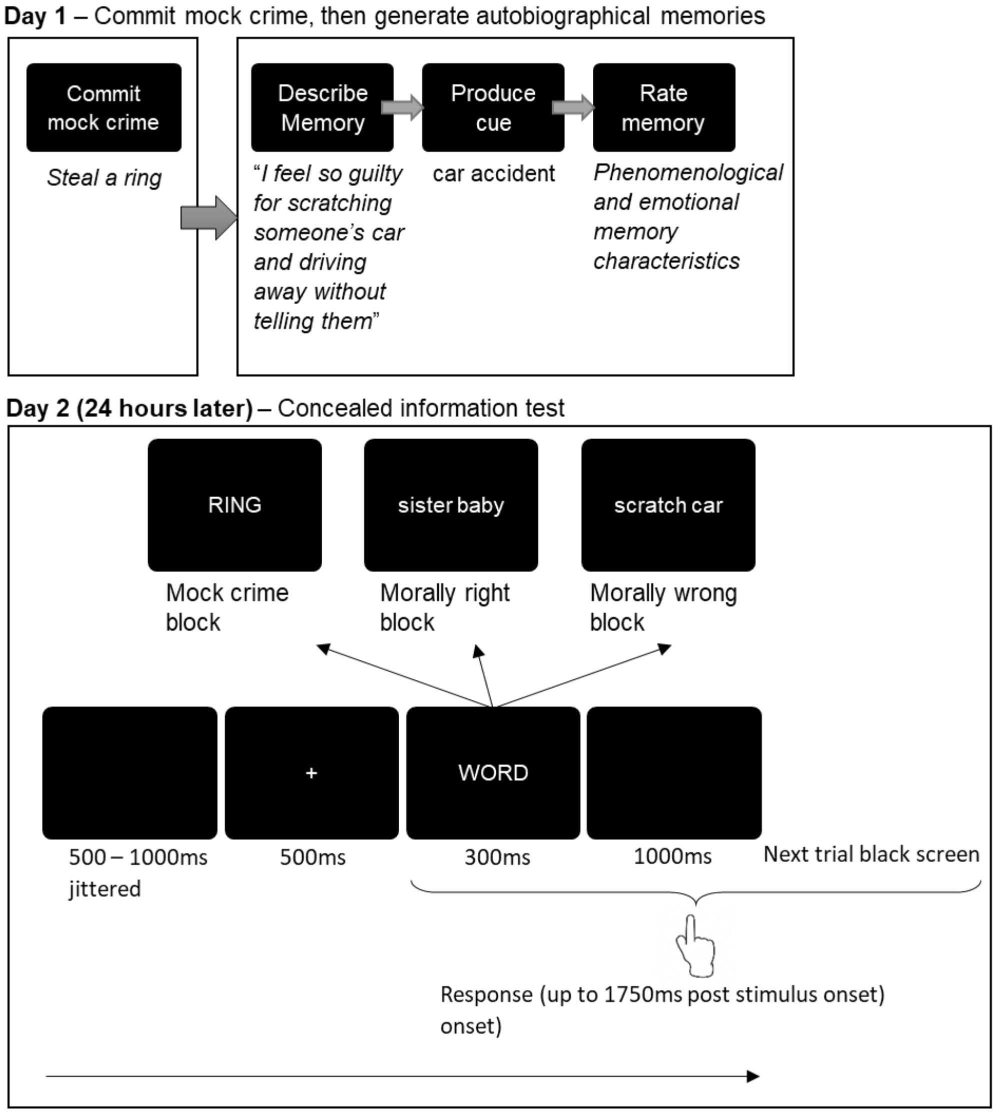


Figure 4.1. Illustration of one trial in the critical procedural phases in the experiment.

4.1.3 EEG data recording, pre-processing, and extraction

The EEG data were pre-processed and analysed using the EEGLAB toolbox for Matlab and self-written code using a standard pre-processing timeline (similar to Hu et al. 2015) and in-line with the pre-registered parameters. In brief, this involved re-referencing the EEG data to the average of the mastoids and filtering it with a 0.3Hz (at -3dB) high-pass filter. This parameter is different from experiment 1 (0.1Hz was used), but consistent with previous research using the CIT paradigm where a 0.3Hz high-pass filter is recommended since it improves memory detection (Hellerstedt et al., 2021; Hu et al., 2015). Next, the data was segmented into 1.75 second epochs (including a 1s pre-stimulus period, of which -200 to 0ms were used to baseline correct the ERPs). Epochs were time-locked to the onset of each word/phrase in the CIT. These epochs were concatenated and corrected for non-brain artefacts using extended info-max ICA using Runica from the EEGLAB toolbox, with default extended-mode training parameters (Delorme & Makeig, 2004). After removal of noise components, corrected data was subsequently low pass filtered digitally at 40 Hz. Any trials that still contained visible artefacts after filtering were manually removed.

Similar to the first chapter, both ERP amplitudes and oscillatory power were derived from the epochs. The same parameters as the first study were used for time-frequency decomposition of the data. To remove edge artefacts from analysis, the epochs were truncated to -625 to 1300ms for all further analyses.

4.1.4 EEG statistical analysis

A 3 (Instruction Group: Standard-guilty, suppression-guilty, innocent) x 3 (Memory Type: Mock crime, morally right, morally wrong) x 3 (Stimulus Type: Probe, target, irrelevant) design was used to divide the epochs into different conditions. On average, there were 47 trials for probes, 43 trials for targets, and 237 trials for irrelevant stimuli, regardless of memory type or instruction group after pre-processing.

Based on a large body of previous research, specific hypotheses regarding the timing and electrode location of P300 and LPN effects justified a focal analysis of these ERP peaks. Measures that were extracted in the focal analysis were further used for individual guilt-classification, since a key application of the CIT is to detect if a suspect is guilty of the crime. This focal analysis may however overlook differences between conditions that could be occurring in other time regions or electrode sites. Therefore, a complementary exploratory whole-head analysis was also conducted using data-driven cluster-based permutation test method, similar to experiment 1.

Focal analysis of ERP peaks. Individual EEG trials were further lowpass filtered with a 6 Hz cut-off (and baseline-corrected again) in order to increase signal-to-noise ratio. The peaks of interest were measured by identifying maximum and minimum mean ERP amplitudes (measured in µV) using a sliding time-window at the individual level as is typical in the CIT literature. For the P300 peak, the mean amplitude of the most positive 100ms time-window between 300ms to 800ms post-stimulus onset at the Pz electrode site was used. To measure the Late Posterior Negativity (LPN, e.g. Hu et al., 2015), I used the mean amplitude in the most-negative 100ms segment following the P300 latency until 1500ms post-stimulus onset. A combined peak-to-peak difference measure was also calculated as the difference between P300 and LPN (P300 – LPN) peak amplitudes. Then, ANOVAs and t-tests were conducted using these three measures to test for specific hypotheses.

Individual guilt classification. Both bootstrap-based and permutation-based classification was conducted to test for reliable/significant recognition effects within each individual, as there is a debate in the literature regarding which method is more effective for guilt classification (see chapter 2 for a review of these methods; Rosenfeld & Donchin, 2015; Zoumpoulaki et al., 2015). These analyses used the individual EEG trials that had been further lowpass filtered with a 6 Hz cut-off (and baseline corrected), in order to increase signal-to-noise ratio, similar to the focal analysis of ERP peaks.

The probe vs. irrelevant P300 and P300-LPN peak-to-peak measures[6] were computed using the sliding windows described in the previous section ("Focal analysis") and used for guilt classification of each participant, for each memory type. For the bootstrap tests, a 90% threshold was used, and one-tailed $p < .1$ was used for permutation tests to classify an

---

[6] Note that although we had planned in the pre-registration to also conduct separate classification on the LPN also, this analysis was not included because the LPN peaks were rather weak or even absent in the group ERPs, meaning that classification of individuals would not be very meaningful (since classifying on basis of an absent/weak component would be based on noise rather than signal).

individual as guilty, in line with previous research (Rosenfeld et al., 2016). To assess the performance of these two classification techniques, a threshold-independent ROC analysis was conducted, and areas under ROC curves were computed for each participant, for each memory type. See chapter 2 for a review of this method. The AUCs were computed for the standard-guilty vs. innocent and suppression-guilty vs. innocent ROC curves to determine how well these groups could be distinguished based on their ERPs (as is typical in this literature, e.g., Hu et al., 2015).

Cluster-based permutation tests. For the whole-head analysis, the EEG trials were not subject to additional filtering in order to ensure that both ERP and time-frequency analysis was conducted on the exact same data. All 28 scalp electrodes and all time points from 0ms to 1300ms were included and submitted to nonparametric cluster-based permutation tests. This approach was used for both ERP amplitudes and oscillatory power derived from time-frequency decomposition, wherein frequency (ranging from 4-30Hz) was added as a third dimension along with electrodes and time-points. This analysis was conducted using FieldTrip toolbox for Matlab (Oostenveld et al., 2011) using the same parameters as in Experiment 1 unless otherwise specified.

Planned comparisons. The exact same analysis plan was implemented for both target vs. irrelevant and probe vs. irrelevant comparisons. First, pairwise comparisons were conducted to determine if the P300, LPN, and peak-to-peak components were enhanced for probes/targets compared to irrelevant items for each memory type in each instruction group. Second, to test if the components were affected by instruction group, a 2 (Instruction Group) x 2 (Stimulus Type) mixed-design factorial analysis was conducted for each memory type. For example, the probe vs. irrelevant difference for mock crime memories was compared across standard-guilty and innocent groups, and then in a separate analysis across suppression-guilty and innocent groups. Finally, to test if these components were modulated by the memory type, 2 (Memory Type) x 2 (Stimulus Type) repeated measures factorial design was conducted in each instruction group. For example, probe vs. irrelevant was compared between morally right and wrong memories within the standard-guilty group.

4.2 Results

4.2.1 Behavioural results

The differences in memory characteristics between the different memories is first reported, followed by an analysis of the accuracy and reaction time in the CIT. Finally, results from the post-experimental questionnaire measuring participants' self-reported performance in the CIT are reported.

Phenomenological and emotional characteristics of autobiographical memories. As in experiment 1 and as described in the pre-registration document, morality, guilt, shame, pleasure, and arousal were chosen for further analyses (see Appendix B.1 for results). However, as the memories were not rated after the CIT, only the differences in characteristics between memory types when they were described by the participants before the CIT was analysed. The results were largely in line with pre-registered predictions and findings from experiment 1: Participants felt more guilty, ashamed, upset, felt less pleasure, and rated their actions as more morally wrong after describing both morally wrong and mock crime memories compared to after describing morally right memories. Interestingly, although both mock crime and morally wrong memories did not differ in perceived moral *wrongness* which was not in line with what was predicted, participants indicated feeling more guilty, ashamed, upset, and felt less pleasure after describing morally wrong than the mock crime memory, as expected.

Additionally, morally wrong memories were rated as more distant, less vivid, and less intentional than morally right memories, in line with experiment 1 and previous findings in the literature (Kouchaki & Gino, 2016). Although this analysis was unplanned, it was conducted to compare the results from experiment 1 (Chapter 3) and experiment 3 (Chapter 5). See Appendix B.1. for further details regarding these analyses.

Accuracy and reaction time. Behavioural performance in the concealed information test is shown in Table 4.1.

Table 4.1. Descriptive statistics showing accuracy (in proportion correct responses) and reaction times (in milliseconds) for each condition in the CIT.

		Probe		Irrelevant		Target	
		Accuracy	RT	Accuracy	RT	Accuracy	RT
Standard Guilty	Mock Crime	99.44 (.01)	1151 (224)	99.29 (.01)	1185 (170)	89.44 (.13)	1236 (122)
	Morally Right	98.56 (.02)	1216 (129)	99.33 (.01)	1211 (136)	92.11 (.09)	1260 (55)
	Morally Wrong	99.22 (.01)	1221 (100)	99.42 (.01)	1213 (91)	89.77 (.11)	1264 (36)
Suppression Guilty	Mock Crime	98.70 (.02)	1138 (224)	99.00 (.02)	1144 (216)	89.8 (.12)	1241 (87)
	Morally Right	97.8 (.05)	1159 (157)	92.30 (.09)	1177 (152)	92.3 (.09)	1262 (36)
	Morally Wrong	98.9 (.02)	1175 (132)	99.36 (.01)	1177 (124)	90.4 (.10)	1259 (43)
Innocent	Mock Crime	99.17 (.02)	1130 (188)	99.17 (.01)	1174 (131)	88.03 (.13)	1232 (84)
	Morally Right	99.1 (.02)	1205 (103)	98.32 (.03)	1199 (90)	88.7 (.11)	1244 (54)
	Morally Wrong	98.62 (.03)	1175 (152)	98.76 (.02)	1195 (112)	88.23 (.11)	1236 (66)

An omnibus 3 (Instruction Group: Standard-guilty, suppression-guilty, and innocent) x 3 (Memory Type: Mock crime, morally right, morally wrong) x 3 (Stimulus Type: Probe, irrelevant, target) mixed-design ANOVA was conducted to test if accuracy and reaction times were different across conditions. None of the factors interacted with instruction group (all *Fs* (4,110) < .56, *ps* > .62), so all results presented below are reported collapsed across the groups. The memory type x group interactions for both reaction time and accuracy were also not significant; both *Fs*(4,110) < 1.8, *ps* > .15. In the pre-registration document, we predicted no significant differences in accuracy or reaction time between conditions, but unexpectedly these differences did exist (see below). Furthermore, as shown in Table 1, the accuracy for targets were lower (around 90%) than previous research (around 95%), whereas probes and irrelevants had very high accuracy rates (around 98%). This is not entirely surprising because the responded target required overriding the prepotent "no" response for probes and irrelevants.

Longer reaction times for moral memories than mock crime memories. There were no differences in accuracy between the memory types (the main effect was not significant, $F(2,110) = 0.17$, $p = 0.85$) but there were differences in reaction times between memory type, $F(2,110) = 3.88$, $p = 0.035$, $\eta_p^2 = 0.066$. Three follow-up repeated measures ANOVAs were then conducted comparing RTs between pairwise memory types. Results indicated that reaction times were significantly longer for morally right than mock crime memories $F(2,110) = 6.183$, $p = 0.016$, $\eta_p^2 = 0.098$; and reaction times for morally wrong memories were also longer than mock crime memories, but this difference did not reach the threshold for significance, $F(2,110) = 3.76$, $p = 0.057$, *partial* $\eta^2 = 0.062$. There was no significant difference in RT between morally right and morally wrong memories, $F(2,110) = .056$, $p = .814$, *partial* $\eta^2 = .001$.

Lower accuracy and longer reaction times for targets compared to probes and irrelevants. There was a significant main effect of stimulus type for both reaction time, $F(2,110) = 49.79$, $p <.001$, $\eta_p^2 = 0.475$, and accuracy, $F(2,110) = 23.06$, $p < .001$, $\eta_p^2 = 0.295$. Three follow-up repeated measures ANOVAs were then conducted for each measure, comparing RTs and accuracy between pairwise stimulus types. Results revealed that accuracy was lower for targets than both probes $F(2,110) = 49.06$, $p <.001$, $\eta_p^2 = 0.463$, and irrelevants $F(2,110) = 53.43$, $p <.001$, $\eta_p^2 = 0.484$, meaning that participants were more likely to make errors in detecting and responding to the target word. Additionally, participants were slower to react to targets than both probes $F(2,110) = 24.64$, p <.001, $\eta_p^2 = 0.302$ and irrelevants, $F(2,110) =$

23.58, p <.001, $\eta_p^2 = 0.293$. Also, reaction times were significantly longer for probes than irrelevants, $F(2,110) = 8.53$, $p = .001$, $\eta_p^2 = .128$.

Post-experimental questionnaire. This analysis was unplanned at the time of pre-registration but provides important information regarding participants' experiences during the concealed-information test. See Table 4.2. for descriptive statistics regarding this questionnaire. A 2 (Group: Standard-guilty, suppression-guilty) x 3 (Memory Type: Mock crime, morally right, morally wrong) mixed ANOVA was conducted to investigate differences between conditions in automaticity and motivation to beat the test. There was a significant main effect of memory type for automaticity, $F(1,36) = 4.72$, $p = 0.018$, $\eta_p^2 = 0.116$. Three follow up repeated-measures ANOVAs, each comparing two memory types were conducted. Results indicated that mock crime memories ($M = 5.34$, $SD = 1.66$) came to mind more automatically than morally right memories ($M = 5.15$, $SD = 1.73$), $F(1,36) = 8.85$, $p = 0.005$, $\eta_p^2 = 0.193$. Although morally wrong memories ($M = 4.52$, $SD = 1.82$) numerically came to mind more automatically than morally right memories, this difference was near, but did not cross, the threshold for significance, $F(1,36) = 3.36$, $p = .07$, $\eta_p^2 = .083$. There was no significant difference in automaticity ratings between mock crime and morally wrong memories, $F(1,36) = .63$, $p = .43$, $\eta_p^2 = .017$. There were no significant differences in the motivation to beat the test between different memory types, $F(1,36) = 2.39$. $p = .10$. These differences were regardless of Group, as there were no memory type x group interactions (both Fs < 1.31, *p*s > .27).

Unique to the suppression-guilty group, additional questions pertaining the difficulty and confidence in suppressing memories in the CIT were asked. Specifically, participants in this group rated how intrusive the memories were in the first half, second half, and across the whole task. They were also asked how confident they were with the suppression strategy. To test this, a 3 (Memory Type: Mock crime, morally right, morally wrong) repeated-measures

ANOVA was conducted for each measure. However, there were no significant differences in intrusiveness or confidence in strategy between the different types of memories for any measure, all *Fs* (1,19) < 1.78, *ps* > .18.

Table 4.2. Descriptive statistics of the average scores from the post-experimental questionnaire.

	Standard-Guilty		
	Mock Crime	Morally Right	Morally Wrong
Automaticity	5.83(1.42)	4.55(1.94)	5.55(1.95)
Motivation	5.55(1.09)	4.83(1.58)	5.27(1.27)
	Suppression-Guilty		
Automaticity	4.90(1.77)	4.50(1.76)	4.80(1.70)
Motivation	5.50(1.24)	5.15(1.23)	5.25(1.16)
Confidence in strategy	4.00(1.45)	4.15(1.57)	3.85(1.31)
Intrusion first half	3.70(1.72)	3.45(1.43)	4.05(1.73)
Intrusion second half	3.35(1.66)	3.75(1.37)	4.00(1.52)
Intrusion total	3.65(1.34)	3.75(1.52)	4.20(1.73)

Note. All measures were scored on a 7-point Likert-scale. *SD* is in brackets.

4.2.2 ERP results

The results of the focal analysis of ERP peaks is presented first, followed by a whole-head analysis of ERPs using cluster-permutation tests. The analyses were conducted in line with the pre-registered analysis plan. Results from the target vs. irrelevant comparison is reported first in each section, followed by the probe vs. irrelevant comparison.

Focal Analysis of ERP Peaks. See Figure 4.2 for an illustration of the grand average ERPs from the mid parietal site and individually estimated target–irrelevant P300 and probe–irrelevant P300 differences. See Table 3 for results of the pairwise t-tests that were conducted to assess if the P300 for each memory type for each group for both target vs. irrelevant and probe vs. irrelevant comparisons. Results of the LPN and the peak-to-peak P300 measures were also analysed and is reported in the Appendix B.2. to maintain brevity in the main text. In short, the LPN peaks were either weak or completely absent, and the peak-to-peak results mirrored results from the P300 measure.

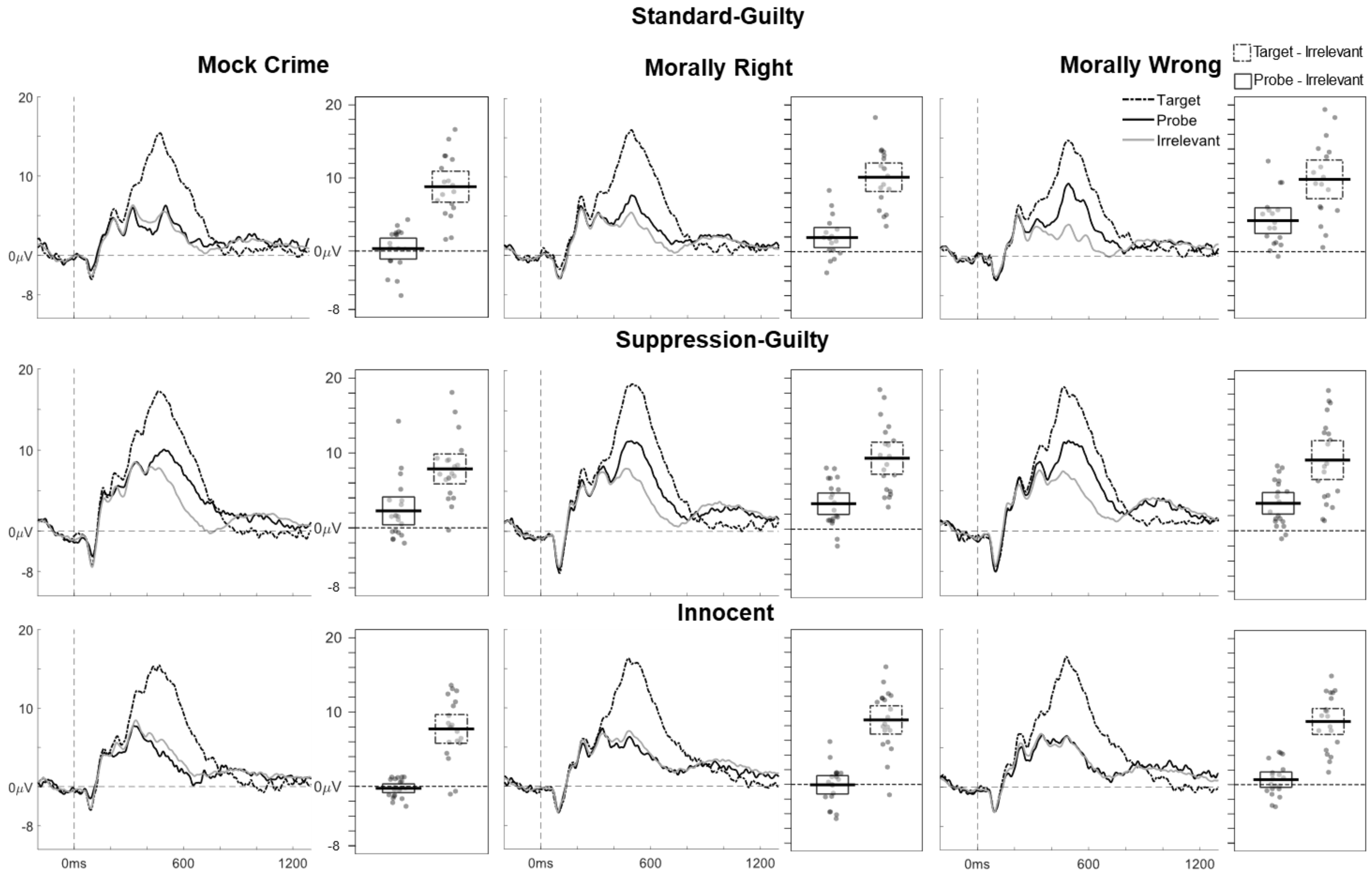


Figure 4.2. Focal analysis of ERPs from the CIT for each memory type in each group. Within each panel, the graph on the left shows grand-average ERP waveforms from the Pz electrode site for target, probe, and irrelevant stimuli. The dot plot on the right shows individually estimated amplitudes of probe-irrelevant and target-irrelevant P300 effects at Pz. The scatter dots show P300 differences between pairwise conditions for each individual. The thick lines show the group means and the boxes depict the 95% confidence interval of the group means.

Table 4.3. Results from several pairwise t-tests comparing both target vs. irrelevant and probe vs. irrelevant stimulus types for the P300.

		Target vs. Irrelevant P300			Probe vs. Irrelevant P300		
		t	*p*	*Cohen's d*	*t*	*p*	*Cohen's d*
Standard-guilty	Mock Crime	**8.77**	**<.001**	**1.69**	0.43	0.67	0.09
	Morally Right	**11.05**	**<.001**	**2.21**	**2.98**	**0.008**	**0.51**
	Morally Wrong	**7.9**	**<.001**	**1.81**	**5.11**	**<.001**	**0.94**
Suppression-guilty	Mock Crime	**8.25**	**<.001**	**1.45**.	**2.54**	**0.019**	**0.49**
	Morally Right	**9.26**	**<.001**	**1.54**	**4.98**	**<.001**	**0.63**
	Morally Wrong	**7.61**	**<.001**	**1.58**	**5.33**	**<.001**	**0.71**
Innocent	Mock Crime	**8.42**	**<.001**	**1.5**	-0.83	0.41	-0.07
	Morally Right	**9.74**	**<.001**	**1.76**	-0.05	0.96	-0.02
	Morally Wrong	**10.21**	**<.001**	**2.07**	1.38	0.19	0.17

Note. Significant results are in bold. *Standard-guilty N = 18, Suppression-guilty N = 20, Innocent N = 20.*

Target vs. Irrelevant. As expected, the paired t-tests revealed that targets elicited a greater P300 than irrelevants for each memory type, in each group (Table 4.3). This target-irrelevant P300 was not affected by instruction group in any memory type block (all *Fs* < 1.27, *ps* > .28). Regardless of instruction group, the target-irrelevant P300 effect was not significantly different for either the mock crime or morally right blocks compared to the morally wrong block (both *Fs* < 1.89, *ps* > .18). The target-irrelevant P300 in the mock crime block was however greater than in the morally right block in both the standard-guilty *F*(1,17) = 4.46, *p* = .05, *partial* η^2 = .208), and suppression-guilty (*F*(1,19) = 4.77, *p* = .042, *partial* η^2 = . 201) groups, whereas this comparison was not significant in the innocent group (*F*(1,19) = 2.97, *p* = .10, *partial* η^2 = .136).

Probe vs. Irrelevant. As predicted, the P300 was not found to be enhanced for any memory probe type when compared to irrelevant P300s in the innocent group (see Table 3 for statistics). Unexpectedly however, probes elicited greater P300 amplitudes than irrelevants for both morally right and morally wrong memories in both guilty groups. For mock crime memories the probe vs. irrelevant P300 was significant in the suppression-guilty but not in the standard-guilty group, contrary to the predictions.

Comparing the probe-irrelevant P300 difference across groups, it was found that this effect was significantly larger in the standard-guilty group than the innocent group for both morally right ($F(1,36) = 5.14$, $p = .03$, *partial* $\eta^2 = .125$) and morally wrong ($F(1,36) = 13.89$, $p = .001$, *partial* $\eta^2 = .278$) memories. The probe-irrelevant P300 difference was however not significantly different between standard-guilty and innocent groups for mock crime memories ($F(1,36) = .54$, $p = .47$, *partial* $\eta^2 = .01$). All three memory types elicited larger probe-irrelevant P300 effects in the suppression-guilty group than the innocent group: Mock crime $F(1,38) = 7.21$, $p = .01$, *partial* $\eta^2 = .159$); morally right $F(1,38) = 14.39$, $p = .001$, *partial* $\eta^2 = .275$); morally wrong $F(1,38) = 11.03$, $p = .001$, *partial* $\eta^2 = .234$). There were no significant differences between standard-guilty and suppression-guilty groups for morally wrong or right memories (both $Fs < 2.21$, $p = .145$). Contrary to predictions, mock crimes elicited a tendency towards a greater probe-irrelevant P300 effect in the suppression-guilty than the standard-guilty group, but this effect did not reach significance, $F(1,36) = 3.02$, $p = .089$, *partial* $\eta^2 = .08$.

In the standard-guilty group, morally wrong memories elicited a larger probe-irrelevant P300 difference than both mock crime $F(1,17) = 10.01$, $p = .006$, *partial* $\eta^2 = .37$, and morally right ($F(1,17) = 12.03$, $p = .003$, *partial* $\eta^2 = .41$) memories. This interaction effect was not significant in this group when comparing morally right and mock crime memories, $F(1,17) = 2.38$, $p = .14$, *partial* $\eta^2 = .12$. There were no significant differences in this effect between memory types in either the suppression-guilty or innocent groups (all $Fs < 1.25$, $ps > .28$).

Focal analysis summary. In sum, as expected, targets elicited greater P300 amplitudes than irrelevants for all memory types in all instruction groups. The probe-irrelevant P300 was elicited by both morally right and morally wrong, but not mock crime memories in the standard-guilty group, whereas all memory types in the suppression guilty group elicited the P300. Expectedly, in the innocent group, none of the probes elicited the P300. Both morally right and morally wrong memories elicited a greater probe-irrelevant P300 in both guilty groups compared to the innocent group, whereas the mock crime P300 was greater for suppression vs. innocent, but not standard vs. innocent groups. Additionally, in the standard-guilty group, morally wrong memories elicited greater probe-irrelevant P300 than both mock crime and morally right memories, and morally right memories elicited greater P300 than mock crime memories.

Individual guilt classification. The bootstrap and permutation tests were conducted at an individual-level for all participants in the experiment using the P300 base-to-peak and peak-to-peak measures to detect if each participant had reliable/significant differences between probes and irrelevant items, which would indicate guilt. Results from both the bootstrap and permutation tests are presented in parallel to be able to compare the results across methods (see table B.5. in Appendix B.3. for individual-level classification results). The peak-to-peak measure indicated identical classification rates as the P300 peak and is thus reported in the Appendix B.3.. A summary of the P300 based classification results is described below.

In the standard-guilty group, the highest number of participants classified as guilty with the P300 measure was for morally wrong memories (14/18 guilty using both bootstrap and permutation), followed by mock crime memories (8/18 guilty using both bootstrap and permutation techniques), followed by morally right memories (Bootstrap: 8/18, permutation: 7/18 guilty) memories. Similarly in the suppression-group, probes for morally wrong memories resulted in the highest number of participants that were classified as guilty (13/20 using both bootstrap and permutation tests), followed by morally right memories (Bootstrap: 13/20, permutation: 11/20), followed by mock crime memories (8/20 using both methods). Importantly, in the innocent group, classification rates were low: 0/20 for mock crime, 3/20 for morally right, and 5/20 for morally wrong memories using both methods were classified as guilty.

Threshold-independent ROC analyses were conducted by individually comparing the guilty groups to the innocent group, for each memory type, for both bootstrap and permutation results (see Table 4.4. for the AUC results). When comparing the standard-guilty group to the innocent group, both bootstrap- and permutation-based classification was better than chance for both morally right and morally wrong memories, but for mock crime memories the AUC indicated that classification was not significantly better than chance using either technique. Similarly, guilt classification using both techniques was better than chance for both the morally-relevant memories in the suppression-guilty group, whereas for mock crime memories, the permutation test AUC was significant, but the bootstrap AUC was not significant (but very close to the alpha threshold).

Table 4.4. Results from the ROC analysis for the base-to-peak measure

	Standard vs. Innocent					
	Bootstrap			Permutation		
	AUC	*SE*	*p*	*AUC*	*SE*	*p*
Mock Crime	0.63	0.09	0.16	0.61	0.10	0.23
Morally Right	**0.70**	**0.08**	**0.033**	**0.70**	**0.09**	**0.039**
Morally Wrong	**0.78**	**0.08**	**0.003**	**0.82**	**0.07**	**0.001**
	Suppress vs. Innocent					
	Bootstrap			Permutation		
	AUC	*SE*	*p*	*AUC*	*SE*	*p*
Mock Crime	0.68	0.09	0.053	**0.71**	**0.08**	**0.021**
Morally Right	**0.82**	**0.07**	**0.001**	**0.81**	**0.07**	**0.001**
Morally Wrong	**0.73**	**0.08**	**0.013**	**0.78**	**0.07**	**0.002**

Note. Statistical significance is indicated in bold.

Guilt classification summary. Both bootstrap and permutation based classification techniques provided very similar results. Overall, significantly more individuals were classified as guilty in the guilty groups than the innocent group for all memory types. This classification performance was efficient for both moral memories for both guilty groups, whereas for mock crime memories, the suppression vs., but not standard vs. innocent comparison had efficient guilty classification. Furthermore, more individuals were classified as guilty with morally wrong memories than both mock crime and morally right memories in both guilty groups.

Global analysis. Results from the cluster-based permutation tests of ERPs are illustrated in Figure 4.3.

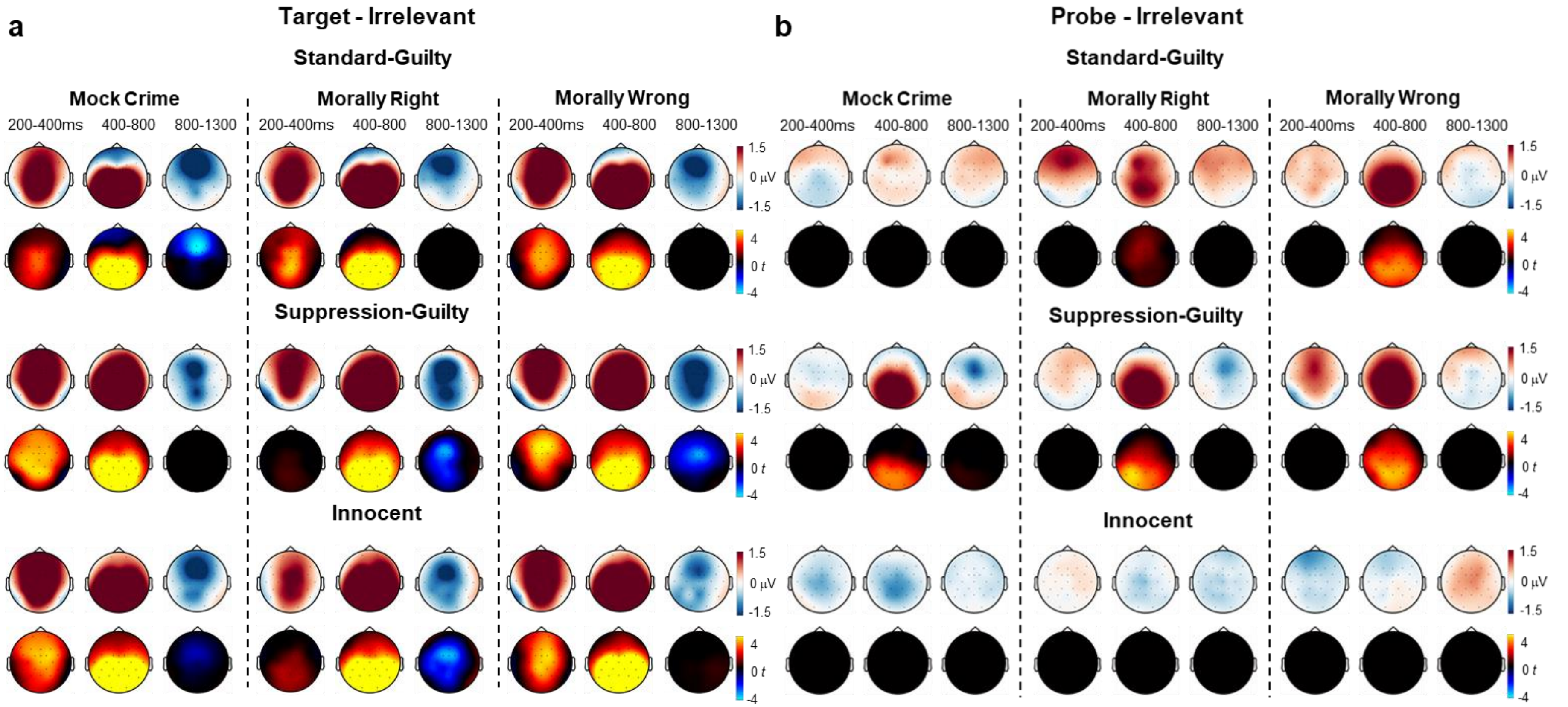


Figure 4.3. Global analysis of ERPs in the CIT. Topographical maps of amplitude differences (top rows, blue/white/red colourmap) and t-values for the differences (bottom rows, cold/black/hot colour map) for Target - Irrelevant (left) and Probe-Irrelevant (right) differences for each memory type in each instruction group.

Target vs. Irrelevant. The targets elicited more positive ERPs than irrelevants for each memory type block in all groups (all *ps* < .001), ranging from around 200ms to 800ms, spread across the central and parietal regions, corresponding to the P300 effect. There were also significant negative clusters later in the epoch, in central and frontal regions, corresponding to the LPN. These negative clusters were significant from around 650 to 1100ns (*p* = .005) and around 1130 to 1300ms (*p* < .001) in the standard-guilty and from around 1100 to 1300ms in the innocent group (*p* = .024) for the mock crime memory. For morally right memories, the clusters were significant in both the suppression-guilty group from around 900 to 1300 (*p* = .006) and innocent group from around 920 to 1300 (*p* = .001). For morally wrong memories, these clusters were significant only in the suppression-guilty group from around 900 to 1300 (*p* = .002).

There were no significant clusters when comparing the target-irrelevant difference between instruction groups for any memory type block.

Interaction analyses comparing differences in target-irrelevant ERP effects between memory types within each instruction group indicated that morally wrong memories elicited a greater target-irrelevant P300 effect than both morally right (around 500 to 650ms; *p* = .008) and mock crime memories (around 500 to 600ms; *p* = .012) in the standard-guilty group. There were no significant clusters when comparing this effect for morally right and mock crime memories in the standard-guilty group. These interaction comparisons did not reveal significant clusters when comparing memory types in any other group.

Probe vs. Irrelevant. Significant positive clusters corresponding to the probe-irrelevant P300 effect were present in both “guilty” groups for both the morally relevant memory types. In the standard-guilty group, this P300 cluster was significant from 390 to 550ms for morally right memories (*p* = .006), and from 400 to 760ms for morally wrong memories (*p* = <.001). In the suppression-guilty group, the P3 cluster was significant from 430 to 800ms for morally right (*p* = .002), and from 400 to 750ms for morally wrong memories (*p* <.001). This effect was present for the mock crime in the suppression-guilty group from around 400 to 800ms (*p* = .003), but it was not significant in the standard-guilty group. As expected, there were no crime relevant probe-irrelevant P300s for any memory type in the innocent group. There were no significant negative clusters (LPN effects) for the probe-irrelevant comparison in any group.

The probe-irrelevant P300 in both guilty groups was significantly greater compared to the same effect in the innocent group for all memory types (all *ps* < .001). There were no significant clusters when comparing probe-irrelevant P300s between standard and suppression guilty groups for any memory type.

Interaction analyses indicated that morally wrong memories elicited a greater P300 than both morally right (*p* = .004) and mock crime (*p* = .01) memories both from around 400 to 600ms in the standard-guilty group. There were no significant clusters when comparing morally right and mock crime memories. These interaction comparisons did not reveal significant clusters when comparing memory types in any other group.

Global analysis summary. In line with the focal analysis, global analyses revealed a positive cluster around 400-800ms, maximal in parietal regions, for all target-irrelevant comparisons, corresponding to the P300 (see Table 4.5). This positive cluster was also found for probe-irrelevant comparison. This effect was significant for morally right and morally wrong, but not mock crime, memories in the standard-guilty group, but was present for all memory types in the suppression-guilty group. In the innocent group, there were no positive clusters for probe-irrelevant comparison. Furthermore, for both target-irrelevant and probe-irrelevant comparisons, the P300 related positive cluster was greater for morally wrong than both morally right and mock crime memories, and morally right memories had a greater P300 than mock crime in the standard guilty group. See Table 4.5. for a comparison of results of focal and global analyses results for the P300 measure.

Negative clusters potentially corresponding to the LPN were found for some memory types in some instruction groups in the target-irrelevant comparison, whereas for probe-irrelevant comparisons, no such negative cluster was present.

Table 4.5. Comparison of two analyses strategies for different ERP components from the intrusions analysis

		Target vs. Irrelevant P300		Probe vs. Irrelevant P300	
		Focal Analysis	*Global Analysis*	*Focal Analysis*	*Global Analysis*
Standard-guilty	Mock Crime	✓	✓	✗	✗
	Morally Right	✓	✓	✓	✓
	Morally Wrong	✓	✓	✓	✓
Suppression-guilty	Mock Crime	✓	✓	✓	✓
	Morally Right	✓	✓	✓	✓
	Morally Wrong	✓	✓	✓	✓
Innocent	Mock Crime	✓	✓	✗	✗
	Morally Right	✓	✓	✗	✗
	Morally Wrong	✓	✓	✗	✗

Note. A tick mark indicates that the component was statistically significant in that analysis, whereas a cross indicates that the component was not statistically significant. Please note that the results are summarised as components here only for illustrative purposes, see text and discussion for more detail.

4.2.3 EEG oscillation results

Like in the ERP analysis, cluster analysis of EEG oscillation power differences between conditions was conducted separately for the target vs. irrelevant and probe vs. irrelevant effects. Main effects and interactions between the experimental factors in the 3 (Group; Standard-guilty, suppression-guilty, innocent) x 3 (Memory Type; Mock crime, morally right, morally wrong) design across all scalp electrodes and timepoints with a cluster-based threshold to control for multiple comparisons (see methods section).

Target vs. Irrelevant. Results from the cluster-based permutation tests of this analysis is illustrated in Figure 4.4.

Overall, the results showed an early theta-alpha synchronisation effect followed by a later alpha-beta desynchronisation effect for targets compared to irrelevant items in most comparisons. In the standard-guilty group, target words in both mock crime and morally right memory blocks showed a synchronisation effect in the theta – alpha bands (around 200 to 700ms, cluster was significant for both memory types at p <.01), followed by a desynchronisation effect in the alpha-beta bands from around 400 to 1000ms (mock crime p = .023, morally right p < .001). Target words in the morally wrong block did not have any significant clusters in this group but note that there was a tendency for an effect in the raw power data, but it did not cross the significance threshold for cluster analysis. In both the suppression-guilty and innocent groups, all three memory blocks elicited both the earlier theta-alpha synchronisation and the later alpha-beta desynchronisation effects for targets compared to irrelevants. All synchronisation effects were from around 100 to 700ms, all *ps* < .01. All desynchronisation effects were from around 400 to 1000ms, all *ps* < .01.

Interaction comparisons within each group, such as comparing probe-irrelevant effects between memory types in a group did not reveal significant clusters. Between-group interaction comparisons, such as comparing probe-irrelevant effects for a memory type between groups also did not reveal significant clusters.

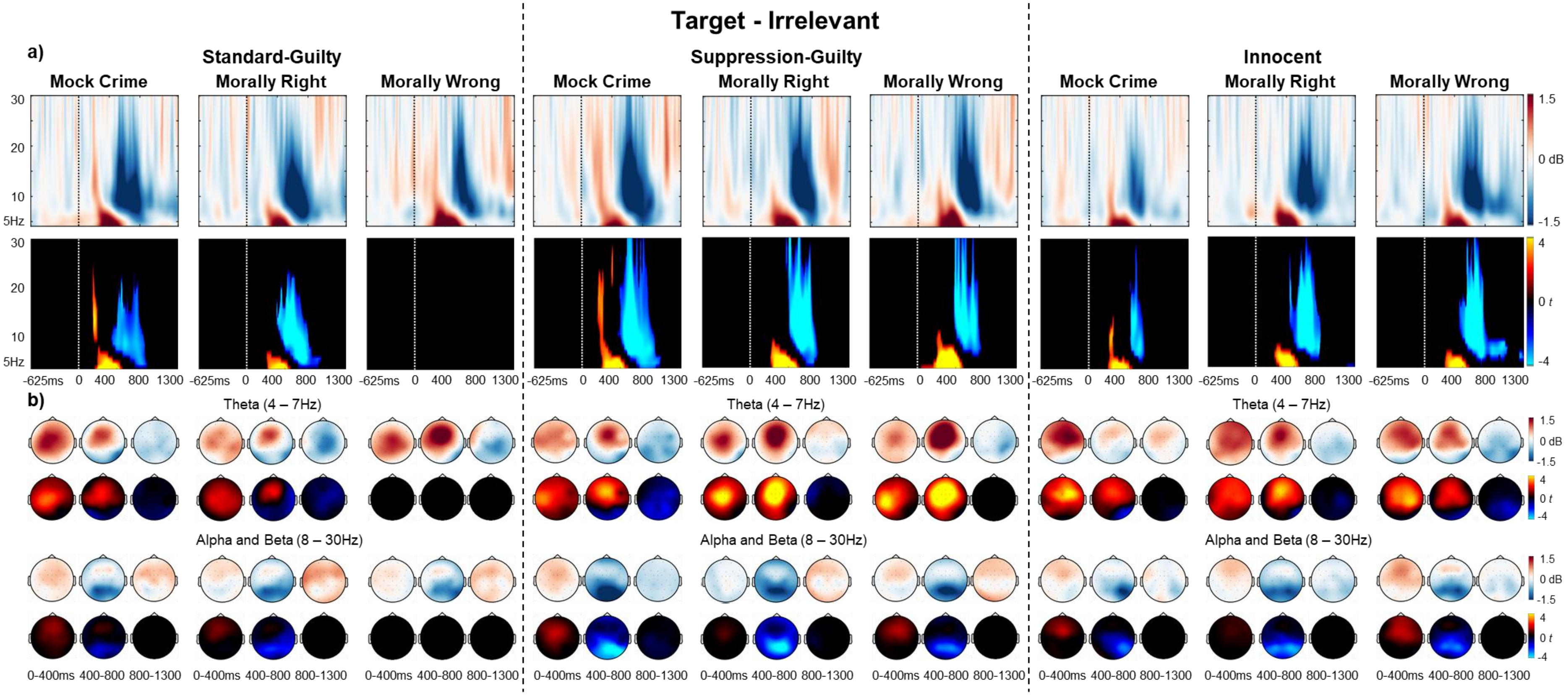


Figure 4.4. Results of the cluster-based permutation tests conducted on time-frequency data of target-irrelevant differences. a) Time-Frequency plots of target-irrelevant differences in power (top row, blue/white/red colourmap) and t-values for the differences (bottom row, cold/black/hot colourmap). These plots are from the Pz electrode. Only significant clusters ($p < .025$ at the cluster level) are shown in the t-statistic time-frequency plots. The pre-stimulus interval was not analysed statistically. b) Topographic maps of target-irrelevant differences in the theta (4-7Hz, top two rows) and alpha/beta (8-30Hz, bottom two rows) frequency bands. Within each frequency band, the top row (blue/white/red colourmap) provides a topographical illustration of raw power differences and the bottom row (cold/black/hot colour map) shows t-values for the differences. Only significant clusters ($p < .025$ at the cluster level) are shown in the t-statistic topographic maps and there is consequently only a partial correspondence between the power and the t-value plots. Note that the same scales (-1.5 to 1.5 dB and -4 to 4 t) have been used for both figures a) and b).

Probe vs. Irrelevant. Results from the cluster-based permutation tests of this analysis is illustrated in Figure 4.5.

The synchronisation and desynchronisation effects found for target vs irrelevants were also found for probe vs irrelevants, but these effects were generally weaker and only present in the guilty groups, in line with predictions. In the standard-guilty group, only morally wrong memories had a cluster that survived the significance threshold after false positive correction. This cluster reflected a significant desynchronisation effect across the whole frequency band (4 – 30Hz), from around 400 to 1300ms (p <.001), whereas a significant synchronisation effect was not found. Similar broadband negative clusters were present in the Suppression-Guilty group for all three memory types. For the mock crime, this cluster was significant from around 500 to 1130ms (p = .003), for morally right memories it was significant around 530 to 1030ms (p = .006), and for morally wrong memories it was significant around 450 to 840ms (p = .003). Morally wrong memories also had a significant theta (4 -7 Hz) synchronisation from around 110 to 720 (p = .01). There were no significant clusters in the innocent group.

There were no significant clusters when comparing the standard-guilty and suppression-guilty groups for any memory type (i.e., no interaction between guilty group and memory type for the probe-irrelevant effects). When comparing the probe-irrelevant oscillation effect between standard-guilty and innocent groups, a positive cluster was found primarily in the theta-band (4-10Hz) around 0 to 650ms, and a negative cluster was found in the 400 to 650ms primarily in the alpha-beta bands (6-30 Hz) for morally wrong memories. This comparison did not reveal any significant clusters for the other two memory types. When comparing the suppression-guilty to the innocent group, all three memory types had a significant negative cluster from around 400ms to 1000ms across the whole frequency range

(4-30Hz). Both morally right and morally wrong memories also had a significant positive cluster from around 0 to 650ms in the theta range (4-12Hz).

Comparing probe-irrelevant effects between memory types within a group did not reveal any significant clusters.

Oscillation analysis summary. In sum, an early theta power synchronisation followed by a later alpha-beta desynchronisation effect was found for target-irrelevant comparison in almost all memory types in all instruction groups. For the probe-irrelevant comparison, the later alpha-beta desynchronisation was more reliably found than the early theta synchronisation. Similar to the P300 ERP results, these oscillation effects were found for some memory types in both "guilty" groups, but no differences were found for any memory type in the innocent group. Specifically, in the standard-guilty group, the later alpha-beta desychronisation was present only for morally wrong memories. In the suppression-guilty groups, the alpha/beta effect was present for all memory type, but morally wrong memories also had the early theta synchronisation. When comparing probe-irrelevant differences between the guilty and innocent groups, all memory types had a significant probe-irrelevant, alpha-beta desynchronisation for suppression-guilty vs. innocent comparison, but for standard-guilty vs. innocent comparison, only morally wrong memories induced this alpha-beta desynchronisation effect.

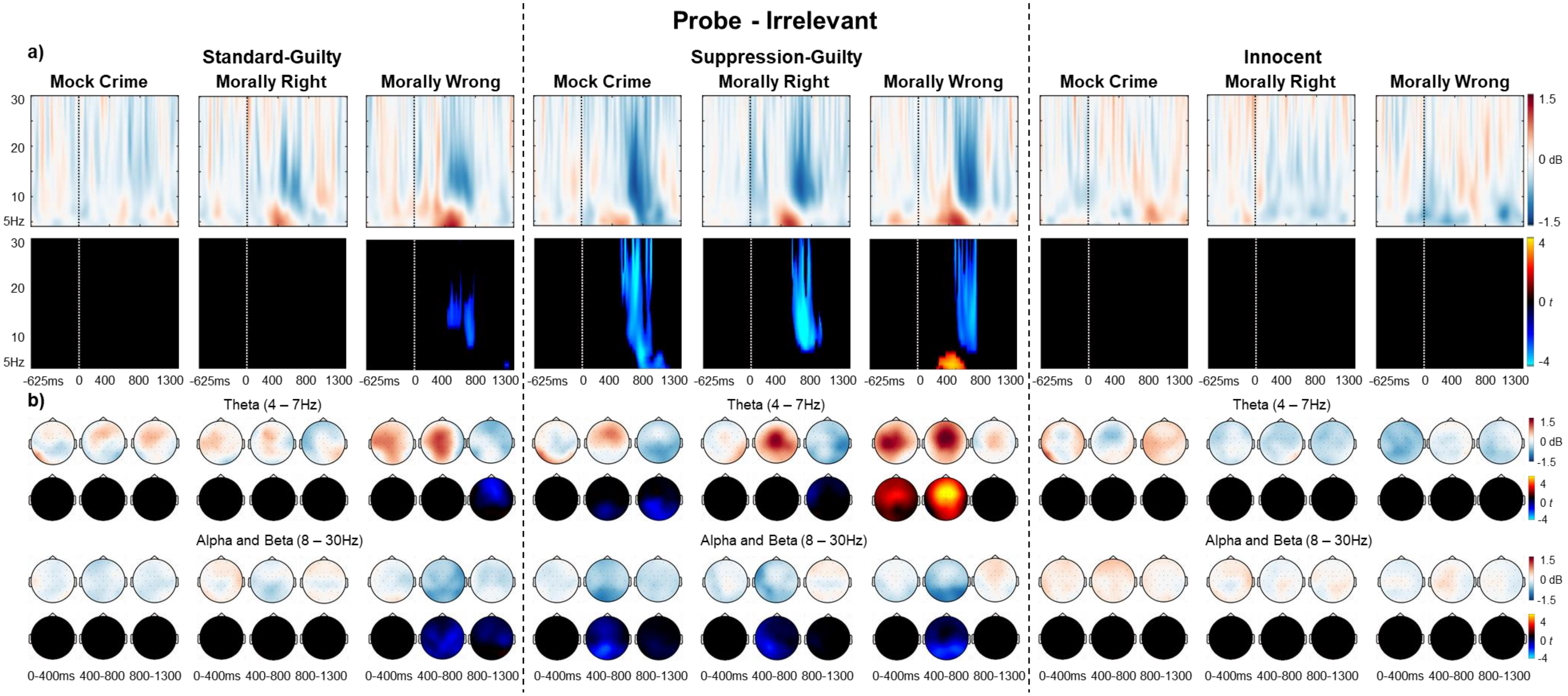


Figure 4.5. Results of the cluster-based permutation tests conducted on time-frequency data of probe-irrelevant differences. a) Time-Frequency plots of probe-irrelevant differences in power (top row, blue/white/red colourmap) and t-values for the differences (bottom row, cold/black/hot colourmap). These plots are from the Pz electrode. Only significant clusters ($p < .025$ at the cluster level) are shown in the t-statistic time-frequency plots. The pre-stimulus interval was not analysed statistically. b) Topographic maps of probe-irrelevant differences in the theta (4-7Hz, top two rows) and alpha/beta (8-30Hz, bottom two rows) frequency bands. Within each frequency band, the top row (blue/white/red colourmap) provides a topographical illustration of raw power differences and the bottom row (cold/black/hot colour map) shows t-values for the differences. Only significant clusters ($p < .025$ at the cluster level) are shown in the t-statistic topographic maps and there is consequently only a partial correspondence between the power and the t-value plots. Note that the same scales (-1.5 to 1.5 dB and -4 to 4 t) have been used for both figures a) and b).

4.3 Discussion

The present study tested the effectiveness of using direct suppression to beat the concealed-information test for morally-relevant autobiographical memories in comparison against mock crime memories, which are used in the standard paradigm in prior literature. Previous research has found that the P300 elicited by probing memories of a lab mock crime can be suppressed in the concealed-information test (Bergström et al., 2013; Hu et al., 2015). Although these memories are rich in sensor-motor detail, such memories could however be more susceptible to suppression than real-world autobiographical memories (Ben-Shakhar & Nahari, 2018; Osugi & Ohira, 2017; Rosenfeld et al., 2017; Rosenfeld, 2020; Ward & Rosenfeld, 2017). The present study therefore is the first investigation to test if direct suppression can reduce the P300 response elicited by autobiographical memories of our past immoral actions, and how this compares to P300s of both mock crime and morally right autobiographical memories. Interesting ERP differences were found between the different memory types, but contrary to predictions, suppression did not reduce P300 effects for any type of memory.

Behavioural performance in the concealed-information tests were largely similar to previous experiments. Participants in this study were highly accurate at responding to stimuli in the CIT. The accuracy rate for responding to targets in this study was lower (around 85% on average) than probes and irrelevants (around 98%); Most previous studies have only one block of the CIT, which takes around 20 minutes to complete (Hellerstedt et al., 2021; Hu et al., 2015; Rosenfeld, 2019), whereas this study had three blocks, leading to longer testing times (around 60 minutes). Participants therefore committed more errors while responding to the target words because it required switching from the prepotent “no” response for probes and irrelevants to a “yes” for the target word. Importantly though, there was no difference in target accuracy between the moral memory type and instruction group conditions, so any differences in neural activity between experimental conditions may not be due to differences in behavioural responses to stimuli. Participants responded faster (around 50 milliseconds on average) for mock crime compared to moral memories. This potentially is because only one word was used as stimuli in the mock crime block whereas two words were used as stimuli in the moral memory blocks, so processing stimuli in the moral blocks would take more time than processing stimuli in the mock crime blocks. Nonetheless, importantly, the accuracy and reaction time was not notably different across the moral memory type and instruction group conditions meaning that any neural differences between these memory type and instruction type was not due to behavioural performance.

Since both, the focal and global analyses of ERPs revealed similar results, these are considered together in the following discussion (see Table 4.5 for a comparison of the results from these two strategies). For the focal analysis, discussion here focuses on the base-to-peak P300 effect but note that the P300-LPN peak-to-peak measure produced the same results, indicating that both measures were equally effective to measure concealed knowledge against some other arguments[7]. The first three first pre-registered hypotheses were largely confirmed by the results, indicating a successful replication of expected ERP effects found in previous research. First (H1), target words elicited greater P300 amplitudes than irrelevants in all groups, for all memory types. The global analyses revealed that this P300 effect was maximal in parietal regions but also present across the whole scalp, and significant from 400-800ms post-stimulus onset (see Figure 4.3). This topographical and temporal distribution of the effect is consistent with previous findings (e.g., Hellerstedt et al., 2021). Therefore,

---

[7] There is some evidence indicating that the peak-to-peak P300 is more effective for detecting concealed information (Soskins, Rosenfeld, & Niendam, 2001; Ward & Rosenfeld, 2017), but these results do not provide evidence for this idea. Against pre-registered expectations (H6), the LPN was not consistently found in the present study and will therefore not be discussed further. Thus, the results discussed here only pertains to the P300 components and apply to both base-to-peak and peak-to-peak measures.

participants in every group successfully elicited P300s in response to the low-probability salient target stimuli regardless of the memory type being tested in each particular block.

If a P300 is elicited by probes in the concealed-information test, it is considered to be an index of automatic recognition of those probes, so that if the concealed information is present in the brain of participants, then it will be reflected by a greater probe-irrelevant P300 (Meijer et al., 2014). The innocent group consisted of participants who did not have any knowledge of the mock crime, nor did they have any stimuli in the moral memory blocks that were relevant to their personal memories. Therefore, as expected (H3), the probe-irrelevant P300 was absent in all blocks of the CIT for innocent participants. In direct comparison, and in line with expectations (H2), participants in the standard-guilty group elicited a significant probe-irrelevant P300 for both morally right and morally wrong memories, indicating that the personal moral memories of participants were successfully detected by the concealed information test (see Figures 4.2 and 4.3, and Table 4.5).

However, against expectations (H2), the mock crime probe did not elicit a significant P300 compared to irrelevants in the standard guilty-group, in either the focal or global analyses (see Figures 4.2 and 4.3 and Table 4.5). An intuitive interpretation would be that participants in this group simply did not recognise the word “ring”. However, two sources of evidence from the present study indicate otherwise. Firstly, in a post-experimental stimuli questionnaire, 100% of the participants in the standard-guilty group reported that the word “ring” was indeed the stolen item and indicated that they had recognised the crime-relevant word in the mock crime block. Secondly, and more notably, participants reported that the memory of the mock crime came to mind more automatically than morally right memories in the CIT, and automaticity was not significantly different from morally wrong memories. Interestingly, although the probe-irrelevant P300 was not found in the group level, at the individual level an equal number of participants in this group were classified as guilty for mock crime memories and morally right memories (8/18), whereas more participants were classified as guilty for morally wrong memories (12/18). A closer inspection of the descriptive statistics (see Figure 4.2) indicates that some participants also had the reversed pattern for the probe vs. irrelevants, which could have influenced group level statistics. Therefore, it is probably not the case that the mock crime probe was not recognised, but rather that group-level P300 was not reliably found for mock crimes.

Critically, the key theoretical hypotheses H4 and H5 predicted that memory suppression would reduce P300 amplitudes in the CIT for all memories albeit at varying degrees, suggesting that even though participants had guilty knowledge, they could appear innocent. In contrast to expectations, all three memory types had a significant probe-irrelevant P300 in the suppression-guilty group (see Figures 4.2 and 4.3 and Table 4.5. Therefore, suppression attempts failed at attenuating the P300 response to probe items, indicating that suppression may not be a successful countermeasure for beating the P300-based concealed information test. This is a direct contradiction to the research indicating the memory suppression may be an effective strategy to appear innocent the P300-based CIT test (Bergström et al., 2013; Hu et al., 2015). Although this finding was unexpected in terms of the pre-registered hypotheses, there are some situations in previous research where memory suppression has failed, and in some cases, also led to *increased* activation of the associated memory (see Stramaccia et al., 2020). It is not entirely clear why suppression was successful in previous CIT research but failed in the present study. Although speculative, I discuss two potential avenues of reasoning that can help understand this finding.

One line of reasoning is that the reduction in P300s for participants who attempted suppression in previous research (e.g., Hu et al., 2015) could be due to certain experimental design parameters of the CIT. It has been argued that these designs are conductive to suppression-induced attenuation of the P300 response (see Rosenfeld, 2020). Their arguments are based on evidence that greater demand on cognitive resources during the task reduces P300 amplitudes (see Polich, 2007). They consider two specific experimental design parameters to provide evidence for their arguments. Firstly, the traditional CIT uses the 3-stimulus-protocol (3SP), which requires participants to make real-time decisions whether the stimulus is a target, probe, or irrelevant. Meixner and Rosenfeld (2011) argue that the cognitive resources required to make such stimulus-based decisions could make it more conducive for suppression to succeed. However, Hu et al., (2015) used the complex-trial-protocol, which does not require stimulus-based decision making in real-time, when the P300 occurs, and still found that suppression was successful. This could be because of the second change in experimental design is the target vs. non-target ratio of trials in the CIT. Hu et al., (2015) had a 50-50 ratio, so participants made target vs. non-target decisions around half of the trials in the study, which would require greater cognitive load than if a lower ratio was used. Recently, Ward and Rosenfeld, (2017) almost identically replicated the experimental design of Hu et al.'s (2015) study, but only changed the 50-50 target vs. non-target ratio to 20-80, thereby reducing the cognitive resources demanded by the task. Interestingly, like in the present study, they also found that the suppression-guilty group elicited the P300 in response to the probe, indicating that suppression was not effective. Therefore, they argue that when cognitive load demanded by the CIT paradigm is lowered, suppression does not affect the P300 response.

In the present study, a similar 20-80 target vs. non-target ratio was used, which may have reduced the cognitive load (compared to Hu et al.'s 2015 study) and led to failed suppression of the P300, as found and explained by Ward and Rosenfeld, (2017). However, the present study used the 3SP which has previously been considered to increase task demand and cognitive load (Meijer et al., 2014). It could be that the 20-80 target vs. target ratio particularly reduces task demand, more than having to make the target decision in real-time, when the P300 is elicited (3SP), than when the target is shown and responded to after the P300 is elicited (CTP). Therefore, this 20-80 target vs. non-target ratio may be important to ensure suppression fails during the CIT. On the other hand, one could also logically argue that because memory suppression involves significant recruitment of top-down inhibitory processes, increase in task demands and cognitive resources could make suppression more difficult and would likely induce conditions where suppression would fail rather than succeed (also noted by Rosenfeld et al., 2017). In support of this claim, cognitive load was found to reduce our ability to suppress memories in the think/no-think task (Noreen & De Fockert, 2017). So even though task demand may reduce P300 amplitudes, ability to engage cognitive control could also suffer, leading to larger rather than reduced P300 amplitudes. Nevertheless, future research needs to clearly tease out the effects of task demands on the P300 in the CIT. This can be achieved by directly comparing combinations of different parameters (3SP vs. CTP, and 50:50 vs. 20:80 target vs. non-target ratios). This will help us understand the clear effect of task demand and cognitive load on suppression ability in the CIT and generally for the P300 response.

Another explanation for the lack of suppression effects could be because the P300 in the CIT is thought to be most closely related to a neurophysiological manifestation of the orienting response, rather than directly indexing retrieval of memories, and this orientation response may not be affected by suppression like memory retrieval (see Chapter 6 for further discussion). Importantly, recent research indicates that the P300-based CIT is sensitive to the saliency of stimuli (klein Selle, Gueta, Harpaz, Deouell, & Ben-Shakhar, 2021), such that highly salient stimuli had larger probe-irrelevant P300 amplitudes compared to low salient stimuli in the CIT. Furthermore, a recent study used the name of the participant as the probe in the CIT and found that suppression attempts *increased* rather than *decreased* the P300 response to the probe, compared to standard-guilty instructions (Rosenfeld et al., 2017).

Interestingly, similar effects were also found in the think/no-think paradigm, such that strong reminders of memories could lead to failed suppression, and can in fact lead to reactivation and rehearsal of the memory and potentially increase the strength of the memory trace (see Engen & Anderson, 2018). Therefore, in the present study, attempts to suppress the memory may have enhanced the orienting response, leading to greater P300 amplitudes.

In relation to previous research, both Bergstrom et al., (2013) and Hu et al., (2015) used only the mock-crime memory in one block of the CIT which is relatively less salient than personal autobiographical memories of past immoral or moral acts. In the present study, a blocked design was used with the three memory types presented in three different blocks. The two moral memory blocks had highly salient stimuli which elicited an orienting response, but the mock crime block stimuli were less salient and referred to everyday objects. Although speculative, this failed suppression may have carried over this effect to mock crime blocks and thus the P300 was found for all three memory types in the suppression-guilty group. Moreover, in the standard group, morally wrong memories had greater probe-irrelevant P300 amplitudes than morally right and mock crime memories, and the target-irrelevant P300s also followed this pattern, therefore providing evidence that the P300 may be especially sensitive to the saliency of probes, in line with klein Selle et al.'s (2021) findings.

Understanding block order effects in the present study could enlighten these interpretations further, for instance, under this account, mock crime memories shown in the first block of the CIT may be more suppressible compared to if where they were shown in the final block in the present study. This could be analysed in the future to further understand why suppression was unsuccessful in the present study but not Hu et al.’s (2014) study. For future research, one potential design would be to consider a larger study, wherein mock crimes and morally wrong memories are tested in a non-blocked fashion. An additional between-groups variable of memory type could be considered to tease out the exact roles of memory suppression for differently salient stimuli. Furthermore, a more applied question would be to simply test autobiographical morally wrong memories in a stand-alone concealed-information test with an *optimal* paradigm that reduces task demands. Therefore, these future studies would hopefully help us clearly identify the role of suppression in beating the concealed-information test.

Nevertheless, these two probable explanations are speculative as the current design is limited in its scope to tease out the role of task-demand, saliency, and recognition – which seem to work in tandem, but may not be equally susceptible to memory suppression to elicit the P300 in response to the probes. Therefore, further research is needed to clearly understand why suppression could be ineffective in the concealed-information test, especially when probed with emotional memories. It may therefore be that a complex combination of different cognitive processes are working to elicit the P300 in response to probes in the concealed information test, which are not all equally susceptible to suppression.

The next hypothesis (H7) predicted that individual guilt classification would have similar patterns of results as the group level results. Indeed, the results were similar to the group level analyses. More participants were classified as guilty in both guilty groups compared to the innocent group for all memory types. However, the classification performance was not significantly better than chance for mock crime memories in the standard-guilty group but was better than chance for the other two memory types in the standard-, and all memory types in the suppression-guilty group, in line with group-level results (see Table 4.4). Additionally, as discussed in Chapter 2, there is a debate regarding the effectiveness of bootstrap vs. permutation methods in individual guilt classification, with some researchers arguing that permutation tests are better than bootstrap, while others argue there is no difference in effectiveness between the two techniques (Rosenfeld & Donchin, 2015; Zoumpoulaki et al., 2015). Using both methods and ROC analysis, the present study indicates that both techniques perform equally well at classifying individuals as either guilty or innocent.

The next two hypothesis were exploratory and concerned the differences in brain oscillation data obtained from the CIT. Expectedly, the oscillation results mirrored the patterns found in ERP data. We predicted greater parietal theta power for probes/targets vs. irrelevants, indexing recognition, similar to the P300 response (H8). Finally, we predicted that attempts at memory suppression would lead to increased right-frontal beta power (H9). The results do not provide clear evidence for these hypotheses. Clearly, suppression was not successful, so the right-frontal beta power was not found in the present study, and H9 was not evidenced (see Figure 4.5). There was also non-significant evidence for H8; rather there was a direct replication of previous research investigating the oscillatory correlates of the concealed-information test which could index guilty knowledge (Hellerstedt et al., 2021). Targets vs irrelevant oscillatory data indicated a strong theta power synchronisation early in the epoch (around 300-700ms post stimulus onset), and this effect seems to be strongest in central regions. This effect was followed by a large alpha/beta power desynchronisation later in the epoch (around 400–1000ms post stimulus onset). Such target-irrelevant differences in power are also found in oscillation data for odd-ball tasks, which the CIT is largely based on (Bernat, Malone, Williams, Patrick, & Iacono, 2007; Keller, Payne, & Sekuler, 2017).

For probe-irrelevant comparisons, the alpha/beta power desynchronisation was more prominent than the theta power synchronisation, and broadly matched the P300 ERP results. Only morally wrong memories had the alpha/beta desynchronisation in the standard-guilty group, whereas all memory types had this effect in the suppression-guilty group. The theta synchronisation was significant only for morally wrong memories in the suppression-guilty group. Therefore, the present results are highly consistent with previous research and provide further evidence for these oscillatory effects as an index of guilty knowledge. The results also resemble the general patterns for the raw oscillatory power found in the “think” condition in experiment 1, wherein memories were voluntarily retrieved, indicating that these early theta and later alpha/beta effects could reflect memory recognition/retrieval processes.

In sum, the present study was able to broadly replicate the increase in P300 response for crime-relevant probe stimulus compared to irrelevant stimuli when participants attempted to conceal their information. Critically though, participants that used direct suppression as a strategy to conceal their memories were not successful as evidenced by extant probe-irrelevant P300 responses for these participants despite their attempts to suppress retrieval. These conclusions were supported by both ERP and oscillation data (see Table 4.5). One reasoning for this finding in the present study could be that the task demands on cognitive load is reduced and led to failed suppression, in line with some arguments. It is also plausible that the P300 reflects the orienting process, which is more difficult to suppress, and increases the P300 amplitude for especially for highly salient stimuli. Future studies are therefore needed to tease out the role of task demands and saliency of stimuli in eliciting the P300 for personal and emotional memories in the CIT. Nevertheless, the present study does provide some evidence for the merit of this concealed-information test in forensic memory detection as it was resistant to suppression of emotional memories.

Chapter 5 – Assessing the effectiveness of different strategies of memory control on emotion regulation using behavioural measures

The first experiment in the thesis (Chapter 3) provided new evidence of the neurocognitive mechanisms of direct suppression, and also indicated that morally wrong memories may be more difficult to suppress. The next experiment reported in Chapter 4 investigated an application of memory suppression in forensic memory detection. The experiment presented here explored another application of autobiographical memory control in everyday life. Two facets of memory control are considered; 1) Can memory control in general be used to achieve emotion regulation? and 2) Which memory control strategies other than direct suppression may be beneficial for emotion regulation?

Firstly, I examined the role of memory control in regulating negative emotions associated with our past immoral acts. As reviewed in Chapter 1 (section 1.3). Memory control processes are considered to be fundamental for emotion regulation and good mental health (Engen & Anderson, 2018; Mary et al., 2020). Recent research has indicated that these processes involves top-down control mechanisms inhibiting mnemonic content present in memory systems in the brain, and that the same control mechanisms also inhibit brain systems involved in emotional processing, especially for emotionally negative memories (Gagnepain et al., 2017). A consequence of such cognitive control was manifested in reduced negative emotional responses to aversive memories. Neural measures of memory control were also related to decreased physiological emotional response to reminders of negative memories (Legrand et al., 2020). Research in this area to date has focused on “offline” consequences of emotion regulation wherein the phenomenological and/or physiological emotional responses to unwanted stimuli are first measured before the memory control task when memories are learned, and then compared to emotions felt after the task when participants are reminded of the memories in a surprise memory test (see Engen & Anderson, 2018). The present study implemented the use of a novel “online” phenomenological measure of emotional responses to reminders to measure the relatively *immediate* effects of memory control on emotion. To the best of my knowledge, this measure is used for the first time in the think/no-think task and provides a new perspective on how emotions are experienced by participants directly after retrieving/controlling their memories. The experimental design of the present study was thus adapted to study these questions.

The design of the present study was very similar to experiment 1 (Chapter 3), with some exceptions. In short, participants first reported autobiographical memories of their past morally relevant acts and created unique reminders for each memory. They also provided self-reports of emotional ratings for each memory. Participants then tried to repeatedly retrieve or avoid retrieval of their memories in response to these reminders in a think/no-think task. Critically, participants rated how they felt after each attempt of retrieval or control. This rating was meant to measure a relatively more intuitive than contemplative emotional response. To this end, participants were asked to consider general feelings of positivity or negativity, and to rate any minor changes in these feelings. Then, after the think/no-think task, a surprise memory test was conducted like in experiment 1, which also included self-reported emotional ratings. Importantly therefore, two behavioural measures were collected in this study: 1) The immediate effect of memory control/retrieval on emotion on a trial-by-trial basis, which provided a measure to track general feelings of positivity or negativity during the think/no-think task; and 2) The subsequent consequence of control/retrieval attempts on emotions measured during the later surprised memory test, when there was no intention to control the memory. The former measure of emotion is new to think/no-think literature, whereas the latter was tested in experiment 1 (Chapter 3), and also in prior research (e.g., Gagnepain et al., 2017; Legrand et al., 2020).

Secondly, three different strategies to control memories were used in the no-think condition of the control task, to test differences in effectiveness of emotion regulation between the three strategies. Three different groups completed the task under different instructions for how to avoid retrieval of no-think memories, whereas all other instructions remained the same. First, the traditional *direct suppression* group were asked to purge memories from awareness without distracting themselves with other thoughts, in line with instructions provided in experiments 1 (Chapter 3) and 2 (Chapter 4) in this thesis. Second, participants were instructed to substitute thoughts of the memory with vivid visual imagery, i.e., *thought substitution*, which has been investigated before in the think/no-think literature (Bergström et al., 2009b; Hertel & Calcaterra, 2005; Joormann et al., 2009; Noreen & Ridout, 2016a; Stramaccia et al., 2020), but has not been directly tested before in the context of autobiographical memory control and emotion regulation. Third, participants were instructed to re-interpret their memories and imagine better alternative outcomes to the existing memory, known as (upward) *counterfactual thinking* (De Brigard et al., 2019).

These memory control strategies can also be mapped on to Gross' emotion regulation theory (see Gross, 2015) that explains similar cognitive techniques to reduce negative emotions. Specifically, the concepts of attentional deployment and cognitive change are particularly relevant to the present experiment. Thought substitution falls into the family of "attentional deployment", which is a conscious, goal-directed shift in attention in order to feel better. The most relevant link is potentially "distraction", where attention is redirected from a negative situation (or memory in this case) to a different situation (stimulus features in direct suppression, or mental imagery in thought substitution). Similarly, direct suppression would also fall under distraction, as the attention is redirected to the features of the stimulus. Alternatively, it could map onto avoidance, a type of situation selection, where there is absolutely no engagement with the unwanted situation (memory in this case). Then, counterfactual thinking potentially directly reflects the family of cognitive change, focussing on reappraisal, where contents of the situation is changed in order to feel better. Therefore, this experiment also has direct implications for emotion regulation theory.

As aforementioned, evidence exists that direct suppression can help regulate unpleasant emotions (Gagnepain et al., 2017; Legrand et al., 2020; Stramaccia et al., 2020; see also experiment 1). There is some debate however whether thought substitution may be more effective for inducing forgetting of unwanted memories than direct suppression, especially under conditions where direct suppression may fail, either due to individual or stimuli-related factors (Hertel & Calcaterra, 2005; Stramaccia et al., 2020), but recent evidence indicates that thought substitution may also fail for people who fail at direct suppression (Noreen & Ridout, 2016b, 2016a). However, these ideas have not been tested with unpleasant autobiographical memories yet, especially those memories that could be particularly difficult to control, like morally wrong memories. Counterfactual thinking has been studied in the context of emotion regulation and was found to help improve negative affect associated with unpleasant memories (e.g., Stanley, Parikh, et al., 2017), but has not been used as a memory control strategy in the think/no-think paradigm. Therefore, the present study aims to directly compare the effectiveness of these strategies on both immediate and consequential measures of emotion regulation.

In terms of immediate “online” effects, it was first predicted that reminders of morally wrong memories would be associated with more negative emotions than morally right memories when participants were thinking of those memories. Based on the previous findings that people can quite successfully prevent intrusions of unwanted memories into awareness (e.g. experiment 1 and previous literature Benoit et al., 2015; Hellerstedt et al., 2016; Levy & Anderson, 2012) it was expected that all strategies would be somewhat effective at online emotion regulation, which would manifest as more positive emotions in response to reminders of morally wrong memories when participants were avoiding retrieval compared to thinking of such memories. For morally right memories, this pattern was predicted to reverse so that positive emotions associated with thinking of these memories would reduce if participants successfully stopped those memories from coming to mind. That is, successful memory control predicts relatively more neutral emotional responses in the no-think condition and relatively more extreme emotional responses in the think condition, with those extreme responses being either more positive or negative depending on the type of memory. However, the degree of emotion regulation effectiveness could differ across strategies, which would be indicated by this predicted pattern being more pronounced for those strategies that were more successful.

Memory control-induced forgetting was also measured in the present study using the same analysis strategy as in experiment 1. Memory descriptions were collected both before and after the think/no-think task for each strategy group, and the experimenter and independent coders rated the three memories on specificity and similarity. Based on results from experiment 1, I predicted that descriptions of morally wrong memories will be less similar after compared to before the think/no-think task than morally right memories, regardless or retrieval or control, indicating that descriptions of morally wrong memories may be overall more susceptible to change. Based on experiment 1 findings, I also predicted that these memory descriptions would not be modulated by think/no-think condition when participants used a direct suppression strategy, suggesting that direct suppression is not an effective strategy for inducing forgetting of these morally relevant memories (at least as implemented in this design). If thought substitution or counterfactual thinking are more effective strategies than direct suppression for inducing forgetting of avoided memories, then similarity and/or specificity of memory ratings should be particularly low for memories in the no-think condition in these strategy groups. This effect may also be modulated by the moral nature of the memories, in that morally wrong memories may be particularly intrusive and may therefore be either more difficult to forget (e.g. Stanley et al., 2018), or alternatively may be more susceptible to inhibition due to their intrusiveness (e.g. Anderson & Hanslmayr, 2014) and may therefore be more likely to be impaired on the final test.

5.1 Method

5.1.1 Participants

One hundred and thirty-five students at the University of Kent completed the study in exchange for course credits. Participants were assigned to groups in a random order, as different experimenters collected data for different groups. Fifty-three participants were in the direct suppression group (M_{age} = 19.81 years, SD_{age} = 3.70), 35 participants in the thought substitution group (M_{age} = 19.34 years, SD_{age} = 2.07), and 47 participants in the counterfactual thinking group (M_{age} = 19.19 years, SD_{age} = 1.28). Although it was originally planned to collect an equal number of participants across groups, doing so was not possible due to the COVID-19 pandemic. The design of the study was exploratory, with no published research using an online emotion scale, so the effect sizes was not clear prior to beginning data collection. Hence, I collected as much data as resources permitted, with the aim of collecting around 40-50 participants.

5.1.2 Design, Materials and Procedure

The procedure was nearly identical to the think-no/think EEG study reported in chapter 3. The present experiment was conducted in two sessions, 24 hours apart. The main differences between the two experiments occurred in the second session, in the think/no-think task.

Session 1. The first session was identical to experiment 1. This session was conducted in a lab using the online Qualtrics survey software. Participants were instructed to think of 22 different autobiographical memory one at a time (10 morally wrong, 10 morally right, one birthday, and one holiday), and type a description in a text box provided to them on a computer screen using a keyboard. To aid recollection of morally wrong actions, they were instructed to think of memories where they lied, cheated, physically or emotionally harmed someone, or any other act they considered was morally wrong. Similar examples were also provided for remembering past morally right actions, such as memories where they were truthful, helped someone physically or emotionally, or other morally right actions. Participants had three minutes to think of and write about a specific memory in two to three sentences by describing their actions, the persons involved, the location, and how they felt. They were instructed to avoid writing about events that easily blend with other memories and to describe each event in as much detail as possible.

After describing each memory, participants were instructed to think of a unique and specific personal title (henceforth referred to as the "cue" in this thesis) that would help them recollect the same exact memory in the next session when cued with the cue. They were also instructed to avoid cues that could evoke multiple memories. The cues were split into a set of cues for morally wrong memories, one set of cues for morally right memories, and one set of filler cues.

After participants generated the cue, they rated the memory for age (in years), vividness (How well do you remember the event? *1 = not well at all, 5 = extremely well),* intentionality (How intentional were your actions? *1 = not intentional at all, 5 = extremely intentional*) and morality (How morally right or morally wrong were the actions you performed during the event? *1 = very morally wrong, 7 = very morally right).* These phenomenological ratings have been used previously in moral autobiographical memory research (see Stanley et al., 2017). The participants were then asked to report emotional affect and valence using two scales: a modified version of the I-PANAS-SF: "Indicate the extent to which you feel this way at the present moment of the event" (e.g., Guilty, *1 = very slightly or* *not at all, 5 = extremely*); and SAM, which is a pictorial scale that measures emotional pleasure and arousal. See Appendix A.1. for an example of the questionnaire.

Session 2. In session two, they completed the think/no-think task. Similar to experiment 1, the cues that the participants generated in session one were used as stimuli in this session to cue the autobiographical memories. To that end, an initial cue-test phase was conducted to ensure that participants could remember the autobiographical memory associated to the cue, similar to experiment 1 (see Appendix A.2.).

Strategy-practice phase. Unique to this experiment, a "strategy-practice" phase was introduced after the cue-test phase. In this phase, a cue appeared for 10s, preceded by a fixation cross for 1s. Only cues assigned to the no-think condition were presented. Each group was given different instructions to control their memories:

The thought substitution group were asked to imagine themselves in a mental safe space. Participants were provided three different examples of "safe spaces". For instance, if the safe space was being at the beach, participants were asked to try hearing and seeing the waves crashing, seagulls flying, feeling the warmth of the sun, etc. Importantly, they were instructed that instead of thinking about the memory associated to the cue, they should just keep thinking about the safe place and keep their thoughts completely focused on the safe place while the cue is on the screen. The other two safe spaces considered being under a starry sky, or in a forest. Participants chose one safe space for the duration of the experiment and were asked to imagine experiencing that safe space each time a cue came up on the screen in the "strategy-practice" phase. The counterfactual group were instructed to imagine a more positive alternative outcome to the memory associated with the cue on the screen. For instance, if the original memory refers to when they bullied their friend and it led to a big fight, then they should imagine that the both of them reconciled instead of fighting. Importantly, participants were instructed to imagine that same counterfactual scenario each time that particular cue appeared throughout the experiment. The direct suppression group practiced memory suppression for each cue using the same exact instructions as in previous experiments in this thesis (see prior chapters). Next, they extensively practiced the think/no-think task. The practice sessions were the same as experiment 1 (see Appendix A.2.)

Think/no-think phase. Similar to experiment 1, the 10 morally wrong and 10 morally right cues were pseudo-randomly assigned to the think and no-think conditions in equal proportions, resulting in five cues in each condition (morally wrong think, morally wrong no-think, morally right think, and morally right no-think).

In each trial, a white fixation cross was first presented on a black background for 1000ms, followed by a cue presented in either green or red for 5000ms. The presentation time of the cue was increased by 1000ms to provide more time for controlling their memories. As usual, participants were asked to retrieve the associated memory if it was shown in green, but if it was shown in red (no-think), they had to either directly suppress, imagine a safe space (thought substitution), or think of the counterfactual event for that particular memory depending on their group.

After each cue, a black screen was shown for 200ms, followed by an emotion rating scale for 1500ms, which showed “1 2 3”, wherein participants responded using keyboard button press (*1 = Negative, 2 = Neutral, 3 = Positive*). This scale was similar to that used to measure intrusions in experiment 1, but in this experiment, it was changed to measure emotions that participants felt immediately after each retrieval/control attempt. Participants were asked to withhold responses until the rating scale was displayed, and then indicate their response by pressing buttons on the keyboard. Participants were instructed to indicate how they felt at that moment, and to indicate any minor fluctuations of emotion at that point of time. This rating was meant to be relative, so they had to press the corresponding button if they felt relatively more positive and negative and press neutral only if they were unable to tell how they felt at that time.

The software was programmed to first present all 20 cues pseudo-randomly, one at a time, then after all cues were presented, the order of presentation was pseudo-randomised again and presented one at a time again, and this process was repeated 12 times, leading to a total of 20 memories x 12 repetitions = 240 trials. Participants were given a short break after each set of 40 trials. The software was also programmed to present cues in one condition (either think or no-think) not more than three times in a row.

5.3 Results

The analysis strategy used here was nearly identical to the behavioural analysis strategy of experiment 1 as reported in chapter 3. The immediate effect of memory control on emotions associated to the different memory types, compared across the first and second halves of the TNT task was investigated first (in contrast to experiment 1, which investigated intrusion ratings with the same analysis strategy of comparing across think/no-think repetitions). Then, I report the change in autobiographical memory descriptions from time 1 to time 2 as an effect of the think/no-think task as quantified by experimenter ratings. Finally, the change in participants' ratings of autobiographical memory characteristics between time 1 and time 2 (i.e., before and after the TNT task) as an effect of think/no-think condition is described.

5.3.1 Real-time emotion ratings

Trial-by-trial emotion ratings provided by participants during the think/no-think task were extracted for think/no-think and morally right/wrong conditions separately for the first (repetitions 1-6) versus second block (repetitions 7-12) of trials. The effect of memory control strategy on emotions over repeated attempts at retrieval/suppression of memories was also tested. Responses recorded using the emotion ratings in the TNT task were divided into different conditions and averaged to produce a mean emotion rating per condition (see Appendix C.1. for means and standard deviations across all experimental conditions). This measure was thus a continuous variable ranging from -1 (Negative), 0 (Neutral), +1 (Positive). See Figure 5.1 for an illustration of the descriptive statistics from this analysis when collapsed across the different strategy groups, which clearly shows that morally right memories were generally associated with positive emotions and morally wrong memories were associated with more negative emotions as would be expected. When participants were attempting to stop retrieval, their emotional responses generally became more neutral as was also expected.

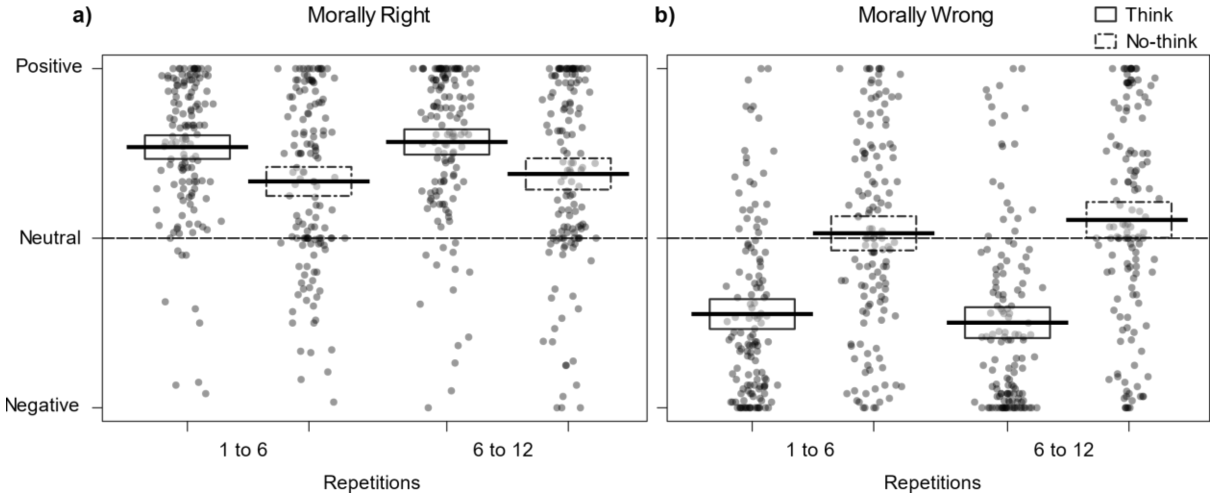


Figure 5.1. Average emotion ratings during the TNT trials, compared across instruction type and block for a) morally right and b) morally wrong memories, but collapsed across strategy group. The scatter dots show the mean emotion rating for a condition of each individual. The thick lines show the condition means and the boxes depict the 95% confidence interval of the means.

To test whether the online emotion measure was significantly modulated by the independent variables, a 3 (Group: Direct suppression, thought substitution, counterfactual thinking) x 2 (Instruction Type: Think, no-think) x 2 (Memory Type: Morally right, wrong) x 2 (Block: First half, second half) mixed-design ANOVA was conducted with this measure as a dependent variable. A significant main effect of instruction type was found. Participants reported feeling overall more positive when not-thinking ($M = .21$, $SD = .49$) than thinking ($M = .04$, $SD = .31$) of their memories; $F(1,132) = 20.42$, $p < .001$, partial $\eta^2 = .13$. Main effects of memory type and block were also significant. As one would expect, participants felt more positive after seeing reminders for morally right memories ($M = .45$, $SD = .32$) than morally wrong memories ($M = -.20$, $SD = .39$); $F(1,132) = 258.75$, $p < .001$, partial $\eta^2 = .66$. Participants also felt better overall in the second ($M = .13$, $SD = .29$) than in the first half ($M = .11$, $SD = .28$) of the task, $F(1,132) = 4.74$, $p = .037$, partial $\eta^2 = .035$. These latter two main effects did not interact with group (both Fs < .99, ps > .37), but there was a significant interaction between group and instruction type (reported in detail at the end of this section).

There were significant interactions between instruction type and memory type, $F(1,132) = 115.89$, $p < .001$, partial $\eta^2 = .47$, and between instruction type and block $F(1,132) = 18.06$, $p < .001$, partial $\eta^2 = .13$, but the memory type x block interaction was not significant $F(1,132) = 1.95$, $p = .17$, partial $\eta^2 = .01$. Importantly, there was a significant three-way instruction-type x memory-type x block interaction, $F(1,132) = 12.33$, $p = .001$, *partial* $\eta^2 = .09$. None of these effects interacted with group (all Fs < 2.38, ps > .097). Therefore, intention to retrieve or avoid retrieval affected the emotions experienced by participants differently for morally right and wrong memories, and this changed over time in the think/no-think task.

A follow up 2 (Instruction Type) x 2 (Block) repeated measures ANOVA was conducted for morally right and morally wrong memories separately to investigate how participants felt over repeated attempts at thinking vs. not-thinking about the different types of memories. There were main effects of instruction type for both types of memories. Overall, regardless of strategy used to prevent retrieval, participants felt better after not thinking ($M = .06$, $SD = .59$) than thinking ($M = -.48$, $SD = .51$) about morally wrong memories; $F(1,132) = 85.88$, $p < .001$, partial $\eta^2 = .39$, whereas they felt worse after not thinking ($M = .35$, $SD = .50$) than thinking about morally right memories ($M = .55$, $SD = .41$), $F(1,132) = 11.62$, $p = .001$, partial $\eta^2 = .08$. Memory control thus makes people’s emotional experiences more neutral, which helps reduce negative emotions associated with morally wrong memories, but also reduces positive emotions associated with morally right memories.

A main effect of block was found for morally right memories; participants felt better in the second than in the first half of the task, $F(1,132) = 5.99$, $p = .016$, partial $\eta^2 = .043$. The interaction between TNT instruction type and block was not significant, suggesting that this emotional change across repetitions was not dependent on whether participants were retrieving or stopping retrieval of those memories, *F*(1,132) = 0.38, *p* = .056, partial η^2 = .003.

In contrast, for morally wrong memories, there was a significant TNT instruction type x block interaction, *F*(1,132) = 23.85, *p* < .001, partial η^2 = .15. The main effect of block was not significant, *F*(1,132) = 0.98, *p* =.32, partial η^2 = .007. To follow up, a repeated-measures ANOVA testing the difference in emotion ratings between first and second half of the task was conducted separately for think and no think conditions for morally wrong memories. Interestingly, results revealed that participants felt worse in the second half than in the first half of the task while thinking about morally wrong memories, *F*(1,132) = 7.49, *p* = .007, partial η^2 = .053, whereas they felt better in the second half than in the first half of the task while not-thinking about these memories, *F*(1,132) = 14.79, p <.001, partial η^2 = .10. These results thus suggest that repeated attempts to control morally negative memories were successful in improving the emotional response elicited by cues to those memories, regardless of the memory control strategy.

The only differences between strategy groups were a significant instruction type x group interaction effect, *F*(2,132) = 29.16, *p* < .001, partial η^2 = .31. Participants felt better when not-thinking than thinking of their memories (overall, collapsed across morally right and wrong memories) for both counterfactual thinking ($M_{No\text{-}think}$ = .48, $SD_{NoThink}$ = .45, M_{Think} = -.04, SD_{Think} = .27), *F*(1,46) = , *p* < .001, partial η^2 = .48, and thought substitution ($M_{No\text{-}think}$ = .39, $SD_{No\text{-}think}$ = .42, M_{Think} = .06, SD_{Think} = .25), *F*(1,34) = 14.62, *p* < .001, partial η^2 = .30, groups. Interestingly, direct suppression affected emotions in the opposite direction; participants felt overall worse after suppressing their memories ($M_{No\text{-}think}$ = -.15, $SD_{No\text{-}think}$ = .34) than thinking of them (M_{Think} = .09, SD_{Think} = .35), *F*(1,52) = 12.61, *p* < .001, partial η^2 = .20. This effect seemed to be primarily driven by how memory control affected morally right memories. Closer inspection of the descriptive statistics (see Figure 5.2) indicates that participants felt less positive after direct suppression of morally right memories, whereas both thought substitution and counterfactual thinking did not reduce positive affect related to morally right memories. Furthermore, although direct suppression improved negative emotions for morally wrong memories compared to retrieving the memories, this difference was numerically smaller than the effect found for thought substitution and counterfactual thinking.

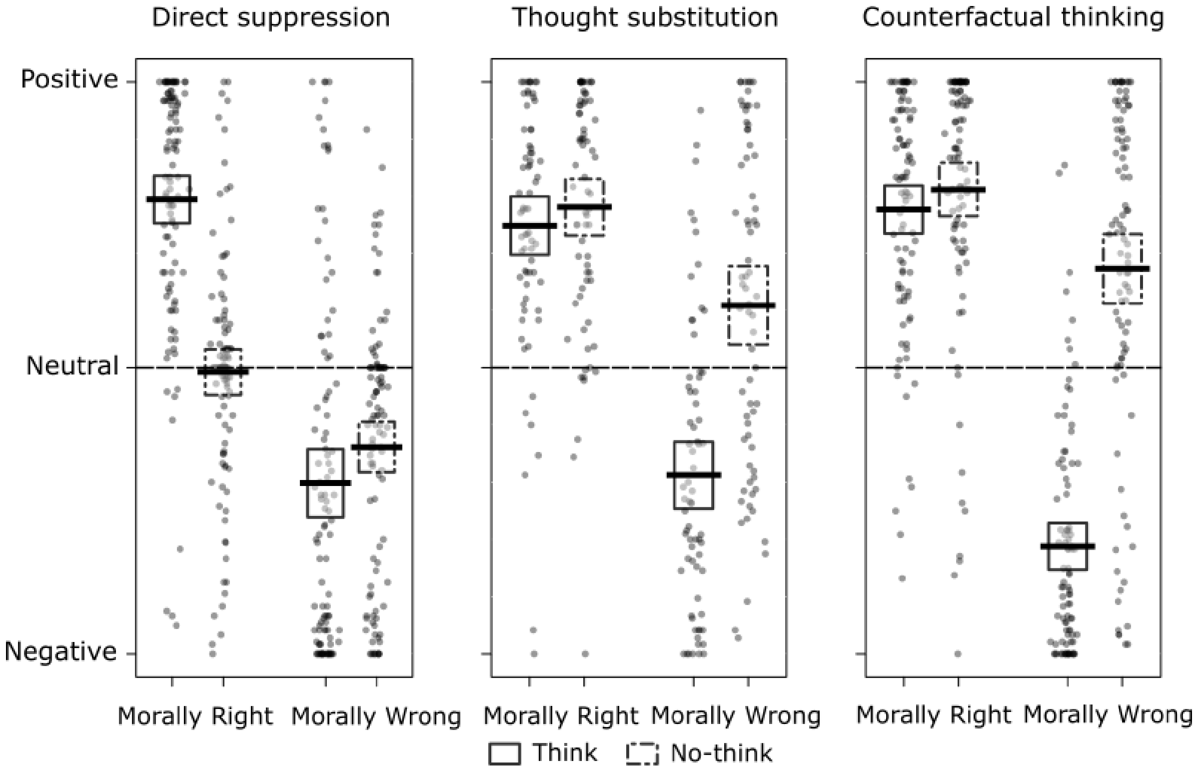


Figure 5.2. Average emotion ratings during the TNT trials, compared across moral memory type and TNT instruction type, for each strategy group. The scatter dots show the mean emotion rating for a condition of each individual. The thick lines show the condition means and the boxes depict the 95% confidence interval of the means.

In sum, comparing online emotion ratings across the two blocks showed that emotions became more neutral as participants repeatedly performed the think/no-think task, in that participants felt less positive after repeatedly thinking of their past morally right acts, but contrastingly, participants felt more positive after repeated attempts at memory control of morally wrong memories. These effects were found regardless of memory control strategy. Participants reported feeling more positive while thinking compared to not-thinking of morally right memories, whereas controlling memories of morally wrong acts led to more positive feelings than thinking of them. The only differences between strategy groups was mainly driven by think-related effects for morally right memories. Both thought substitution and counterfactual thinking groups reported feeling better while not-thinking than thinking about their memories, but the direct suppression group felt worse while not-thinking than thinking about their memories.

5.3.2 Autobiographical memory descriptions analysis

Using the same coding instructions as chapter 3, the experimenter rated all the written memory descriptions that were provided before and after the think/no-think task, and independent coders rated 50% of the memories for this study. There was a .73 (ranging from .69 - .75) correlation between the experimenter and independent ratings. See Appendix C.2. for average of independent ratings. The results broadly replicated findings from chapter 3, but there were some differences between groups in effect sizes.

Similarity. A 3 (Group: Direct suppression, thought substitution, counterfactual thinking) x 2 (Instruction Type: Think, no-think) x 2 (Memory Type: Morally right, morally wrong) mixed ANOVA was first conducted on the ratings of how similar the memory descriptions were that participants provided before and after the think/no-think task (see Table 5.2), to test for general changes in memory as a function of strategy, instruction, and memory type.

Table 5.1. Average similarity ratings across instruction and memory types for each strategy group.

	Direct Suppression		Thought Substitution		Counterfactual Thinking	
	Right	Wrong	Right	Wrong	Right	Wrong
Think	4.06(.41)	3.83(.49)	4.11(.54)	3.91(.51)	4.15(.55)	3.73(.70)
NoThink	4.10(.40)	3.72(.52)	4.10(.38)	3.86(.64)	4.12(.56)	3.61(.69)

Note. Similarity was rated on a 5-point scale (1 = Not at all similar, 5 = Extremely similar). Standard deviation is reported in brackets.

Memories in the no-think condition were numerically rated as less similar than memories in the think condition, but this difference was not significant as a main effect, $F(1,132) = 1.92$, $p = .17$, nor did this think/no-think instruction effect interact with group $F(2,132) = .39$, $p = .68$. There was a highly significant main effect of memory type, with morally right memory descriptions on average being more similar across times 1 and 2 than the morally wrong memories, $F(1,132) = 90.54$, $p = <.001$, partial $\eta^2 = .41$, and a significant memory type x group interaction was also found, suggesting that this difference between right and wrong memory descriptions varied as a function of the strategy people used to prevent retrieval, $F(2,132) = 4.14$, $p = .018$, partial $\eta^2 = .059$. Neither the instruction type x memory type not the instruction type x memory type x group interactions were significant (both *Fs* < 2.26, *ps* > .14).To follow up on the memory type x group interaction, three 2 (Group) x 2 (Memory Type) mixed ANOVAs were conducted to test for pairwise group differences in memory change as a function of whether those memories were of morally wrong or right events. The results indicated that the difference in similarity between right and wrong memories was greater for the counterfactual thinking group compared to both direct suppression ($F(1,98) = 4.25$, $p = .042$, partial $\eta^2 = .042$) and thought substitution ($F(1,80) = 7.35$, $p = .008$, partial $\eta^2 = .084$) groups. However, when comparing the direct suppression and thought substitution groups, the interaction effect was not significant, suggesting that differences in similarity between right and wrong memories were more equivalent across these two groups, $F(1,86) = 1.07$. $p = .304$.

Follow up repeated measures ANOVAs were conducted to test whether the right and wrong memories differed on similarity scores within each group. In all groups, morally wrong memories were rated as less similar than morally right memories: Direct suppression: $F(1,52) = 35.37$, $p < .001$, partial $\eta^2 = .41$; thought substitution: $F(1,34) = 7.88$, $p = .008$, partial $\eta^2 = .188$; counterfactual thinking $F(1,46) = 74.62$, $p < .001$, partial $\eta^2 = .619$. Thus, the significant interaction with group for this effect was due to a difference in effect sizes between the three groups with the counterfactual thinking group showing a particularly large difference in similarity between morally right and wrong memories, but all groups showed the same directional pattern.

Specificity and valence. Next, a 3 (Group: Direct suppression, thought substitution, counterfactual thinking) x 2 (Instruction Type: Think, no-think) x 2 (Memory Type: Morally right, morally wrong) x 2 (Administration Time: Before TNT, after TNT) mixed ANOVA was conducted for ratings or specificity and valence differences between pre- and post-TNT descriptions (see Table 5.3 for raw scores). This analysis aimed to investigate if the specificity and valence of memory descriptions changed as a result of the think/no-think instruction, the strategy people used to avoid retrieval for no-think items, and the type of memory they were cued to recall.

Table 5.2. Average ratings of specificity and valence across instruction and memory types. Ratings were provided by the experimenter.

	Direct Suppression				Thought Substitution				Counterfactual Thinking			
	Right		Wrong		Right		Wrong		Right		Wrong	
	Pre-TNT	Post-TNT	Pre-TNT	Post-TNT	Pre-TNT	Post-TNT	Pre-TNT	Post-TNT	Pre-TNT	Post-TNT	Pre-TNT	Post-TNT
	Specificity											
Think	3.05(.56)	2.96(.52)	2.98(.62)	2.99(.57)	2.90(.71)	2.81(.66)	2.72(.72)	2.80(.59)	2.92(.74)	2.90(.71)	2.95(.67)	2.88(.70)
NoThink	2.97(.57)	2.98(.57)	3.01(.56)	3.02(.53)	2.37(.73)	2.88(.77)	2.80(.77)	2.79(.58)	2.95(.67)	3.03(.68)	2.88(.65)	2.93(.63)
	Valence											
Think	4.23(.35)	4.20(.30)	3.53(.49)	3.56(.49)	4.16(.35)	4.22(.41)	3.61(.47)	3.66(.34)	4.10(.34)	4.12(.31)	3.61(.37)	3.52(.43)
NoThink	4.21(.36)	4.19(.41)	3.46(.56)	3.49(.45)	4.22(.51)	4.23(.36)	3.64(.42)	3.63(.48)	4.13(.30)	4.13(.30)	3.65(.40)	3.57(.43)

Note. Specificity was rated on a 5-point scale (1=Not at all specific to 5 = Extremely specific), emotional valence was measured on a 7-point scale (1=Very Negative, 4 = Not negative or positive, 7 = Very Positive). Standard deviation is reported in brackets.

No significant effects of any factors were found for the specificity ratings, (*Fs* < 1.88, *ps* > .17). For valence, the main effects of instruction type and its interaction with administration type and group were not significant (all *Fs* < 2.18, *ps* > .12). However, there was a main effect of memory type, wherein morally wrong memories were more negatively valenced than morally right memories, F (1,132) = 134.37, $p < .001$, partial $\eta^2 = .504$. A significant group x memory type x administration time interaction was also found: F (2,132) = 3.33, $p = .039$, partial $\eta^2 = .048$, indicating that strategies affected the change in valence as a result of think/no-think task and this change was different across the two memory types. All other effects of memory type were not significant (all *Fs* < 1.07, *ps* > .348).

To follow up on this interaction, three 2 (Group) x 2 (Memory Type) x 2 (Administration Time) mixed ANOVAs were conducted to compare the memory type x administration time interaction between the three groups. This interaction was significant for direct suppression and counterfactual thinking groups, F (1,98) = 5.88, $p = .017$, partial $\eta^2 = .057$; but for both direct suppression and thought substitution, and thought substitution and counterfactual thinking comparisons, the memory type x administration time interactions were not significant (both *Fs* < 1.64; *ps* > .20). Therefore, the change in valence due to the think/no-think task was different across memory types, and this effect differed across direct suppression and counterfactual thinking groups only.

Then, a 2 (Memory Type) x 2 (Administration Time) repeated measures ANOVA was conducted in both the direct suppression and counterfactual thinking groups. In the direct suppression group, neither the main effect of administration time nor the interaction between memory type and administration time was significant (both *Fs* < 2.29, *ps* > .14). In the counterfactual thinking group, a trend-level effect was found for the main effect of administration time, such that memory descriptions were more negatively valenced after compared to before the think/no-think task: *F*(1,46) = 3.47, *p* = .069, partial η^2 = .07. The interaction between memory type and administration time was also trend-level significant: *F*(1,46) = 3.41, *p* = .07, partial η^2 = .069. As these interactions were not significant, no follow up tests were conducted. Essentially, compared to direct suppression group, the counterfactual thinking group had more negatively valenced descriptions after compared to before the think/no-think task, especially for morally wrong memories.

5.3.4 Phenomenological and emotional characteristics of autobiographical memories

The same self-reported memory characteristics (collected at the time of memory generation and then in the final test) were analysed as in previous studies in the thesis: Memory age, vividness, intentionality, morality, guilt, shame, pleasure, and arousal. Results indicated a successful replication of the other two studies.

Like in experiment 1, an initial supplementary analysis (not reported here, see Appendix C.3.) tested differences in self-reported characteristics between morally wrong and right memories before the think/no-think task. These results verified that memories were very similar across the think/no-think conditions at the time they were assigned to these conditions (i.e., that the random assignment had succeeded), regardless of the instruction group. The results also replicated previous findings that morally wrong memories are rated to have occurred in the more distant past (Escobedo & Adolphs , 2010; chapters 3 and 4), are less vivid (Kouchaki & Gino, 2016; chapters 3 and 4), and their actions as less intentional (chapters 3 and 4) than morally right memories.

For the key analyses, a difference score was computed between self-reports collected in the first session before the think/no-think task, and the second session after the think/no-think task, and this was used as a dependent variable to test the effect of the think/no-think manipulation on phenomenological characteristics of memories. A 3 (Strategy Group: Direct suppression, thought substitution, counterfactual thinking) x 2 (Instruction Type: Think, no-think) x 2 (Memory Type: Morally right, morally wrong) mixed ANOVA was conducted with the difference scores (see Table 5.4 for results).

Table 5.3. Results of omnibus ANOVAs, assessing the effects of the TNT task on phenomenological and emotional characteristics of different memory types.

	Phenomenological memory characteristics											
	Memory Age			Vividness			Intention			Morality		
	F	*p*	η^2_p	*F*	*p*	η^2_p	*F*	*p*	η^2_p	*F*	*p*	η^2_p
IT (T vs. NT)	0.52	0.47	0.00	1.22	0.27	0.01	0.39	0.53	0.00	0.31	0.58	0.00
ITxGroup	1.12	0.33	0.02	0.66	0.52	0.01	0.61	0.55	0.01	0.18	0.84	0.00
MT (MW vs. MR)	**6.17**	**0.014**	**0.045**	**4.92**	**0.028**	**0.036**	1.93	0.17	0.01	2.20	0.14	0.02
MTxGroup	0.55	0.58	0.01	1.62	0.20	0.02	1.04	0.36	0.02	0.29	0.75	0.00
ITxMT	0.01	0.93	0.00	0.22	0.64	0.00	0.17	0.68	0.00	1.17	0.28	0.01
ITxMTxGroup	1.36	0.26	0.02	0.13	0.88	0.00	0.55	0.58	0.01	0.70	0.50	0.01

	I-PANAS-SF						*SAM*					
	Ashamed			Guilty			Pleasure			Arousal		
	F	*p*	η^2_p	*F*	*p*	η^2_p	*F*	*p*	η^2_p	*F*	*p*	η^2_p
IT (T vs. NT)	0.30	0.58	0.00	0.90	0.34	0.01	0.04	0.85	0.00	0.10	0.75	0.00
ITxGroup	0.39	0.68	0.01	0.77	0.47	0.01	0.85	0.43	0.01	2.15	0.12	0.03
MT (MW vs. MR)	**25.30**	**<.001**	**0.161**	**23.14**	**<.001**	**0.149**	1.26	0.26	0.01	0.10	0.75	0.00
MTxGroup	0.26	0.77	0.00	0.27	0.76	0.00	0.38	0.68	0.01	2.15	0.12	0.03
ITxMT	0.06	0.82	0.00	0.22	0.64	0.00	1.96	0.16	0.02	0.01	0.92	0.00
ITxMTxGroup	0.50	0.61	0.01	2.99	0.054	0.04	0.16	0.85	0.00	0.95	0.39	0.01

Note. IT = Instruction Type, MT = Memory Type, T = Think, NT = No-Think, MW = Morally Wrong, MR = Morally Right. h^2_p = partial eta sq. (effect size). I-PANAS-SF = International – Positive and Negative Affect Scale – Short Form, SAM = Self-Assesssment Manikin. I-PANAS-SF = International – Positive and Negative Affect Scale – Short Form. Significant effects ($p < .05$) are shown in bold. *N = 135*.

The results showed that the strategy participants used did not interact with any other factor in influencing changes in self-reported characteristics of the memory. Replicating results from experiment 1 (chapter 3), neither the main effects of think/no-think instruction type, nor think/no-think instruction type x right/wrong memory type interactions were significant for any measure. However, there were significant main effects of memory type for some measures: Participants reported feeling less ashamed and guilty for morally wrong memories after the TNT task, but there were no such changes in guilt and shame for morally right memories (see Figure 5.3).

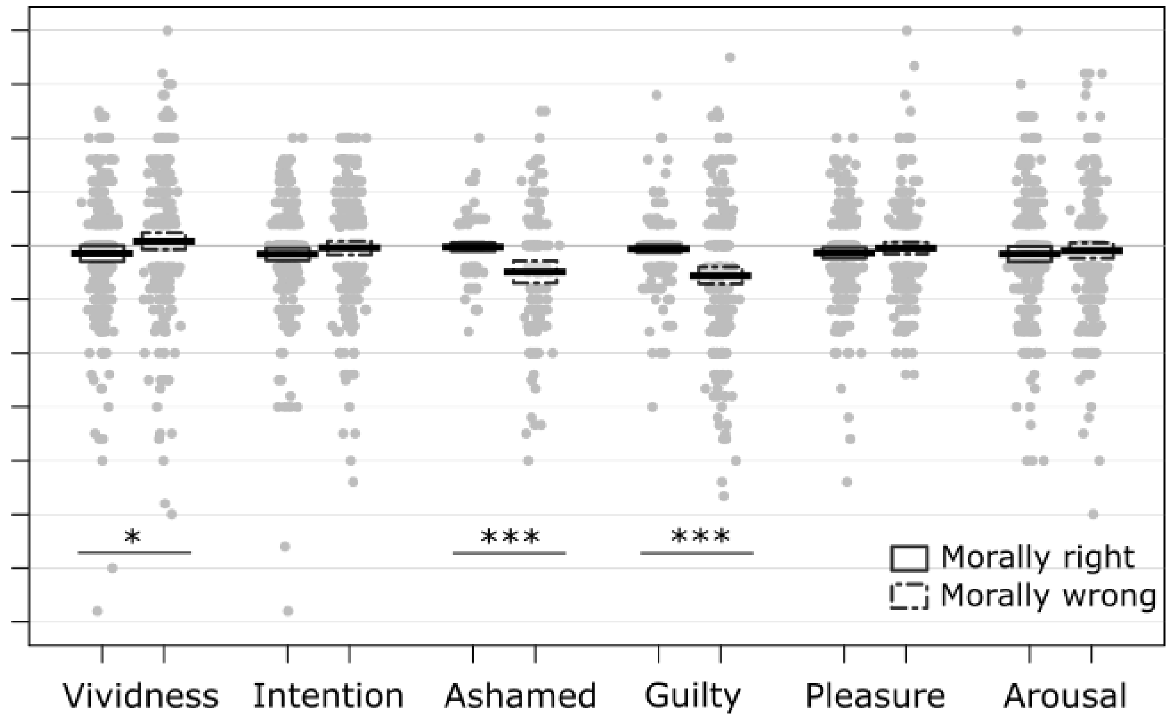


Figure 5.3. Self-report difference scores (before TNT – after TNT) compared across memory type. The scatter dots show difference scores of each individual for each. The thick lines show the means and the boxes depict the 95% confidence interval of the means. *p<.05 ***p<.001.

The self-reports thus showed a direct replication of study 1 results: Negative emotions of guilt and shame are reduced as a result of both thinking and not-thinking of morally wrong memories, but not morally right memories. However, unique to this experiment, participants reported a greater memory age of morally right memories (M_{change} = -2.54 years, SD_{change} = .79) after compared to before the think/no-think task, but for morally wrong memories the change in memory age was not significant (M_{change} = -.08 years, SD_{change} = .74). Furthermore, morally wrong memories were rated as more vivid after the TNT task, whereas morally right memories were rated as less vivid after the TNT task.

5.4 Discussion

Different lines of research have indicated that three specific cognitive strategies may be used to control unwanted memories: Direct suppression (Anderson & Hanslmayr, 2014; Engen & Anderson, 2018), thought substitution (Engen & Anderson, 2018; Joormann et al., 2009), and counterfactual thinking (De Brigard et al., 2019; De Brigard et al., 2013). The present study compared these different memory strategies in terms of effectiveness at regulating negative emotions related to autobiographical memories of morally relevant actions. Most previous research has investigated how emotions associated with memories change after compared to before the memory control task. The present study extended on and replicated some of these findings with morally relevant autobiographical memories by collecting measures of emotional responses to memory reminders both “online” while participants were engaging in memory control, and “offline” to measure after-effects of control. Crucially, the present study provides new evidence that memory control can have an immediate effect on unpleasant memories by reducing the negative emotions associated with them, and that thought substitution and counterfactual thinking strategies may be more advantageous than direct suppression for this measure of emotion regulation. In contrast, subsequent effects of control on memories were generally weaker, but some findings indicated after-effects of control on memory descriptions.

There were interesting effects of memory control on immediate emotions measured during the think/no-think task, after each attempt of retrieval/control. The following effects were found for all strategies and did not differ between strategies. For morally right memories, there was a general improvement in positive feelings from initial attempts to latter attempts, regardless of retrieval or control. As one would expect, participants indicated that thinking of morally wrong memories made them feel worse, and repeatedly thinking of these memories worsened the negative emotions associated with these memories. Crucially though, participants reported feeling better during latter attempts compared to initial attempts of memory control for morally wrong memories. Therefore, repeatedly thinking of past morally wrong actions worsened the already existing negative emotions for such memories, but in contrast, repeatedly applying cognitive control over these memories reduced the negative emotions associated with the memories. The present results thus indicate that repeated memory control provides a distinct advantage for regulating negative emotions associated with morally wrong autobiographical memories, especially in contrast to repeatedly retrieving such unpleasant memories, which exacerbates negative emotions.

This improvement in feelings for negative memories could be a result of success in controlling intrusions of the unwanted memory. As found in previous research (see experiment 1; Benoit et al., 2015; Levy & Anderson, 2012), memory intrusions tend to be high at first attempts of memory suppression, but the intrusion frequency decreases over time, showcasing an improvement in ability to control the unwanted, automatic recollection of memories into consciousness. The overall improvement in negative feelings over repeated attempts at memory control could thus be attributed to a decrease in frequency of the unwanted memory coming to mind in face of the reminder. This real-time indication of improvement in memory control is often found in studies that test direct suppression mechanisms, and is generally indexed by reduced intrusions over repeated attempts (e.g., Benoit et al., 2015; Levy & Anderson, 2012). The present study presents the first evidence, to the best of my knowledge, of this real-time improvement of memory control using thought substitution and counterfactual thinking strategies, as indexed by an improvement in feelings over repeated attempts at controlling autobiographical memories.

There were differences between control strategies in how they affected the memories in general, regardless of the initial or latter attempts at control/retrieval. Both thought substitution and counterfactual thinking induced more *positive* feelings compared to simply retrieving their memories, whereas direct suppression induced more *negative* feelings than retrieving their memories. Interestingly, both thought substitution and counterfactual thinking strategies induced more positive feelings for both morally right and wrong memories compared to retrieving these memories. In line with this, direct suppression reduced negative feelings for morally wrong memories compared to retrieving them, but this improvement was qualitatively lesser than the improvement found in thought substitution and counterfactual thinking groups. Furthermore, for morally right memories that were directly suppressed, there was a notable decrease in positive feelings. Moreover, even though direct suppression induced improvement in negative emotions related to morally wrong memories, the mean score was still *emotionally negative*, whereas counterfactual thinking and thought substitution improved negative emotions such that the mean score was emotionally *positive* as a result of memory control (see Fig 5.3). Therefore, the results indicate that direct suppression can be helpful for upregulating negative emotions, but counterfactual thinking and thought substitution may provide an added benefit by inducing strong positive emotions, regardless of whether the to-be-controlled memory is negative or positive in nature, at least immediately after attempts at control and using the specific strategies implemented here.

Direct suppression is thought to work by inhibiting memory traces from entering consciousness, while attention is maintained on the cue and by pushing the memory contents from memory (Engen & Anderson, 2018). The cue in itself may not be especially positive in nature, unlike thinking of a safe space or better alternative versions of the past. Therefore, it not surprising that participants felt better when thinking of morally right compared to inhibiting those memories, which almost neutralised the positive emotions for these memories. Directly suppressing morally wrong memories potentially reduced intrusions and negative emotions, but this still did not make participants feel *positive* because the distribution of scores were still centred around the negative ratings of the scale. On the other hand, both thought substitution and counterfactual thinking as implemented here required the person to redirect attention from the unwanted memory to an imagined scenario, with one being a safe space such as lying on a beach (thought substitution; Hertel & Calcaterra, 2005; Joormann et al., 2009) and the other being a better version of the to-be-controlled memory (counterfactual thinking; De Brigard et al., 2013), which were inherently positive in nature, unlike a word on the screen.

The results indicated that diverting attention from the unwanted memory to the alternate imagined positive scenarios overpowered emotions associated with both morally right and morally wrong memories. For morally right memories, participants indicated that thinking of these alternate scenarios made them feel more positive than thinking of their actual memories. Moreover, these strategies provided an added benefit for morally wrong memories, as controlling memories led to positive feelings compared to retrieving the memories, which was more negative feelings. So, thought substitution and counterfactual thinking may not only be effective in reducing memory intrusions and neutralising emotions but can also have a functional role in inducing positive feelings, at least in healthy participants, in comparison to direct suppression.

These findings provide important evidence for the debate regarding whether thought substitution can be more effective than direct suppression as a memory control strategy (see Noreen & Ridout, 2016a, 2016b; Stramaccia et al., 2020), and also indicates counterfactual thinking can be considered as an additional strategy for controlling memories. Recent evidence indicates that thought substitution was not effective at inducing forgetting for individuals with dysphoric tendencies (Noreen & Ridout, 2016b, 2016a), who are thought to find direct suppression difficult (e.g., Sacchet et al., 2017). One potential reason for this finding is that they used neutral substitutes in these studies, which may not be effective for inducing positive emotions, as found in the present study. Some evidence indicates emotional valence of the substitute may affect how likely thought substitution may succeed, such that positive substitutes may have a greater chance of succeeding in inducing forgetting (Joormann et al., 2009; Noreen & Ridout, 2016a). This study also supports this idea because a positive safe space was found to induce more positive feelings during thought substitution. The present results therefore provide support to the idea that a positive valence of the substitute/alternate imagination increases the potency of thought substitution in regulating negative emotions.

The results from this study also indicate that counterfactual thinking potentially provides an advantage over simple retrieval, or even direct suppression, of people's unpleasant memories in regulating emotions at the time of retrieval/control. The extant research has investigated the role of counterfactual thinking in emotions by comparing emotion ratings collected before to after a counterfactual thinking paradigm (De Brigard et al., 2019). Interestingly, although they find that counterfactual thinking reduced negative emotions associated with autobiographical memories, remembering the negative memories reduced negative emotions to a greater extent than counterfactual thinking. In line with these findings, the present results also indicated that specific feelings of guilt and shame were reduced after, compared to before, the think/no-think task for both retrieved and controlled memories. Hence, counterfactual thinking may be particularly beneficial over simple retrieval for regulating negative emotions at the time of retrieval/control, whereas the after-effects of counterfactual thinking and simple retrieval are both effective at reducing negative emotions of unpleasant memories. Therefore, it may be that counterfactual thinking mechanisms work differently on emotions at the time of counterfactual thinking, compared to when the person encounters the reminder after memory control. This reasoning is supported by another piece of evidence derived from analysing the memory descriptions provided by participants. Descriptions of morally wrong memories were rated as more negatively-valenced after compared to before the think/no-think task, for the counterfactual group but not thought substitution or direct suppression groups. Although this effect was small it is worth noting due to its relation to previous literature. This opposing effect on emotions during, compared to after, counterfactual thinking could be because of the type of counterfactual thoughts that participants were instructed to generate. Participants imagined a better alternative to their memories (upward counterfactual thinking), which explains why people felt better at the time of thinking of a positive scenario in the present study, but interestingly, research has indicated that people report feeling worse after upward counterfactual thinking when they recall the memory again, potentially due to the regret that is associated with thinking of a better outcome (Allen et al., 2014; Kahneman & Miller, 1986), and the present results agree with this finding.

This after-effect on worsened negative valence of memory descriptions however was also found for simple retrieval in the counterfactual thinking group. One interpretation is that participants may have used counterfactual thinking even during “think” trials accidentally (treatment diffusion), leading to similar effects across conditions. Nevertheless, it is important to consider these interpretations with caution because these effect sizes were small and these specific results of memory descriptions could be false positives. Moreover, even though the descriptions were more negatively-valenced, specific feelings of guilt and shame associated to the memories were reduced rather than exacerbated due to counterfactual thinking (and retrieval).

Such after-effects of thought substitution and direct suppression were not apparent in the memory descriptions for these groups. However, guilt and shame were reduced after compared to before the think/no-think task, in both thought substitution and direct suppression groups, also in line with the counterfactual group, and with De Brigard et al.'s (2019) findings. Therefore, memory control seems to be effective at reducing general negative feelings at the time of controlling unpleasant memories, and consequently also improves specific emotions of guilt and shame when faced with reminders of the memories.

In light of emotion regulation theory (Gross, 2015) therefore, there seems to be a clear advantage of reappraisal (cognitive change) in regulating negative emotions, especially over distraction and/or avoidance. Indeed, cognitive reappraisal is the current gold-standard for emotion regulation in practice and theory (see Engen & Anderson, 2018). Notably, it may be that using highly potent and positive distractors could be more beneficial, maybe as much as cognitive reappraisal, for regulating immediate negative emotions. It may be therefore interesting for this research to be applied in contexts outside memory control, where people in real-time situations can use reappraisal or positive distraction to improve their current emotional state.

Interestingly, not only memory control, but actively remembering such unpleasant memories seemed to also reduce guilt and shame associated with the memories, also in line with findings from experiment 1. Contrastingly, participants indicated that they felt overall negative during the time of thinking of immoral memories and felt worse over repeated attempts at retrieving these unpleasant memories. In this "think" condition, participants were instructed to voluntarily bring the contents of the memory to mind and pay more attention to the details of the memory. This attention to negative details of the memory probably increased over repeated thinking in the think/no-think task, leading to an exacerbation of negative feelings at the time of retrieval. Curiously, some research indicates that this attention to salient negative information can reduce negative emotions associated with unpleasant memories (De Brigard et al., 2019; Ochsner & Gross, 2005; Todd et al., 2012). An interpretation of my findings, therefore, is that paying attention to the negative aspects could have potentially led to negative feelings at the time of retrieval, but these negative emotions may reduce when faced with a reminder post-experiment. This argument is also in line with De Brigard et al.'s (2019) interpretation because they also found both retrieval and counterfactual thinking reduced negative emotions for negative autobiographical memories post compared to pre-test and provide a similar interpretation. However, repeated thinking of negative information in some situations can be indicators of poor mental health, such as ruminative tendencies or obsessive thoughts in anxious individuals, which could exacerbate negative emotions (e.g., Nolen-Hoeksema, 2000; Yapan et al., 2020), so the interpretation that increased attention to negative information can reduce unpleasant emotions is potentially context/situation-dependent and needs to be further investigated in the future.

It is important to note that, as mentioned in experiment 1, a baseline condition was not present to test if any effects on emotions measured post-think/no-think task. Therefore, it is not clear if the reduction in guilt and shame seen in the final test is due to the cognitive processes of thinking/no-thinking or due to habituation or fatigue (see Chapter 6 for further discussion).

Most research in the past has compared the effects of direct suppression and thought substitution in how effective they are at inducing forgetting, and some research indicates that thought substitution is more effective for populations that may find direct suppression difficult (Stramaccia et al., 2020). There was no suppression/control-induced forgetting however in the present study for any strategy, so it was not possible to ascertain whether these different strategies differ in how they induce forgetting of morally relevant autobiographical memories. In fact, the present results replicated experiment 1, such that for all three strategies, pre- and post- think/no-think descriptions of morally wrong memories were rated as less similar than morally right memories, regardless of whether they were controlled or retrieved. Interestingly, the effect size was greater for counterfactual thinking than direct suppression and thought substitution, but no significant difference in effect size between thought substitution and direct suppression was found. It is not straightforward to conclude that these differences between groups were due to memory control, because this greater dissimilarity in descriptions for morally wrong compared to morally right memories was found in both think and no-think conditions in the counterfactual thinking group. One possibility could be that there was treatment diffusion, such that participants were tempted to use counterfactual thinking even during "think" trials, leading to a greater effect for this group but not for direct suppression and thought substitution groups. Nevertheless, the interpretation of this effect is the same as experiment 1, indicating that morally wrong memories may simply be described more differently across time and could be more susceptible to change.

A 3-point Likert scale was used to measure online emotions with the aim to gather relatively more intuitive sense of how the participant is feeling at that time. The downside to this decision however is that emotions usually require more fine-grained measurements rather than absolute scales as measured in this study. In the future, a 5-point Likert scale such as the SAM scale could be used. Alternatively, one could use a more continuous measure, such as holding down a button for varied duration. If the button press is considerably short, that means they feel very negative, but holding for longer durations would reflect feeling more neutral/positive. Such options would therefore provide fine-grained measures of emotion rather than more absolute positive vs. neutral vs. negative responses.

In sum, the present results indicate that memory control may be advantageous for reducing negative emotions felt in response to reminders of unpleasant memories of our past immoral actions at the time of preventing retrieval of the memories, regardless of the memory control strategy. This ability to control emotions associated with memories improves over repeated attempts, evidenced by a weakening of negative emotions during latter compared to early attempts of control. Thought substitution and counterfactual thinking may provide an advantage over to direct suppression for generally improving positive feelings. Therefore, this study has clear implications that memory control can be useful in everyday life for immediate uplifting of emotions, especially by repeated attempts at control, with imagining positive scenarios prevailing over direct suppression at least for emotion regulation.

Chapter 6 – General discussion

The primary aims of the thesis was to investigate neurocognitive mechanisms involved in controlling unwanted autobiographical memories of morally relevant personal actions and its consequences on memory and emotion. Here, memory control refers to stopping retrieval when confronted with a reminder of a particular memory. Our current understanding of the mechanisms and consequences of memory control is largely based on research of our ability to control simple memories that are encoded in a lab-environment, such as words and pictures (e.g., Anderson & Hanslmayr, 2014; Anderson & Hulbert, 2021). More recent research has studied memories that contain more dimensions, such as using emotional words or pictures (Depue et al., 2007; Gagnepain et al., 2017). However, it is important to understand how these results apply to real-world complex memories that we deal with in everyday life, i.e., autobiographical memories. Although there is some research on autobiographical memory control, these studies mostly focus on forgetting as a measure of ability to control memories (Noreen & Macleod, 2013; Noreen et al., 2016; Stephens et al., 2013). The present thesis therefore aimed to broaden our understanding of autobiographical memory control using real-time behavioural and EEG measures.

Studying autobiographical memories also provides an opportunity to understand how the nature of memories can modulate our ability to control them. Most research that has investigated the link between emotion and memory control has investigated more general negative and positive emotions (e.g., Davidson et al., 2020; Engen & Anderson, 2018; van Schie et al., 2013). There is reason to believe that specific memories of our past immoral actions could be strong motivators for us to engage memory control, based on theories arguing that people will be particularly motivated to avoid thinking of certain types of memories, such as those that threaten our positive self-concept (Anderson & Hanslmayr, 2014; Kouchaki & Gino, 2016; Stanley & De Brigard, 2019; Stanley et al., 2018). Interestingly, some evidence indicates that memory suppression can be used to beat forensic memory detection tests, so that a criminal can appear innocent and evade incrimination (Bergström et al., 2013; Hu et al., 2015), but these studies are based on relatively emotionally-neutral autobiographical memories. Due to these issues, morally relevant autobiographical memories were specifically investigated in the present thesis to also understand whether memories of one's own immoral acts are particularly difficult to avoid retrieving in both typical memory control tasks and in forensic memory detection tests.

The neurocognitive mechanisms of autobiographical memory retrieval and control is also relatively unknown, with only one study using fMRI methods (Noreen et al., 2016), to the best of my knowledge. However, there is a rich literature on the mechanisms underlying intentional control of simpler memories (Benoit et al., 2015; Bergström et al., 2007, 2009; Gagnepain et al., 2017; Waldhauser et al., 2012, 2015). The present thesis therefore aimed to shed light on the neural dynamics of autobiographical memory retrieval, suppression, and unwanted intrusions.

6.1 Review of empirical findings

Three empirical studies were conducted in order to investigate these research questions. The key results from these studies will be discussed in the following paragraphs.

In the first experiment (Chapter 3), the think/no-think paradigm was adapted for morally relevant autobiographical memories. The aim was to replicate and extend on findings from previous think/no-think research focusing on simpler memories (e.g., Anderson et al., 2004; Bergström et al., 2009a; Gagnepain et al., 2017; Levy & Anderson, 2012). The predominantly investigated strategy of memory control is direct suppression, which involves applying conscious control to purge memories from awareness (Anderson & Hanslmayr, 2014; Engen & Anderson, 2018; Stramaccia et al., 2020). Using this strategy also provides an opportunity to test how difficult these autobiographical memories are to control, as indexed by memory intrusion ratings provided after each attempt of suppression (Benoit et al., 2015; Hellerstedt et al., 2016; Levy & Anderson, 2012).

To the best of my knowledge, autobiographical memory intrusions have not been measured using the think/no-think paradigm before, and furthermore, the EEG correlates of autobiographical memory retrieval and suppression was also relatively unknown prior to this research. Therefore, experiment 1 provided an exploratory account of these concepts using self-report measures of intrusions, and ERP and EEG oscillatory correlates of retrieval, suppression, and intrusions. There is existent evidence that details of autobiographical memories (Noreen & Macleod, 2013; Stephens et al., 2013), can be forgotten, induced by direct suppression, which we aimed to replicate. As aforementioned, morally wrong memories are theorised to be more difficult to control (Stanley et al., 2018), and thus be more susceptible to change as a result (Kouchaki & Gino, 2016), in line with motivated forgetting theories (Anderson & Hanslmayr, 2014; Stanley & De Brigard, 2019). Therefore, another aim was to investigate how the moral nature of memories modulated the ability to control memories, indicated by behavioural and EEG measures.

The strongest and most reliable results were derived from real-time measures of memory control collected during the think/no-think task. Firstly, participants in this study reported in real-time that they were largely successful at avoiding the memories from coming to mind (around 70-80% of the time on average). When the contents of memories did intrude into awareness despite attempts at suppression, the frequency of intrusions was high during the initial trials of the task, but intrusion frequency reduced over repeated attempts at memory suppression. This result is a replication of the pattern of intrusions observed for think/no-think studies of simpler memories (e.g., Benoit et al., 2015; Levy & Anderson, 2012). Importantly this is the first evidence from real-time measures that retrieval of autobiographical memories can be avoided, and even if they do intrude, people can consciously reduce the frequency of intrusions over time by repeatedly attempting suppression.

EEG correlates of success in autobiographical memory suppression were also found and these results converged with previous findings using simpler stimuli. A positive deflection of ERPs is often found in parietal regions around 500-800ms post-onset of the reminder for memories that are retrieved compared to avoided memories as a result of suppression (e.g., Bergström et al., 2007, 2009b; Depue et al., 2013). This ERP component is thought to index recollection (see Rugg & Curran, 2007), and is thus absent when retrieval is avoided as a result of suppression success. In line with these findings, retrieval of autobiographical memories also elicited more positive ERPs than suppression in these same temporal and spatial regions. Oscillatory data from the EEG also further substantiated these results: A strong, sustained power desynchronisation for no-think compared to think memories was found in the theta band (4 -7 Hz) beginning around 500-1000ms in parietal regions. This is often found in previous research and is argued to index avoidance of retrieval (Legrand et al., 2020; Quaedflieg et al., 2020; Waldhauser et al., 2015). These pieces of evidence together therefore provide more objective, neural evidence that retrieval of autobiographical memories can be avoided by direct suppression.

In later time periods, around 1500-3500ms post-stimulus onset, an ERP positivity for retrieval vs. suppression was present, and the topography of this effect was more frontal-central. As the earlier, more parietal positivity indexes avoided retrieval, this later positivity is interpreted to reflect processes involved in maintaining control over the memory such that it does not enter conscious awareness. This interpretation is based on previous research that find this later sustained think vs. no-think positivity is linked it to greater suppression-induced forgetting (Depue et al., 2013; Hanslmayr et al., 2009). Brain oscillations in the alpha/beta band also supported this claim, as suppressed items had a significant power desynchronisation compared to retrieved items, which has been found in previous studies (Legrand et al., 2020; Waldhauser et al., 2015). Interestingly, when participants experienced an intrusion, this oscillatory correlate of maintaining control over the avoided memory was not as sustained as when suppression was successful. Therefore, it could be that intrusions occurred as a failure of sustained control over memory, in line with findings from another behavioural study (van Schie & Anderson, 2017).

Both behavioural and ERP results indicated that morally wrong memories were more difficult to suppress than morally right memories. Participants reported that they experienced more intrusions when attempting to suppress morally wrong compared to morally right memories. An ERP component was found around 500-700ms in frontal-central areas for morally right but not morally wrong memories. The mechanisms underlying this difference however is not conclusive from this experiment. This effect could reflect mechanisms involved in cognitive control of memories, or a simple difficulty in avoidance of retrieval of morally wrong memories compared to morally right memories.

The first experiment of the present thesis thus elucidated the neural correlates of direct suppression of autobiographical memories, and also found that memories that are morally wrong in nature are more difficult to suppress than morally right memories.

Furthermore, the experiment validated that, our understanding of neural mechanisms involved in directly suppressing simpler memories derived from previous research (Bergström et al., 2009a; Depue et al., 2013; Waldhauser et al., 2015) applies to complex emotional memories that we may deal with in everyday life.

The next two experiments in the theses considered relatively more applied research questions. In Chapter 4, an experiment was devised to test if direct suppression can be used to beat an EEG-based forensic memory detection test. The concealed-information test in the forensic domain presents a series of stimuli, which includes one crime-relevant stimuli (probe), one target word, and other crime-irrelevant stimuli (see Rosenfeld, 2020). If a P300 ERP is elicited in response to the crime-relevant stimulus, then the suspect is considered to have recognised that stimulus and consequently incriminated as guilty of the crime (Rosenfeld, 2018). Some recent evidence has indicated that the suspect can use direct suppression to attenuate the P300 response such that the suspect appears innocent because the person does not seem to recognise the crime-relevant stimulus (Bergström et al., 2013; Hu et al., 2015). However, in these experiments, participants were tested on memories of a mock crime that they committed in a lab-environment. Such memories may not entirely reflect memories of real-world crime, which are more emotional and potentially more difficult to conceal using direct suppression. Therefore, the second experiment in the thesis investigated the ability to directly suppress P300 responses from the concealed information test of mock crimes, morally right, and morally wrong autobiographical memories (see Chapter 3 for full detail).

In contrast to our predictions, direct suppression was not effective for beating the concealed-information test in this study. Three sources of evidence substantiate this claim. Firstly, and most importantly, the suppression group elicited a significant probe vs. irrelevant P300 for all three types of memories similar to the standard-guilty group. Secondly, oscillatory brain correlates - potentially reflecting recognition of relevant stimuli - was present for probes vs. irrelevants for both suppression- and standard- guilty groups. Thirdly, individual guilt classification analysis indicated that more individuals in both suppression- and standard-guilty groups were classified as guilty than in the innocent group.

It is not entirely clear from the results why suppression failed in version of the CIT used in Chapter 4. One reason could be that previous studies that found suppression-induced reduction in the P300 induced greater task-demand on cognitive resources and thus created situations that were more conducive for suppression to succeed (e.g., in Hu et al., 2015), as Ward and Rosenfeld (2017) argue. Specifically, using a 20-80 target vs. non-target ratio may be beneficial for making the CIT resistant to memory suppression.

Another explanation is that because the P300 relies on the orienting response, this could not be suppressed, especially because of highly salient probes being used in the present study. Two recent studies support this idea: Firstly, one study directly provides evidence that the more salient items evoke greater P300 amplitudes than low salient stimuli (klein Selle et al., 2021) in a blocked design; second, suppression failed when participants were probed with their own names, which are one of the most salient stimuli we could come across in everyday life (Rosenfeld et al., 2017). The blocked design used in the present study could have carried over failed suppression effects across the different memory types, which could explain why the mock-crime may be suppressible in a stand-alone CIT block (as Hu et al., 2015 finds), rather than in a blocked CIT design that include more salient stimuli such as moral and personal autobiographical memories, like in the present thesis.

Therefore, the evidence indicates that the P300 in concealed information test may be sensitive to the saliency of stimuli, task demands, and recognition processes, which all may not be equally susceptible to inhibition. Future research is therefore needed to tease out the roles of these different processes in suppressing the P300 response. Nevertheless, the study does provide evidence that this version of the CIT is resistant to suppression and provides support for its practical use.

The final study in the thesis (Chapter 5) regarded the role of memory control in regulating emotions. Two streams of research have provided evidence that memory control and emotion regulation are closely related. One group of researchers has consistently found that participants who suffer from mental health and emotional issues have poorer memory control than emotionally healthy participants (e.g., Stramaccia et al., 2020; Mary et al., 2020). The present thesis used the other approach, which involves manipulating the emotional nature of to-be-controlled memories in healthy participants and measuring emotional changes due to control. Most previous research has considered "offline" measures of emotion, after the think/no-think task in a surprise memory test. There is evidence that memory suppression downregulates hemodynamic activity in the amygdala especially for negative memories (Gagnepain et al., 2017). This is directly related to reduced phenomenological ratings of negative emotions (Gagnepain et al., 2017; Legrand et al., 2020) and reduced emotional physiological responses to negative stimuli (Legrand et al., 2020).

The emotional responses at the time of controlling memories has not yet been investigated to the best of my knowledge, and hence was studied in the present thesis. Moreover, the memory control strategies of direct suppression, thought substitution and counterfactual thinking and their effect on emotion regulation have not yet been compared directly in an experiment, which the present thesis also aimed to explore. To address these research aims, the think/no-think paradigm was used in the third experiment as in experiment 1, with two major design changes. First, an emotion rating scale was introduced in each think/no-think trial to measure the immediate effects of memory control on emotion, for the first time using this paradigm. Second, three different groups of participants were tested, and each group used one memory control strategy.

The results indicated that regardless of the strategy, memory control distinctly reduced the immediate negative emotions associated with morally wrong memories. Continuously retrieving morally wrong memories heightened negative emotions, as evidenced by more reports of negative feelings during the end of the task than in the beginning of the task. Importantly, in contrast, continuously attempting to control memories improved these negative emotions over time, as evidenced by larger reduction of negative emotions during the end of the task compared to the beginning of the task.

Therefore, experiment 3 provided clear evidence that memory control leads to successful emotion regulation at least at the time of control, and repeatedly applying this control can enhance the effectiveness of this emotion regulation. This reduction of negative emotions could be a result of control in the frequency of unwanted intrusions; participants felt more negative during initial attempts at memory control because they were less successful at avoiding intrusions, but as they got better at controlling these intrusions, they also felt better as a consequence. This interpretation is supported by evidence that intrusions also decrease over repeated attempts of control, as found in experiment 1 of the thesis, and other research (e.g., Benoit et al., 2015; Levy & Anderson, 2012).

I argue that although stopping intrusions may make a person feel qualitatively *better*, it may not necessarily make them feel qualitatively *positive*. Participants that directly suppressed their memories only purged the contents from memory by redirecting attention to words on the screen, leading to lessening of negative emotions. On the other hand, substituting unwanted thoughts by imagining a positive safe space not only reduces negative feelings, but may also induce pleasant feelings due to the positive nature of the substitute (see also Joormann et al., 2009; Noreen & Ridout, 2016a). Similarly, by imagining a better alternative to the unwanted memory, not only is the original negative emotions associated to the memory dampened but imagining the alternate positive scenario could make us feel more positive at least at the time of imagination. These interpretations were also supported by the finding that participants felt more positive during thought substitution/counterfactual thinking compared to retrieval. Therefore, even though direct suppression can be effective at regulating negative emotions, both thought substitution and counterfactual thinking could be more advantageous for immediate emotion regulation because they not only reduce intrusions of the original unwanted contents of memory, but also provide a positive imaginary space that can induce pleasant feelings.

The present thesis also considered more “offline” measures of the consequences of memory control on memory and emotion in experiments 1 (Chapter 3) and 3 (Chapter 5). These measures were collected in the form of self-report questionnaires collected before and after the think/no-think task. Participants provided a description of each memory both before and after the task. Suppression-induced forgetting was quantified by comparing these two descriptions. After providing a description, they reported memory characteristics such as memory age, vividness, morality, intention of their actions in the memory, followed by reports of morality-specific emotions of guilt, shame, and general emotional pleasure, and arousal. This allowed to test for any suppression-induced changes in memory characteristics, and if these “offline” measures of emotion could be regulated by comparing self-reports before and after the think/no-think task.

The results from both experiments did not show any notable differences in memory or emotion measures between retrieval and control, against our predictions. However, there were overall differences between morally right and wrong memories: Regardless of whether participants retrieved or controlled their memories, the two descriptions for morally wrong memories were more dissimilar than the two descriptions of morally right memories. Furthermore, the specific emotions of guilt and shame were also reduced for morally wrong, but not morally right memories for both retrieved and control conditions. Therefore, it is not clear if these changes for morally wrong memories are a result of memory control, or both retrieval and memory control, or are time-based or fatigue effects (discussed in a later section of this chapter).

6.2 Broader implications on theory and practice

A common finding in social psychology research and in everyday life is that people often believe that they are morally good (Aquino & Americus, 2002; Freitas, Cikara, Grossmann, & Schlegel, 2018; Strohminger, Knobe, & Newman, 2017), but ironically, behave against their own moral code at a startling frequency (e.g., Hofmann et al., 2014). There are many potential reasons for this discrepancy in thought and behaviour (see literature on moral disengagement, e.g., Moore, 2015), but one promising explanation comes from the theory of motivated forgetting (see Anderson & Hanslmayr, 2014). Essentially, morally wrong memories are considered to be strongly related to the sense of self, usually associated with feelings of guilt, remorse, and/or shame, that are all ingredients to especially motivate us to control such memories and consequently lead to forgetting (Kouchaki & Gino, 2016). In line with this proposition, recent evidence has suggested that people could forget details about their past immoral actions because they are motivated maintain a view that they are morally good and because thoughts of these memories induces unpleasant feelings of guilt and shame (Stanley & De Brigard, 2019).

In the present thesis, when faced with reminders of their past moral actions, participants reported that morally wrong memories were more intrusive than morally right memories, consistent with one aspect of motivated forgetting. However, there was no clear suppression-induced forgetting in the present study, as the theory would predict. Therefore, there was no direct evidence for motivated forgetting in the present study.

Most of the evidence for *forgetting* in support of the account that people are motivated to forget morally wrong memories comes from testing memories either encoded in the lab or imagined events rather than actually experienced events in real life. For instance, details of an honour code for an examination are forgotten for people who go on to cheat in that test (Shu, Gino, & Bazerman, 2011), consumers of products forget immoral details of the product that they consume , such as if it was produced from child labour (Reczek et al., 2018). Therefore, forgetting these details could provide a benefit for such memories that are not necessarily personally relevant autobiographical memories and has applications in those aspects of immoral behaviour. Evidence from this thesis and other research indicates that, suppression-induced forgetting of personally relevant and emotional autobiographical memories is not often found to be complete forgetting of an episode, but rather forgetting of a subset of details in those memories (Noreen & Macleod, 2013; Stephens et al., 2013).

Considering that autobiographical memories may not necessarily be completely forgotten; it is important to reconsider motivated forgetting from a different perspective for such personal and emotional memories. Theoretically, as Stanley and De Brigard (2019) argue, the motivation to “forget” such personal, autobiographical memories is potentially due to the unpleasant emotions of guilt and shame, and consequently to maintain an idea of a good self. I argue therefore that a more suitable approach for personal autobiographical memories would be to consider if aspects other than accuracy of details of the memory could be affected.

Some research has suggested that the phenomenology of memories are more obfuscated for past immoral actions, such that they are perceived to have occurred more distant in the past, are less vivid than moral actions (e.g., Escobedo & Adolphs, 2010;

Kouchaki & Gino, 2016). In line with these findings, participants in all three experiments in the thesis reported their morally wrong memories as more distant in the past (Escobedo & Adolphs, 2010), less vivid (Kouchaki & Gino, 2016), and their actions as less intentional than morally right memories. An interesting finding is that even though morally wrong memories were rated as more distant and less vivid than morally right memories, they were more difficult to suppress than morally right memories when exposed to reminders of the memories. In real life, it is not often that we can avoid reminders of negative memories, so even if a self-serving motivated bias existed for such memories, they could still be more intrusive in face of reminders to the memory, as experiment 1 suggests. Therefore, an important research question that the present thesis also answers is that when faced with the reminder of a memory, the negative emotions associated to the memory can be regulated, even if the episode is not forgotten (see also Engen & Anderson, 2018). To this end, results from this thesis provides promising evidence that the negative emotions associated with morally wrong memories can be improved as a result of memory control, and this could especially be true for feelings of guilt and shame.

Interestingly, the results from the present thesis also indicated that suppression attempts may not always be successful and could be counterproductive (see also Stramaccia et al., 2020). In experiment 1 of the thesis (Chapter 3), there was strong evidence that autobiographical memory retrieval was successfully avoided, from both behavioural measures using self-reports of intrusions, and using EEG measures, such that markers of retrieval were reduced during suppression attempts. Contrastingly though, in experiment 2, attempts at suppressing memories in the concealed-information test was not successful. The concealed-information test involves presenting a crime-relevant stimulus (probe) mixed with irrelevant stimuli. Therefore, it is nearly identical to an odd-ball task. The P300 in fact is often found in such odd-ball tasks and is thought to mainly reflect orientation of attentional resources due to the salience of a stimulus (see Polich, 2007)). The think/no-think task on the other hand does not necessarily evoke such orienting-related processes and could be a clearer measure of retrieval and recognition-related processes. Therefore, the results from these experiments indicate that while memory retrieval can be inhibited, the orienting response cannot as easily be suppressed.

A question that arises is why retrieval can be suppressed but the orienting response cannot be suppressed? This orienting response reflects bottom-up attention processes, which is often considered to be *stimulus-driven* by the novelty and/or saliency of the stimulus. Bottom-up attention is thought to be automatic, reflexive, and less prone to be affected by conscious processes such as cognitive control (see Pinto et al., 2013). Memory retrieval on the other hand, is not as rigid, but is in fact much more malleable and vulnerable to change (Schacter, Coyle, & Harvard Center for the Study of Mind, 1995). Memory retrieval processes are also more *goal-driven,* making them more susceptible to top-down attention and cognitive control mechanisms such as direct suppression.

These results also has implications on our understanding of parietal ERP components. The P300 (particularly the P3b) component in the concealed-information test is found to be especially sensitive to the saliency of stimuli, and is thus interpreted a manifestation of the orienting response and bottom up processes (see also klein Selle et al., 2021). The morally-relevant autobiographical memories used in this study are highly salient because they are strongly related to the sense of self (Stanley & De Brigard, 2019), and could therefore increase the P300 response, and cause failure in suppression. The parietal positivity effect found in the think/no-think task however (experiment 1) could instead index recollection and recognition processes, as it was not present during suppression attempts in experiment 1, unlike the P300 in experiment 2 (see Rugg & Curran, 2007).

I do not mean to argue that recognition does not occur in the concealed-information test, or that the P300 does not reflect recognition or retrieval process, but rather that the saliency of the probe, and in turn the bottom-up orientation response, could potentially override the recognition aspect of the P300. The important idea therefore is that the P300 based concealed-information may rely less on recognition processes, and more on the saliency of the crime-relevant stimulus. Nonetheless, a clear practical implication of this study is that suppression may not be effective as a countermeasure for beating the concealed information test of real world crimes, especially because such stimuli may be salient and potentially also increase rather than decrease the P300 response (see also Rosenfeld, 2017).

Experiment 1 in the present thesis also sheds light on solutions for the multiple comparison problem (see Chapter 2) that researchers face with handling highly correlated, multidimensional EEG data sets. The cluster-based permutation tests are efficient at identifying large differences across time-electrode-frequency dimensions. For instance, the large think vs. no-think ERP and oscillatory effects, or the P300 ERP components found in the present thesis were clear in cluster-based tests. Indeed, these effects are also the most reliable effects. However, a dilemma arises when one expects “smaller” effects in the data that are not as widely distributed in time or space. The logic of cluster-based tests would be to consider small effects as noise and not true effects. However, if these effects have been consistently found in the literature, then it raises the question whether true effects are overlooked by using cluster-based tests. One reason why the cluster tests may not detect these effects could be because these differences arise only in small regions of time and space. For example, ERP peaks such as the N2 and P2 peaks that are thought to represent certain cognitive processes (e.g., Chen et al., 2012; Mecklinger et al., 2009), may be overlooked when there are large differences occurring elsewhere in the dataset.

One could be tempted to visually inspect the data to choose specific regions-of-interest for these smaller components that may not be detected as significant in tests that correct for multiple comparisons. However, doing so introduces issues of experimenter bias and interpreting noise as true data (Kilner, 2013). Such biased methods of selecting specific regions-of-interests were commonly used in the past, and unfortunately some studies in the think/no-think literature also have used such approaches. For instance, the N2 component is strongly argued to reflect cognitive control during memory suppression (e.g. Streb et al., 2016), but studies seem to select these specific regions-of-interests both using previous research and visual inspection. I argue that the most appropriate strategy would be to consider an *a-priori* regions-of-interest based on previous research and pre-register these parameters so that the experimenter is not biased to "visually inspect" the data and interpret these components to on a post-hoc basis. Pre-registration practices are easier in the P300 and concealed information test literature as they are large and reliable effects and were also used in experiment 2 of this thesis. Think/no-think research should also thus adopt this idea, now that there is growing evidence regarding the neural correlates of processes involved in controlling our memories, especially with a focus on understanding the "smaller" effects such as the N2, which if indeed reflects cognitive control, could be important to further our understanding of memory control in everyday life.

6.3 Limitations and future directions

There were limitations in the present thesis, which also provided avenues for further research. Firstly, in experiments 1 and 3, the post-experimental questionnaires indicated that both retrieving and controlling memories led to greater reduction in negative emotions of guilt and shame for morally wrong but not morally right memories. One interpretation of this effect is that any exposure to morally wrong memories is useful at regulating negative emotions associated with thinking of these episodes. However, another reason for this result could be that ratings of guilt and shame especially are at floor for morally right memories, meaning that any reduction in these feelings would not turn out as statistically significant. So, in this case, morally right memories potentially do not act as a good comparison memory type. Furthermore, it could be that participants are fatigued at the end of the study, especially considering the EEG setup and the length of test times. This fatigue could have resulted in overall dampening of emotions regardless of whether they were retrieved or controlled in the think/no-think task.

Adding a baseline conditions in future studies will help understand if this reduction in associated negative moral emotions is due to the retrieval/control issues, or due to other factors. If a similar reduction in negative emotions is found for think, no-think, and baseline, then it would reflect an effect of fatigue, but if only thinking and not thinking of the memories lead to a reduction in negative emotions, then it suggests that the exposure of these memories is helpful at regulating the negative emotions associated with them, as interpreted in the thesis. Furthermore, suppression-induced forgetting is characterised by lower memory accuracy for no-think items compared to a baseline, rather than the think condition. Therefore, even if we did find lower accuracy for no-think compared to think in the present thesis, it would not necessarily qualify as suppression-induced forgetting. The main reason for not adding a baseline condition in experiment was because I could not collect a sufficient number of memories to split them into think, no-think, and baseline conditions, which would impact the power to detect any effects. While piloting the questionnaire to collect morally right and wrong memories, participants indicated that they were struggling to think of more than 10 memories for each memory type. Therefore, these studies were not optimised to test for suppression-induced forgetting or reductions in emotional responses after the think/no-think phase had been completed.

The analysis approach used in the present thesis for answering the question regarding whether morally wrong memories are more intrusive than morally right memories involved comparing the intrusiveness of morally wrong vs. morally right memories that had been classified into these two different categories. Due to the complex nature of these autobiographical memories however, there could be multiple reasons why a morally wrong memory could be more or less intrusive than a morally right memory. In theory, it is often argued that memories of morally wrong actions are more intrusive due to the negative emotions of guilt and shame (e.g., Stanley & De Brigard, 2019). Recently, Knez and Nordhall (2017) used the memory experiences questionnaire (Sutin & Robins, 2007), and found that greater guilt associated with memories of past immoral acts led to lower accessibility of these memories, but they nevertheless contained high sensory detail. Therefore, guilt potentially plays an important role in increasing the potential intrusiveness of these types of memories. Another aspect that could increase a morally wrong memory's intrusiveness is how threatening it is to the self-concept, which is a core reason for why it is believed to be a good motivator for memory control (Stanley, Henne, et al., 2017). Therefore, there could be multiple variables that affect how we remember and control such complex memories.

Considering that I did measure memory characteristics and emotional measures of each memory, it would be interesting to use a multivariate analysis technique to understand what combination of characteristics would make certain memories more difficult to suppress/control. One key aspect that was not measured is how threatening the memory is to the concept of the self, which is considered to be a major reason for the discomfort that potentially motivates us to control such memories, which limited the potential to answer research questions in more detail in this study. Importantly, these factors may not only predict memory intrusiveness, but also how easy it is to regulate the negative emotions associated with such memories. Therefore, such analysis approaches can also be used even for more general negative/positive memories including a range of questions regarding the memory characteristics. An interesting and practically relevant research question would thus arise: What makes a memory more or less intrusive?

Another limitation exists in experiment 3 of the thesis, which found that three different strategies of memory control can be beneficial for regulating negative emotions. Emotions were measured immediately after attempts at memory retrieval/suppression, and one concern is that this measure could be susceptible to demand characteristics. It is intuitively expected that one would feel more negative after thinking (repeatedly) about a morally wrong memory, but not-thinking of these memories should make them feel better. Furthering this argument, in the surprise memory test, the participants do not necessarily know which think/no-think condition the memory belonged to, and negative emotions were reduced for both think/no-think conditions. This discrepancy was interpreted in a theoretical sense, which could be a possibility, but one must also consider this limitation while making such theoretical interpretation.

Future research using more objective measures of memory control and emotions could be used to substantiate these findings further. Functional brain imaging can potentially provide more objective insight into these research questions. For instance, some evidence has found that memory control downregulates amygdala activity for negative emotions associated with unpleasant memories (Gagnepain et al., 2017). Future research could investigate if these online ratings of emotions is indeed reflected in changes in such emotional brain states to provide stronger evidence for the role of memory control in regulation of immediate emotions. Moreover, self-reports of positive/negative emotions could be combined with physiological measures of emotion. For instance, there is also strong evidence that heart rate variability is linked with better emotion regulation, and that it induces oscillatory activity in the brain and affects functional connectivity between regions involved in emotion regulation

(Mather & Thayer, 2018). To my knowledge, there is evidence that heart-rate variability predicts memory control, and cognitive control in general (Gillie, Vasey, & Thayer, 2014). There is also evidence of reduction in physiological emotional responses to aversive stimuli as an after-effect of memory control (Legrand et al., 2020). It may be interesting therefore to measure if memory control can lead to physiological changes in real-time, similar to the pattern of self-reported emotions found in the present study.

6.4 Conclusions

In conclusion, the present thesis provides the first neurocognitive evidence that unwanted retrieval of autobiographical memories of past immoral actions can be avoided using direct suppression from both EEG and behavioural measures. Although there was some indication for neural correlates of cognitive control processes that led to this success in avoidance, further research is needed to solidify these ideas. Importantly, both unwanted intrusions and negative emotions associated to morally wrong memories were regulated over repeated attempts at control. Therefore, there are clear implications for emotion regulation research and for everyday life that controlling such unwanted memories can help us feel better at least while in face of the reminder, especially by repeatedly attempting control. Interestingly, strategies of thought substitution and counterfactual thinking were found to have an advantage in immediate emotion regulation over direct suppression. There was also an indication that unwanted retrieval of memories of past immoral actions is more difficult to control than such retrieval of memories of morally right actions, even though morally wrong memories were perceived to be less vivid and more generally distant than morally right memories. Therefore, this research provides direct evidence that we may be more motivated to control such morally wrong memories. Furthermore, the present thesis tested the effects of suppression in a P300-based forensic memory detection test for emotional memories of past immoral actions for the first time. The results indicated that the concealed-information-test was not susceptible to suppression and thus has clear implications that this test may be useful in real life for detecting guilty knowledge. However, the results also indicated that these tests may rely more on the saliency or cognitive load, rather than recognition, of crime-relevant stimuli. All in all, this thesis also provides a clear foundation for future research on the real-time effects of emotional autobiographical memory control.

References

Adams, S. A. (2008). Using Transactional Analysis and Mental Imagery to Help Shame-Based Identity Adults Make Peace With Their Past. *Adultspan Journal*, *7*(1), 2–12. https://doi.org/10.1002/j.2161-0029.2008.tb00038.x

Addis, D. R., Moscovitch, M., Crawley, A. P., & McAndrews, M. P. (2004). Recollective qualities modulate hippocampal activation during autobiographical memory retrieval. *Hippocampus*, *14*(6), 752–762. https://doi.org/10.1002/hipo.10215

Allen, M. S., Greenlees, I., & Jones, M. V. (2014). Personality, counterfactual thinking, and negative emotional reactivity. *Psychology of Sport and Exercise*, *15*(2), 147–154. https://doi.org/10.1016/j.psychsport.2013.10.011

Anderson, M. C. (2005). The Role of Inhibitory Control in Forgetting Unwanted Memories: A Consideration of Three Methods. *Dynamic Cognitive Processes*, 159–189. https://doi.org/10.1007/4-431-27431-6_8

Anderson, M. C., Bjork, E. L., & Bjork, R. A. (2000). Retrieval-induced forgetting: Evidence for a recall-specific mechanism. *Psychonomic Bulletin and Review*, *7*(3), 522–530. https://doi.org/10.3758/BF03214366

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), 279–292. https://doi.org/10.1016/J.TICS.2014.03.002

Anderson, M. C., & Huddleston, E. (2012). Towards a Cognitive and Neurobiological Model of Motivated Forgetting (pp. 53–120). Springer, New York, NY.

https://doi.org/10.1007/978-1-4614-1195-6_3

Anderson, M. C., & Hulbert, J. C. (2021, January 4). Active Forgetting: Adaptation of Memory by Prefrontal Control. *Annual Review of Psychology*. Annual Reviews Inc. https://doi.org/10.1146/annurev-psych-072720-094140

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

Anderson, M. C., Ochsner, K. N., Kuhl, B., Cooper, J., Robertson, E., Gabrieli, S. W., … Gabrieli, J. D. E. (2004). Neural Systems Underlying the Suppression of Unwanted Memories. *Science, 303*(5655), 232–235. https://doi.org/10.1126/science.1089504

Aquino, K., & Americus, R. (2002). The self-importance of moral identity. *Journal of Personality and Social Psychology*, *83*(6), 1423–1440. https://doi.org/10.1037/0022-3514.83.6.1423

Bastiaansen, M., Mazaheri, A., & Jensen, O. (2012). Beyond ERPs:: Oscillatory Neuronal Dynamics. In *The Oxford Handbook of Event-Related Potential Components*. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780195374148.013.0024

Ben-Shakhar, G., & Nahari, T. (2018). The External Validity of Studies Examining the Detection of Concealed Knowledge Using the Concealed Information Test. *Detecting Concealed Information and Deception*, 59–76. https://doi.org/10.1016/B978-0-12-812729-2.00003-3

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

Benoit, R. G., Davies, D. J., & Anderson, M. C. (2016). Reducing future fears by suppressing the brain mechanisms underlying episodic simulation. *Proceedings of the National Academy of Sciences*. https://doi.org/10.1073/pnas.1606604114

Benoit, R. G., Hulbert, J. C., Huddleston, E., & Anderson, M. C. (2015). Adaptive Top–Down Suppression of Hippocampal Activity and the Purging of Intrusive Memories from Consciousness. *Journal of Cognitive Neuroscience*, *27*(1), 96–111. https://doi.org/10.1162/jocn_a_00696

Bergström, Z. M., Anderson, M. C., Buda, M., Simons, J. S., & Richardson-Klavehn, A. (2013). Intentional retrieval suppression can conceal guilty knowledge in ERP memory detection tests. *Biological Psychology*. https://doi.org/10.1016/j.biopsycho.2013.04.012

Bergström, Z. M., de Fockert, J., & Richardson-Klavehn, A. (2009a). Event-related Potential Evidence that Automatic Recollection Can Be Voluntarily Avoided. *Journal of Cognitive Neuroscience*, *21*(7), 1280–1301. https://doi.org/10.1162/jocn.2009.21075

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

Bergström, Z. M., O'Connor, R. J., Li, M. K. H., & Simons, J. S. (2012). Event-related potential evidence for separable automatic and controlled retrieval processes in proactive interference. *Brain Research*, *1455*, 90–102. https://doi.org/10.1016/j.brainres.2012.03.043

Bergström, Z. M., Velmans, M., de Fockert, J., & Richardson-Klavehn, A. (2007). ERP evidence for successful voluntary avoidance of conscious recollection. *Brain Research*, *1151*, 119–133. https://doi.org/10.1016/J.BRAINRES.2007.03.014

Bernat, E. M., Malone, S. M., Williams, W. J., Patrick, C. J., & Iacono, W. G. (2007). Decomposing delta, theta, and alpha time-frequency ERP activity from a visual oddball task using PCA. *International Journal of Psychophysiology*, *64*(1), 62–74. https://doi.org/10.1016/j.ijpsycho.2006.07.015

Berntsen, D. (n.d.). The Unbidden Past: Involuntary Autobiographical Memories as a Basic Mode of Remembering. https://doi.org/10.1177/0963721410370301

Bian, Z., Yang, R., Yang, X., Liu, Y., Gao, X., & Chen, H. (2021). Influence of negative mood on restrained eaters' memory suppression of food cues: An event-related potentials study. *Appetite, 164,* 105269. https://doi.org/10.1016/j.appet.2021.105269

Burgess, A. P. (2012). Towards a Unified Understanding of Event-Related Changes in the EEG: The Firefly Model of Synchronization through Cross-Frequency Phase Modulation. *PLoS ONE*, *7*(9), e45630. https://doi.org/10.1371/journal.pone.0045630

Castiglione, A., Wagner, J., Anderson, M., & Aron, A. R. (2019). Preventing a Thought from Coming to Mind Elicits Increased Right Frontal Beta Just as Stopping Action Does. *Cerebral Cortex*, *29*(5), 2160–2172. https://doi.org/10.1093/cercor/bhz017

Catarino, A., Küpper, C. S., Werner-Seidler, A., Dalgleish, T., & Anderson, M. C. (2015). Failing to Forget: Inhibitory-Control Deficits Compromise Memory Suppression in Posttraumatic Stress Disorder. *Psychological Science*, *26*(5), 604–616. https://doi.org/10.1177/0956797615569889

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

Cohen, M. X. (2019). A better way to define and describe Morlet wavelets for time-frequency analysis. *NeuroImage*, *199*, 81–86. https://doi.org/10.1016/j.neuroimage.2019.05.048

Cohen, M. X., & Gulbinaite, R. (2014, January 15). Five methodological challenges in cognitive electrophysiology. *NeuroImage*. Academic Press Inc. https://doi.org/10.1016/j.neuroimage.2013.08.010

Compton, R. J. (2003, May 18). The interface between emotion and attention: a review of evidence from psychology and neuroscience. *Behavioral and Cognitive Neuroscience Reviews*. SAGE Publications. https://doi.org/10.1177/1534582303002002003

Conway, M. A., Pleydell-Pearce, C. W., & Whitecross, S. E. (2001). The neuroanatomy of autobiographical memory: A slow cortical potential study of autobiographical memory retrieval. *Journal of Memory and Language*, *45*(3), 493–524. https://doi.org/10.1006/jmla.2001.2781

Daselaar, S. M., Rice, H. J., Greenberg, D. L., Cabeza, R., LaBar, K. S., & Rubin, D. C. (2008). The spatiotemporal dynamics of autobiographical memory: Neural correlates of recall, emotional intensity, and reliving. *Cerebral Cortex*, *18*(1), 217–229. https://doi.org/10.1093/cercor/bhm048

David, O., Kilner, J. M., & Friston, K. J. (2006). Mechanisms of evoked and induced responses in MEG/EEG. *NeuroImage*, *31*(4), 1580–1591. https://doi.org/10.1016/j.neuroimage.2006.02.034

Davidson, P., Hellerstedt, R., Jönsson, P., & Johansson, M. (2020a). Suppression-induced forgetting diminishes following a delay of either sleep or wake. *Journal of Cognitive Psychology*, *32*(1), 4–26. https://doi.org/10.1080/20445911.2019.1705311

Davidson, P., Hellerstedt, R., Jönsson, P., & Johansson, M. (2020b). Suppression-induced forgetting diminishes following a delay of either sleep or wake. *Journal of Cognitive Psychology*, *32*(1), 4–26. https://doi.org/10.1080/20445911.2019.1705311

De Brigard, F., Addis, D. R., Ford, J. H., Schacter, D. L., & Giovanello, K. S. (2013). Remembering what could have happened: Neural correlates of episodic counterfactual thinking. *Neuropsychologia*, *51*(12), 2401–2414. https://doi.org/10.1016/J.NEUROPSYCHOLOGIA.2013.01.015

De Brigard, F., Hanna, E., St Jacques, P. L., & Schacter, D. L. (2019). How thinking about what could have been affects how we feel about what was. *Cognition and Emotion*, *33*(4), 646–659. https://doi.org/10.1080/02699931.2018.1478280

De Brigard, F., Parikh, N., Stewart, G. W., Szpunar, K. K., & Schacter, D. L. (2017). Neural activity associated with repetitive simulation of episodic counterfactual thoughts. *Neuropsychologia*, *106*, 123–132. https://doi.org/10.1016/J.NEUROPSYCHOLOGIA.2017.09.022

De Brigard, F., Szpunar, K. K., & Schacter, D. L. (2013). Coming to Grips With the Past: Effect of Repeated Simulation on the Perceived Plausibility of Episodic Counterfactual Thoughts. *Psychological Science*. https://doi.org/10.1177/0956797612468163

Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*(1), 9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009

Depue, B. E., Banich, M. T., & Curran, T. (2006). Suppression of emotional and nonemotional content in memory: Effects of repetition on cognitive control. *Psychological Science*, *17*(5), 441–447. https://doi.org/10.1111/j.1467-9280.2006.01725.x

Depue, B. E., Burgess, G. C., Willcutt, E. G., Ruzic, L., & Banich, M. T. (2010). Inhibitory control of memory retrieval and motor processing associated with the right lateral prefrontal cortex: Evidence from deficits in individuals with ADHD. *Neuropsychologia*, *48*(13), 3909–3917. https://doi.org/10.1016/j.neuropsychologia.2010.09.013

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

Depue, B. E., Ketz, N., Mollison, M. V, Nyhus, E., Banich, M. T., & Curran, T. (2013). ERPs and Neural Oscillations during Volitional Suppression of Memory Retrieval. *Journal of Cognitive Neuroscience*, *25*(10), 1624–1633. https://doi.org/10.1162/jocn_a_00418

Diana, R. A., Yonelinas, A. P., & Ranganath, C. (2007). Imaging recollection and familiarity in the medial temporal lobe: a three-component model. *Trends in Cognitive Sciences*, *11*(9), 379–386. https://doi.org/10.1016/j.tics.2007.08.001

Dieler, A. C., Herrmann, M. J., & Fallgatter, A. J. (2014). Voluntary suppression of thoughts is influenced by anxious and ruminative tendencies in healthy volunteers. *Memory*, *22*(3), 184–193. https://doi.org/10.1080/09658211.2013.774420

Donkers, F. C. L., & Van Boxtel, G. J. M. (2004). The N2 in go/no-go tasks reflects conflict monitoring not response inhibition. *Brain and Cognition*, *56*(2 SPEC. ISS.), 165–176. https://doi.org/10.1016/j.bandc.2004.04.005

Drew, T. W., McCollough, A. W., & Vogel, E. K. (2006). Event-Related Potential Measures of Visual Working Memory. *Clinical EEG and Neuroscience*, *37*(4), 286–291. https://doi.org/10.1177/155005940603700405

Dudai, Y. (2004). The neurobiology of consolidations, or, how stable is the engram? *Annual*

Review of Psychology, *55*, 51–86. https://doi.org/10.1146/annurev.psych.55.090902.142050

Ehlers, A. (2010). Understanding and treating unwanted trauma memories in posttraumatic stress disorder. *Journal of Psychology*, *218*(2), 141–145. https://doi.org/10.1027/0044-3409/a000021

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

Epstude, K., & Roese, N. J. (2008). The functional theory of counterfactual thinking. *Personality and Social Psychology Review*, *12*(2), 168–192. https://doi.org/10.1177/1088868308316091

Escobedo, J. R., & Adolphs, R. (2010). Becoming a better person: Temporal remoteness biases autobiographical memories for moral events. *Emotion*. https://doi.org/10.1037/a0018723

Fawcett, J. M., Benoit, R. G., Gagnepain, P., Salman, A., Bartholdy, S., Bradley, C., … Anderson, M. C. (2015). The origins of repetitive thought in rumination: Separating cognitive style from deficits in inhibitory control over memory. *Journal of Behavior Therapy and Experimental Psychiatry*, *47*, 1–8. https://doi.org/10.1016/j.jbtep.2014.10.009

Florkowski, C. M. (2008). Sensitivity, specificity, receiver-operating characteristic (ROC) curves and likelihood ratios: communicating the performance of diagnostic tests. *The Clinical Biochemist. Reviews*, *29 Suppl 1*(Suppl 1), S83-7. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/18852864

Freitas, J. De, Cikara, M., Grossmann, I., & Schlegel, R. (2018). Moral Goodness Is the Essence of Personal Identity. *Trends in Cognitive Sciences*, *22*(9), 739–740. https://doi.org/10.1016/J.TICS.2018.05.006

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

Gagnepain, P., Hulbert, J., & Anderson, M. C. (2017). Parallel regulation of memory and emotion supports the suppression of intrusive memories. *Journal of Neuroscience*, *37*(27), 6423–6441. https://doi.org/10.1523/JNEUROSCI.2732-16.2017

Gillie, B. L., Vasey, M. W., & Thayer, J. F. (2014). Heart Rate Variability Predicts Control Over Memory Retrieval. *Psychological Science*, *25*(2), 458–465. https://doi.org/10.1177/0956797613508789

Gilovich, T., & Medvec, V. H. (1995). The Experience of Regret: What, When, and Why. *Psychological Review*, *102*(2), 379–395. https://doi.org/10.1037/0033-295X.102.2.379

Gross, J. J. (2015). Emotion Regulation: Current Status and Future Prospects. *Http://Dx.Doi.Org/10.1080/1047840X.2014.940781*, *26*(1), 1–26. https://doi.org/10.1080/1047840X.2014.940781

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

Hanslmayr, S., Staresina, B. P., & Bowman, H. (2016, January 1). Oscillations and Episodic Memory: Addressing the Synchronization/Desynchronization Conundrum. *Trends in*

Neurosciences. Elsevier Ltd. https://doi.org/10.1016/j.tins.2015.11.004

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

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

Hellerstedt, R., Moccia, A., Brunskill, C. M., Bowman, H., & Bergström, Z. M. (2021). Aging reduces EEG markers of recognition despite intact performance: Implications for forensic memory detection. *Cortex, 140*, 80–97. https://doi.org/10.1016/j.cortex.2021.03.015

Hertel, P. T., & Calcaterra, G. (2005). Intentional forgetting benefits from thought substitution. *Psychonomic Bulletin and Review*. Springer New York LLC. https://doi.org/10.3758/BF03193792

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

Hitchcock, C., Werner-Seidler, A., Blackwell, S. E., & Dalgleish, T. (2017). Autobiographical episodic memory-based training for the treatment of mood, anxiety and stress-related disorders: A systematic review and meta-analysis. *Clinical Psychology Review, 52*, 92–107. https://doi.org/10.1016/j.cpr.2016.12.003

Hofmann, W., Wisneski, D. C., Brandt, M. J., & Skitka, L. J. (2014). Morality in everyday life. *Science*, *345*(6202), 1340–1343. https://doi.org/10.1126/SCIENCE.1251560

Hoormann, J., Falkenstein, M., Schwarzenau, P., & Hohnsbein, J. (1998). *Methods for the quantification and statistical testing of ERP differences across conditions*. *Behavior Research Methods, Instruments, & Computers* (Vol. 30).

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

Hu, X., Bergström, Z. M., Gagnepain, P., & Anderson, M. C. (2017). Suppressing Unwanted Memories Reduces Their Unintended Influences. *Current Directions in Psychological Science*, *26*(2), 197–206. https://doi.org/10.1177/0963721417689881

Hulbert, J. C., & Anderson, M. C. (2018). What doesn't kill you makes you stronger: Psychological trauma and its relationship to enhanced memory control. *Journal of Experimental Psychology: General*, *147*(12), 1931–1949. https://doi.org/10.1037/xge0000461

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

Joormann, J., Hertel, P. T., LeMoult, J., & Gotlib, I. H. (2009). Training Forgetting of Negative Material in Depression. *Journal of Abnormal Psychology*, *118*(1), 34–43. https://doi.org/10.1037/a0013794

Kahneman, D., & Miller, D. T. (1986). Norm Theory. Comparing Reality to Its Alternatives. *Psychological Review*, *93*(2), 136–153. https://doi.org/10.1037/0033-295X.93.2.136

Kappes, A., & Crockett, M. J. (2016). The Benefits and Costs of a Rose-Colored Hindsight.

Trends in Cognitive Sciences, *20*(9), 644–646. https://doi.org/10.1016/j.tics.2016.06.009

Keller, A. S., Payne, L., & Sekuler, R. (2017). Characterizing the roles of alpha and theta oscillations in multisensory attention. *Neuropsychologia*, *99*, 48–63. https://doi.org/10.1016/j.neuropsychologia.2017.02.021

Ketz, N., O’Reilly, R. C., & Curran, T. (2014). Classification aided analysis of oscillatory signatures in controlled retrieval. *NeuroImage*, *85*, 749–760. https://doi.org/10.1016/j.neuroimage.2013.06.077

Kilner, J. M. (2013). Bias in a common EEG and MEG statistical analysis and how to avoid it. *Clinical Neurophysiology*, *124*(10), 2062–2063. https://doi.org/10.1016/j.clinph.2013.03.024

Kirschbaum, C., Pirke, K.-M., & Hellhammer, D. H. (1993). The ‘Trier Social Stress Test’ – A Tool for Investigating Psychobiological Stress Responses in a Laboratory Setting. *Neuropsychobiology*, *28*(1–2), 76–81. https://doi.org/10.1159/000119004

klein Selle, N., Gueta, C., Harpaz, Y., Deouell, L. Y., & Ben-Shakhar, G. (2021). Brain-based concealed memory detection is driven mainly by orientation to salient items. *Cortex, 136*, 41–55. https://doi.org/10.1016/j.cortex.2020.12.010

Knez, I., & Nordhall, O. (2017). Guilt as a motivator for moral judgment: An autobiographical memory study. *Frontiers in Psychology*. https://doi.org/10.3389/fpsyg.2017.00750

Kouchaki, M., & Gino, F. (2016). Memories of unethical actions become obfuscated over time. *Proceedings of the National Academy of Sciences*. https://doi.org/10.1073/pnas.1523586113

Legrand, N., Etard, O., Vandevelde, A., Pierre, M., Viader, F., Clochon, P., … Gagnepain, P.

(2020). Long-term modulation of cardiac activity induced by inhibitory control over emotional memories. *Scientific Reports, 10*(1), 1–19. https://doi.org/10.1038/s41598-020-71858-2

Levy, B. J., & Anderson, M. C. (2002, July 1). Inhibitory processes and the control of memory retrieval. *Trends in Cognitive Sciences*. Elsevier Current Trends. https://doi.org/10.1016/S1364-6613(02)01923-X

Levy, B. J., & Anderson, M. C. (2012). Purging of memories from conscious awareness tracked in the human brain. *Journal of Neuroscience*, *32*(47), 16785–16794. https://doi.org/10.1523/JNEUROSCI.2640-12.2012

Leynes, P. A., Bruett, H., Krizan, J., & Veloso, A. (2017). What psychological process is reflected in the FN400 event-related potential component? *Brain and Cognition*, *113*, 142–154. https://doi.org/10.1016/J.BANDC.2017.02.004

López-Caneda, E., Crego, A., Campos, A. D., González-Villar, A., & Sampaio, A. (2019). The Think/No-Think Alcohol Task: A New Paradigm for Assessing Memory Suppression in Alcohol-Related Contexts. *Alcoholism: Clinical and Experimental Research*, *43*(1), 36–47. https://doi.org/10.1111/acer.13916

Luck, S. J. (Steven J. (2005). *An introduction to the event-related potential technique*. MIT Press.

Lykken, D. T. (1959). The GSR in the detection of guilt. *Journal of Applied Psychology, 43*(6), 385–388. https://doi.org/10.1037/h0046060

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods, 164*(1), 177–190. https://doi.org/10.1016/J.JNEUMETH.2007.03.024

Marks, E. H., Franklin, A. R., & Zoellner, L. A. (2018). Can’t get it out of my mind: A systematic review of predictors of intrusive memories of distressing events. *Psychological Bulletin*. https://doi.org/10.1037/bul0000132

Mary, A., Dayan, J., Leone, G., Postel, C., Fraisse, F., Malle, C., … Gagnepain, P. (2020). Resilience after trauma: The role of memory suppression. *Science*, *367*(6479). https://doi.org/10.1126/science.aay8477

Marzi, T., Regina, A., & Righi, S. (2014a). Emotions shape memory suppression in trait anxiety. *Frontiers in Psychology*, *4*(JAN), 1001. https://doi.org/10.3389/fpsyg.2013.01001

Marzi, T., Regina, A., & Righi, S. (2014b). Emotions shape memory suppression in trait anxiety. *Frontiers in Psychology*, *4*(JAN). https://doi.org/10.3389/fpsyg.2013.01001

Mather, M., & Thayer, J. (2018). How heart rate variability affects emotion regulation brain networks. *Current Opinion in Behavioral Sciences*, *19*, 98. https://doi.org/10.1016/J.COBEHA.2017.12.017

Mayes, A., Montaldi, D., & Migo, E. (2007, March). Associative memory and the medial temporal lobes. *Trends in Cognitive Sciences*. Trends Cogn Sci. https://doi.org/10.1016/j.tics.2006.12.003

Mecklinger, A., & Jäger, T. (2012). Episodic memory storage and retrieval: Insights from electrophysiological measures. In *Neuroimaging of Human Memory: Linking Cognitive Processes to Neural Systems*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199217298.003.0020

Mecklinger, A., Parra, M., & Waldhauser, G. T. (2009). ERP correlates of intentional forgetting. *Brain Research*, *1255*, 132–147.

https://doi.org/10.1016/J.BRAINRES.2008.11.073

Meijer, E. H., Selle, N. K., Elber, L., & Ben-Shakhar, G. (2014). Memory detection with the Concealed Information Test: A meta analysis of skin conductance, respiration, heart rate, and P300 data. *Psychophysiology*, *51*(9), 879–904. https://doi.org/10.1111/psyp.12239

Meixner, J. B., & Rosenfeld, J. P. (2011). A mock terrorism application of the P300-based concealed information test. *Psychophysiology*, *48*(2), 149–154. https://doi.org/10.1111/j.1469-8986.2010.01050.x

Moore, C. (2015). Moral disengagement. *Current Opinion in Psychology*, *6*, 199–204. https://doi.org/10.1016/J.COPSYC.2015.07.018

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, B. D., Muscatell, K. A., & Kensinger, E. A. (2011). Effects of emotion and age on performance during a think/no-think memory task. *Psychology and Aging*, *26*(4), 940–955. https://doi.org/10.1037/a0023214

Murray, J. G., Ouyang, G., & Donaldson, D. I. (2019). Compensation of trial-to-trial latency jitter reveals the parietal retrieval success effect to be both variable and thresholded in older adults. *Frontiers in Aging Neuroscience, 10*(JUL), 179. https://doi.org/10.3389/fnagi.2019.00179

Nader, K., Schafe, G. E., & Le Doux, J. E. (2000). Fear memories require protein synthesis in the amygdala for reconsolidation after retrieval. *Nature*, *406*(6797), 722–726. https://doi.org/10.1038/35021052

Nolen-Hoeksema, S. (2000). The role of rumination in depressive disorders and mixed anxiety/depressive symptoms. *Journal of Abnormal Psychology*, *109*(3), 504–511. https://doi.org/10.1037/0021-843X.109.3.504

Nørby, S., Lange, M., & Larsen, A. (2010). Forgetting to forget: On the duration of voluntary suppression of neutral and emotional memories. *Acta Psychologica*, *133*(1), 73–80. https://doi.org/10.1016/j.actpsy.2009.10.002

Noreen, S., Bierman, R. N., & MacLeod, M. D. (2014). Forgiving You Is Hard, but Forgetting Seems Easy: Can Forgiveness Facilitate Forgetting? *Psychological Science*. https://doi.org/10.1177/0956797614531602

Noreen, S., & De Fockert, J. W. (2017). The role of cognitive load in intentional forgetting using the think/no-think task. *Experimental Psychology*, *64*(1), 14–26. https://doi.org/10.1027/1618-3169/a000347

Noreen, S., & Macleod, M. D. (2013). It’s all in the detail: intentional forgetting of autobiographical memories using the autobiographical think/no-think task. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *39*(2), 375–393. https://doi.org/10.1037/a0028888

Noreen, S., & MacLeod, M. D. (2014). To think or not to think, that is the question: Individual differences in suppression and rebound effects in autobiographical memory. *Acta Psychologica*, *145*(1), 84–97. https://doi.org/10.1016/j.actpsy.2013.10.011

Noreen, S., O’Connor, A. R., & MacLeod, M. D. (2016). Neural correlates of direct and indirect suppression of autobiographical memories. *Frontiers in Psychology*. https://doi.org/10.3389/fpsyg.2016.00379

Noreen, S., & Ridout, N. (2016a). Examining the impact of thought substitution on intentional forgetting in induced and naturally occurring dysphoria. *Psychiatry Research*, *241*, 280–288. https://doi.org/10.1016/j.psychres.2016.04.086

Noreen, S., & Ridout, N. (2016b). Intentional forgetting in dysphoria: Investigating the inhibitory effects of thought substitution using independent cues. *Journal of Behavior Therapy and Experimental Psychiatry*, *52*, 110–118. https://doi.org/10.1016/j.jbtep.2016.04.002

Nyhus, E., & Curran, T. (2010, June). Functional role of gamma and theta oscillations in episodic memory. *Neuroscience and Biobehavioral Reviews*. Neurosci Biobehav Rev. https://doi.org/10.1016/j.neubiorev.2009.12.014

Ochsner, K. N., & Gross, J. J. (2005, May 1). The cognitive control of emotion. *Trends in Cognitive Sciences*. Elsevier Ltd. https://doi.org/10.1016/j.tics.2005.03.010

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J. M. (2011). FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational Intelligence and Neuroscience*, *2011*. https://doi.org/10.1155/2011/156869

Opitz, B., & Cornell, S. (2006). Contribution of Familiarity and Recollection to Associative Recognition Memory: Insights from Event-related Potentials. *Journal of Cognitive Neuroscience*, *18*(9), 1595–1605. https://doi.org/10.1162/jocn.2006.18.9.1595

Osipova, D., Takashima, A., Oostenveld, R., Fernández, G., Maris, E., & Jensen, O. (2006). Theta and gamma oscillations predict encoding and retrieval of declarative memory. *Journal of Neuroscience*, *26*(28), 7523–7531. https://doi.org/10.1523/JNEUROSCI.1948-06.2006

Osugi, A. (2018). Field Findings From the Concealed Information Test in Japan. *Detecting*

Concealed Information and Deception, 97–121. https://doi.org/10.1016/B978-0-12-812729-2.00005-7

Osugi, A., & Ohira, H. (2017). Emotional Arousal at Memory Encoding Enhanced P300 in the Concealed Information Test. *Frontiers in Psychology, 8*, 2334. https://doi.org/10.3389/fpsyg.2017.02334

Parikh, N., Ruzic, L., Stewart, G. W., Spreng, R. N., & De Brigard, F. (2018). What if? Neural activity underlying semantic and episodic counterfactual thinking. *NeuroImage*. https://doi.org/10.1016/j.neuroimage.2018.05.053

Peirce, J. W. (2008). Generating stimuli for neuroscience using PsychoPy. *Frontiers in Neuroinformatics, 2*, 10. https://doi.org/10.3389/neuro.11.010.2008

Pernet, C. R., Latinus, M., Nichols, T. E., & Rousselet, G. A. (2015). Cluster-based computational methods for mass univariate analyses of event-related brain potentials/fields: A simulation study. *Journal of Neuroscience Methods*, *250*, 85–93. https://doi.org/10.1016/j.jneumeth.2014.08.003

Pinto, Y., Leij, A. R. van der, Sligte, I. G., Lamme, V. A. F., & Scholte, H. S. (2013). Bottom-up and top-down attention are independent. *Journal of Vision*, *13*(3), 16–16. https://doi.org/10.1167/13.3.16

Polich, J. (2007, October). Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology*. NIH Public Access. https://doi.org/10.1016/j.clinph.2007.04.019

Quaedflieg, C. W. E. M., Schneider, T. R., Daume, J., & Engel, A. K. (2020). Stress Impairs Intentional Memory Control through Altered Theta Oscillations in Lateral Parietal Cortex. *Journal of Neuroscience, 40*(40), 7739–7748.

Reczek, R. W., Irwin, J. R., Zane, D. M., & Ehrich, K. R. (2018). That's Not How I

Remember It: Willfully Ignorant Memory for Ethical Product Attribute Information. *Journal of Consumer Research*, *45*(1), 185–207. https://doi.org/10.1093/jcr/ucx120

Renoult, L., Davidson, P. S. R., Schmitz, E., Park, L., Campbell, K., Moscovitch, M., & Levine, B. (2015). Autobiographically Significant Concepts: More Episodic than Semantic in Nature? An Electrophysiological Investigation of Overlapping Types of Memory. *Journal of Cognitive Neuroscience, 27*(1), 57–72. https://doi.org/10.1162/JOCN_A_00689

Rim, S. Y., & Summerville, A. (2014). How Far to the Road Not Taken? The Effect of Psychological Distance on Counterfactual Direction. *Personality and Social Psychology Bulletin*, *40*(3), 391–401. https://doi.org/10.1177/0146167213513304

Roach, B. J., & Mathalon, D. H. (2008, September). Event-related EEG time-frequency analysis: An overview of measures and an analysis of early gamma band phase locking in schizophrenia. *Schizophrenia Bulletin*. Schizophr Bull. https://doi.org/10.1093/schbul/sbn093

Rosenfeld, J. P. (2018). *Detecting Concealed Information and Deception: Recent Developments*. Retrieved from http://groups.psych.northwestern.edu/rosenfeld/documents/Rosenfeld,

Rosenfeld, J. P. (2019). P300 in detecting concealed information and deception: A review. *Psychophysiology*, e13362. https://doi.org/10.1111/psyp.13362

Rosenfeld, J. P., & Donchin, E. (2015). Resampling (bootstrapping) the mean: A definite do. *Psychophysiology*, *52*(7), 969–972. https://doi.org/10.1111/psyp.12421

Rosenfeld, J. P., Reinhart, A. M., Bhatt, M., Ellwanger, J., Gora, K., Sekera, M., & Sweet, J. (1998). P300 correlates of simulated malingered amnesia in a matching-to-sample task:

topographic analyses of deception versus truthtelling responses. *International Journal of Psychophysiology, 28*(3), 233–247. https://doi.org/10.1016/S0167-8760(97)00084-6

Rosenfeld, J. P., Ward, A., Drapekin, J., Labkovsky, E., & Tullman, S. (2017). Instructions to suppress semantic memory enhances or has no effect on P300 in a concealed information test (CIT). *International Journal of Psychophysiology*, *113*, 29–39. https://doi.org/10.1016/j.ijpsycho.2017.01.001

Rosenfeld, J. P., Ward, A., Meijer, E. H., & Yukhnenko, D. (2016). Bootstrapping the P300 in diagnostic psychophysiology: How many iterations are needed? https://doi.org/10.1111/psyp.12789

Rowling, J. K. (2000). *Harry Potter and the Goblet of Fire*. Bloomsbury.

Rugg, M. D., & Curran, T. (2007, June 1). Event-related potentials and recognition memory. *Trends in Cognitive Sciences*. Elsevier. https://doi.org/10.1016/j.tics.2007.04.004

Sacchet, M. D., Levy, B. J., Hamilton, J. P., Maksimovskiy, A., Hertel, P. T., Joormann, J., … Gotlib, I. H. (2017). Cognitive and neural consequences of memory suppression in major depressive disorder. *Cognitive, Affective and Behavioral Neuroscience*, *17*(1), 77–93. https://doi.org/10.3758/s13415-016-0464-x

Sahakyan, L., Delaney, P. F., Foster, N. L., & Abushanab, B. (2013). List-Method Directed Forgetting in Cognitive and Clinical Research: A Theoretical and Methodological Review. *Psychology of Learning and Motivation - Advances in Research and Theory*, *59*, 131–189. https://doi.org/10.1016/B978-0-12-407187-2.00004-6

Sakaki, M., Kuhbandner, C., Mather, M., & Pekrun, R. (2014). Memory suppression can help people “unlearn” behavioral responses-but only for nonemotional memories. *Psychonomic Bulletin and Review*, *21*(1), 136–141. https://doi.org/10.3758/s13423-013-

0480-6

Sassenhagen, J., & Draschkow, D. (2019). Cluster-based permutation tests of MEG/EEG data do not establish significance of effect latency or location. *Psychophysiology*, *56*(6), e13335. https://doi.org/10.1111/psyp.13335

Schacter, D. L., Benoit, R. G., De Brigard, F., & Szpunar, K. K. (2015). Episodic future thinking and episodic counterfactual thinking: Intersections between memory and decisions. *Neurobiology of Learning and Memory*. https://doi.org/10.1016/j.nlm.2013.12.008

Schacter, D. L., Coyle, J. T., & Harvard Center for the Study of Mind, B. (1995). Memory distortion : how minds, brains, and societies reconstruct the past, 417.

Schmitz, T. W., Correia, M. M., Ferreira, C. S., Prescot, A. P., & Anderson, M. C. (2017). Hippocampal GABA enables inhibitory control over unwanted thoughts. *Nature Communications*, *8*(1), 1–12. https://doi.org/10.1038/s41467-017-00956-z

Shu, L. L., Gino, F., & Bazerman, M. H. (2011). Dishonest deed, clear conscience: When cheating leads to moral disengagement and motivated forgetting. *Personality and Social Psychology Bulletin*. https://doi.org/10.1177/0146167211398138

Soskins, M., Rosenfeld, J., & Niendam, T. (2001). Peak-to-peak measurement of P300 recorded at 0.3 Hz high pass filter settings in intraindividual diagnosis: complex vs. simple paradigms. *International Journal of Psychophysiology*, *40*(2), 173–180. https://doi.org/10.1016/S0167-8760(00)00154-9

St. Jacques, P. L., Rubin, D. C., LaBar, K. S., & Cabeza, R. (2008). The short and long of it: Neural correlates of temporal-order memory for autobiographical events. *Journal of Cognitive Neuroscience*, *20*(7), 1327–1341. https://doi.org/10.1162/jocn.2008.20091

St Jacques, P. L., & De Brigard, F. (n.d.). Neural Correlates of Autobiographical Memory Methodological Considerations.

Stanley, M. L., & De Brigard, F. (2019). Moral Memories and the Belief in the Good Self. *Current Directions in Psychological Science, 28*(4), 387–391. https://doi.org/10.1177/0963721419847990

Stanley, M. L., Henne, P., Iyengar, V., Sinnott-Armstrong, W., & Brigard, F. De. (2017). I'm not the person I used to be: The self and autobiographical memories of immoral actions. *Journal of Experimental Psychology: General*. https://doi.org/10.1037/xge0000317

Stanley, M. L., Parikh, N., Stewart, G. W., & De Brigard, F. (2017). Emotional intensity in episodic autobiographical memory and counterfactual thinking. *Consciousness and Cognition*. https://doi.org/10.1016/j.concog.2016.12.013

Stanley, M. L., Yang, B. W., & de Brigard, F. (2018). No evidence for unethical amnesia for imagined actions: A failed replication and extension. *Memory and Cognition*. https://doi.org/10.3758/s13421-018-0803-y

Stephens, E., Braid, A., & Hertel, P. T. (2013). Suppression-Induced Reduction in the Specificity of Autobiographical Memories. *Clinical Psychological Science, 1*(2), 163–169. https://doi.org/10.1177/2167702612467773

Storm, B. C., & Nestojko, J. F. (2010). Successful inhibition, unsuccessful retrieval: Manipulating time and success during retrieval practice. *Memory, 18*(2), 99–114. https://doi.org/10.1080/09658210903107853

Stramaccia, D. F., Meyer, A. K., Rischer, K. M., Fawcett, J. M., & Benoit, R. G. (2020). Memory suppression and its deficiency in psychological disorders: A focused meta-analysis. *Journal of Experimental Psychology: General*.

https://doi.org/10.1037/xge0000971

Streb, M., Mecklinger, A., Anderson, M. C., Johanna, L. H., & Michael, T. (2016). Memory control ability modulates intrusive memories after analogue trauma. *Journal of Affective Disorders*, *192*, 134–142. https://doi.org/10.1016/j.jad.2015.12.032

Strohminger, N., Knobe, J., & Newman, G. (2017). The True Self: A Psychological Concept Distinct From the Self: *Https://Doi.Org/10.1177/1745691616689495*, *12*(4), 551–560. https://doi.org/10.1177/1745691616689495

Sutin, A. R., & Robins, R. W. (2007). Phenomenology of autobiographical memories: The Memory Experiences Questionnaire. *Memory*, *15*(4), 390–411. https://doi.org/10.1080/09658210701256654

Tallon-Baudry, C., & Bertrand, O. (1999, April 1). Oscillatory gamma activity in humans and its role in object representation. *Trends in Cognitive Sciences*. Elsevier Current Trends. https://doi.org/10.1016/S1364-6613(99)01299-1

Tanguay, A. N., Benton, L., Romio, L., Sievers, C., Davidson, P. S. R., & Renoult, L. (2018). The ERP correlates of self-knowledge: Are assessments of one's past, present, and future traits closer to semantic or episodic memory? *Neuropsychologia*, *110*, 65–83. https://doi.org/10.1016/j.neuropsychologia.2017.10.024

Todd, R. M., Cunningham, W. A., Anderson, A. K., & Thompson, E. (2012, July 1). Affect-biased attention as emotion regulation. *Trends in Cognitive Sciences*. Elsevier Current Trends. https://doi.org/10.1016/j.tics.2012.06.003

van der Hart, O. (1986). The use of imagery in phase 1 treatment of clients with complex dissociative disorders. *Brown {\&} Fromm*. https://doi.org/10.3402/ejpt.v3i0.8458

van Schie, K., & Anderson, M. C. (2017). Successfully controlling intrusive memories is harder when control must be sustained. *Memory*. https://doi.org/10.1080/09658211.2017.1282518

van Schie, K., Geraerts, E., & Anderson, M. C. (2013). Emotional and non-emotional memories are suppressible under direct suppression instructions. *Cognition and Emotion*, *27*(6), 1122–1131. https://doi.org/10.1080/02699931.2013.765387

Vilberg, K. L., Moosavi, R. F., & Rugg, M. D. (2006). The relationship between electrophysiological correlates of recollection and amount of information retrieved. *Brain Research*, *1122*(1), 161–170. https://doi.org/10.1016/j.brainres.2006.09.023

Wagner, J., Wessel, J. R., Ghahremani, A., & Aron, A. R. (2018). Establishing a right frontal beta signature for stopping action in scalp EEG: Implications for testing inhibitory control in other task contexts. *Journal of Cognitive Neuroscience*, *30*(1), 107–118. https://doi.org/10.1162/jocn_a_01183

Waldhauser, G. T., Bäuml, K. H. T., & Hanslmayr, S. (2015). Brain oscillations mediate successful suppression of unwanted memories. *Cerebral Cortex*. https://doi.org/10.1093/cercor/bhu138

Waldhauser, G. T., Lindgren, M., & Johansson, M. (2012). Intentional suppression can lead to a reduction of memory strength: Behavioral and electrophysiological findings. *Frontiers in Psychology*, *3*(OCT), 401. https://doi.org/10.3389/fpsyg.2012.00401

Ward, A. C., & Rosenfeld, J. P. (2017). Attempts to Suppress Episodic Memories Fail but do Produce Demand: Evidence from the P300-Based Complex Trial Protocol and an Implicit Memory Test. *Applied Psychophysiology and Biofeedback*, *42*(1), 13–26. https://doi.org/10.1007/s10484-016-9348-7

Watson, L. A., Berntsen, D., Kuyken, W., & Watkins, E. R. (2012). The characteristics of involuntary and voluntary autobiographical memories in depressed and never depressed individuals. *Consciousness and Cognition*, *21*(3), 1382–1392. https://doi.org/10.1016/j.concog.2012.06.016

White, K., & Lehman, D. R. (2005). Culture and social comparison seeking: The role of self-motives. *Personality and Social Psychology Bulletin*, *31*(2), 232–242. https://doi.org/10.1177/0146167204271326

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., & Ranganath, C. (2012). Electrophysiological Correlates of Episodic Memory Processes. In *The Oxford Handbook of Event-Related Potential Components*. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780195374148.013.0187

Wimber, M., Rutschmann, R. M., Greenlee, M. W., & Bäuml, K. H. (2009). Retrieval from episodic memory: Neural mechanisms of interference resolution. *Journal of Cognitive Neuroscience, 21*(3), 538–549. https://doi.org/10.1162/jocn.2009.21043

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), 1113. https://doi.org/10.1038/s41598-018-19570-0

Yapan, S., Türkçapar, M. H., & Boysan, M. (2020). Rumination, automatic thoughts, dysfunctional attitudes, and thought suppression as transdiagnostic factors in depression and anxiety. *Current Psychology*, 1–17. https://doi.org/10.1007/s12144-020-01086-4

Zoumpoulaki, A., Alsufyani, A., & Bowman, H. (2015). Resampling the peak, some dos and don’ts. *Psychophysiology*, *52*(3), 444–448. https://doi.org/10.1111/psyp.12363

Appendix

Appendix A: Supplemental information for Experiment 1 (Chapter 3)

A.1. Autobiographical memory generation questionnaire example

Recall one specific event from your personal past where your action(s) was **morally wrong.** Usually, people write about events where they harmed someone physically or emotionally, lied, cheated, were unfaithful, or when they stole something. People also write about events from their past when they felt guilty, shameful, regretful, or sneaky. You can use any one, or all, of the above examples to help you think of events from your past, or you can write about other events that were morally wrong. Once you have come up with an event fitting this description, write about this event in two to five sentences using the space below. You will have a maximum of 3 minutes to complete this task. Please describe your actions, who was involved, where the event took place, and how you felt. Do this without revealing the names of people involved. Please make sure you include a specific time and place. Please also ensure that this is a personal event. **Please note that we require you to write about a SPECIFIC event that took place in a particular location and at a particular time. Avoid writing about events that easily blend into other events. Please write as many details as you can about this event for the next 3 minutes.**

____

____

____

____

____

You are now required to come up with a brief title that describes this event. The title needs to be specific and unique enough such that you will be able to recollect this same exact event on day 2, if this title is presented to you.

It should not overlap with titles you have previously written down in this questionnaire. We encourage you to come up with a title that will be recognisable by only you (or people close to you). Examples of good titles are: *NCUK win, big mouth, cherry drops.* Please try to avoid general titles. Examples of general titles are: *lying, helping, grandma, father, friend, guilty, proud, etc.*

As mentioned above, this will be important for day 2 to help you think of this same event.

Please write down a title that best describes this event (maximum 2 words)

____

How long ago did this event occur?

○ Within the past day

○ Within the past week

○ Within the past month

○ Within the past year

○ Within the past two years

○ Within the past three years

○ Within the past four years

○ Within the past 5 years

○ Within the past 6 years

○ Within the past 7 years

○ Within the past 8 years

○ Within the past 9 years

○ Within the past 10 years

○ More than 10 years ago

Overall, how well do you remember the event?

○ Not well at all

○ Slightly well

○ Moderately well

○ Very well

○ Extremely well

How intentional were your actions?

○ Not intentional at all

○ Slightly intentional

○ Moderately intentional

○ Very intentional

○ Extremely intentional

How morally right or morally wrong was the action performed during the event?

○ Very morally wrong

○ Moderately morally wrong

○ Slightly morally wrong

○ Not morally wrong or right

○ Slightly morally right

○ Moderately morally right

○ Very morally right

I-PANAS-SF

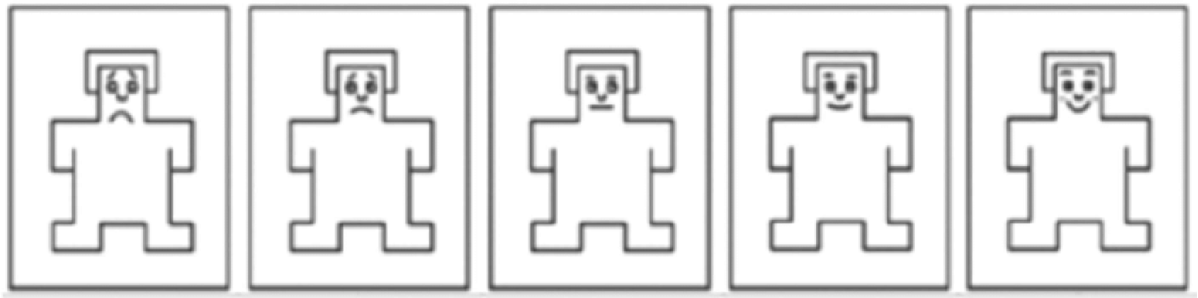
Indicate the extent to which you feel this way at the present moment about the event

	Very slightly or not at all	A little	Moderately	Quite a bit	Extremely
Upset	○	○	○	○	○
Alert	○	○	○	○	○
Ashamed	○	○	○	○	○
Inspired	○	○	○	○	○
Nervous	○	○	○	○	○
Determined	○	○	○	○	○
Guilty	○	○	○	○	○
Active	○	○	○	○	○
Afraid	○	○	○	○	○

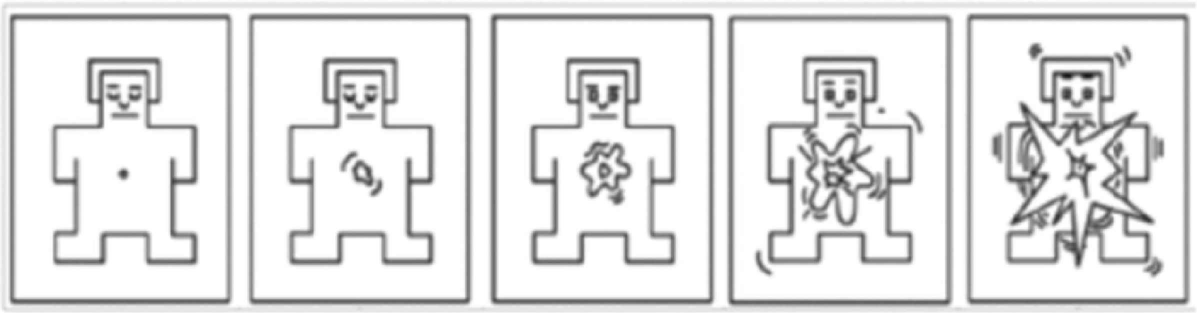
Attentive	○	○	○	○	○

SAM questionnaire

The scale below represents the happy-unhappy dimension, which ranges from a smile to a frown. At one extreme of the happy vs. unhappy scale, you feel happy, pleased, satisfied, contented, and hopeful. If you felt completely *happy* while thinking about the event, then choose the right-most image. The other end of the scale is when you felt completely unhappy, annoyed, unsatisfied, melancholic, despaired, bored. If you felt completely *unhappy* while thinking about the event, then choose the left-most image. The figures also allow you to describe intermediate feelings of *pleasure*, you can do this by choosing any of the images in-between the two extremes. This permits you to make more finely graded ratings of how you felt while thinking about the event.



The scale below represents the excited vs. calm dimension. At one extreme of the scale you felt stimulated, excited, frenzied, jittery, wide-awake, and aroused. If you felt completely *aroused* while thinking about the event, then choose the right-most image. On the other hand, at the other end of the scale, you felt completely relaxed, calm, sluggish, dull, sleepy, unaroused. You can indicate you felt completely *calm* by choosing the left-most image. As with the happy-unhappy scale, you can represent intermediate levels of *arousal* by choosing the images in between the two extremes.



A.2. Experimental phases prior to think/no-think task.

Cue-test phase. The aim of this phase was to test if participants could remember the memory associated to the title. Importantly, these titles were used to aid participants' recall of autobiographical memories in the subsequent think/no-think phases of the experiment. A fixation cross was first presented on a computer screen for 1000ms, followed by a title, displayed for 4000ms. Then, a scale (1 2 3) was shown on the screen indefinitely, until the program recorded a response from the participant. The participants used this scale to indicate how well they can think of the memories related to the titles (*1 = not well at all, 2 = somewhat well 3 = extremely well*), which allowed us to also test if that title evoked the memory. This process of testing the association between titles and memories repeated for all 22 memories generated in session one. The titles were presented in a randomised order.

The titles associated with the memories that participants could not remember (indicated by pressing 1) were first evenly and randomly distributed to think and no-think conditions and then the remaining titles (indicated by pressing either 2 or 3) were assigned to different conditions, using self-written code in PsychoPy programming software (Peirce, 2008).

Practice Think/NoThink phases. The aim of this phase was to allow participants to practice the think/no-think task and the intrusions rating scale. This practice was run in three stages and the two filler titles were used as memory cues in this phase.

In stage one, participants practiced thinking and not-thinking about associated memories without the intrusions rating scale. This stage consisted of 10 trials. The two filler items were presented randomly and evenly across all trials. In the second stage, participants practiced responding to a rating scale: One of three scale labels were presented on the screen for 4000ms, followed by the rating scale (1 2 3). Participants responded to the scale using the 1, 2, 3 number keys on the keyboard, and the next trial began only if the response was accurate. Each label was randomly repeated thrice, leading to a total of nine trials in this stage. In the third stage, participants practiced a combination of the think/no-think task and the rating scale, mirroring the experimental think/no-think phase. The two filler items were assigned to the same condition as in practice stage one. The two filler items were assigned to the same condition as in stage one, but the order of presentation of think/no-think cues was randomised again. This stage consisted of 10 trials. Following the third stage, the main experiment began.

After both stages one and three, the experimenter orally administered a diagnostic questionnaire to ensure that the participants were accurately following the instructions. The experimenter then gave feedback to the participant after each question to reinforce the instructions. This diagnostic questionnaire was also administered at the halfway point of the experimental think/no-think phase.

A.3 Instructions for coding autobiographical memory descriptions

The table below contains descriptions of autobiographical memories that were written at two separate points of time by participants in an experiment, who had to report memories of their past morally right or morally wrong actions.

Each row in the table represents one memory, with the two descriptions of the memories provided side-by-side in a row. Your first task is to rate each description on the following two factors

- Specificity of description
  - Look for details that include “who, what, where, when” - What was the morally relevant act, who was involved, where did it happen, and when did it take place?
  - As the memories are of morally relevant actions, give more weightage to “what” and “who” was involved.
  - Also consider if other specific details are reported in the memory such as a precise object, location, situation, etc. Or if they describe in detail how the action was committed
  - Use this scale: 1 = Not at all specific, 2 = Slightly specific, 3 = Moderately specific, 4 = Very specific, 5 = Extremely specific. 0 = if the memory could not be recalled

- Emotional valence
  - Consider if the emotion of the participant derived from the description is negative or positive: look for specific negative/positive words (example: feeling bad, guilty, etc.)

  - Use this scale: 1 = Very negative, 2 = Moderately negative, 3 = Slightly negative, 4 = Not positive or negative, 5 = Slightly positive, 6 = Moderately positive, 7 = Very positive, 0 = if the memory could not be recalled

Your next task is to compare the two descriptions and indicate how similar they are based on your judgement and using your previous ratings. Please use the following instructions to do this:

- Similarity**:** Please give primary weightage to the specificity of the description, and then the valence.

    - Use this scale: 1 = Not at all similar, 2 = Slightly similar, 3 = Moderately similar, 4 = Very Similar, 5 = Extremely Similar. 0 = if one of the descriptions indicates that the memory could not be recalled

Ratings provided by independent coders

The two tables below present the average ratings provided by the independent coders.

Table A.1. Average similarity ratings provided by independent coders.

	Right	Wrong
Think	3.83(.43)	3.77(.97)
NoThink	3.71(1.09)	3.80(.96)

Note. Similarity was rated on a 5-point scale (1 = Not at all similar, 5 = Extremely similar). Standard deviation is reported in brackets.

Table A.2. Average ratings of specificity and valence across instruction and memory types. Ratings were provided by the experimenter.

	Right		Wrong	
	Pre-TNT	Post-TNT	Pre-TNT	Post-TNT
	Specificity			
Think	2.95(.99)	3.10(1.06)	2.91(.63)	3.08(.65)
NoThink	3.14(1.00)	3.00(1.10)	3.09(1.06)	3.16(1.04)
	Valence			
Think	4.45(.88)	4.48(.84)	3.16(.81)	3.19(.84)
NoThink	4.49(.97)	4.22(.90)	3.30(1.06)	3.28(.88)

Note. Specificity was rated on a 5-point scale (1=Not at all specific to 5 = Extremely specific), emotional valence was measured on a 7-point scale (1=Very Negative, 4 = Not negative or positive, 7 = Very Positive). TNT = Think/no-think Standard deviation is reported in brackets.

A.4. Analysis of autobiographical memory characteristics

First, differences between conditions in phenomenological characteristics was analysed to a) check if the memory type manipulation was successful; b) test the hypothesis that morally wrong memories are perceived to be more distant than morally right memories in terms of vividness and memory age; and c) check if memories were correspondingly assigned into different conditions in terms of phenomenological characteristics. Then, difference in ratings reported in the first session and the second session was analysed.

Memory characteristics and emotions successfully distinguish morally right and wrong memories. A 2 (Instruction Type: Think vs. No-Think) x 2 (Memory Type: Morally Wrong vs. Morally Right) repeated measures ANOVA was conducted with the chosen self-report measures from session one, when participants first described and rated their memories before the think/no-think task. See Table 3 for results from the omnibus ANOVA. There were significant main effects of memory type for several measures: As expected, participants felt more guilty, ashamed, less pleasure, and rated their actions as more morally wrong for morally wrong compared to morally right memories.

There was no main effect of instruction type, nor an interaction between instruction type and memory type for any dependent variable in this analysis, suggesting that memories were assigned reasonably into different conditions. Furthermore, there was no significant difference between memory types in arousal.

As expected, participants indicated that their morally wrong memories were less vivid and occurred more distant in the past than morally right memories. Moreover, participants reported their actions as less intentional in morally wrong than morally right memories. See Figure A.1. and A.2. for an illustration of these results.

Figure A.1. Average ratings compared across memory type during session one as manipulation checks. Note that only Morality was on a scale of 1-7, the other measures were on a scale of 1-5. Error bars represent ±1 SEM. ***p<.001.

Figure A.2. Average ratings compared across memory type during session one for theoretical significance. Error bars represent ±1 SEM. *p<.05 **p<.01 ***p<.001.

Table A.3. Results from omnibus ANOVA using self-report ratings of phenomenological and emotional characteristics of memories as measured in session one (before TNT task).

	Phenomenological memory characteristics											
	Memory Age			Vividness			Intention			Morality		
	F	*p*	η^2_p	*F*	*p*	η^2_p	*F*	*p*	η^2_p	*F*	*p*	η^2_p
IT (T vs .NT)	1.35	0.25	0.04	0.92	0.34	0.03	0.18	0.67	0.01	0.17	0.69	0.01
MT (MW vs. MR)	**36.01**	**<.001**	**0.52**	**4.43**	**0.043**	**0.12**	**62.86**	**<.001**	**0.66**	**791.09**	**<.001**	**0.96**
ITxMT	3.72	0.06	0.10	0.84	0.37	0.03	1.30	0.26	0.05	1.79	0.19	0.05

	I-PANAS-SF						SAM					
	Ashamed			Guilty			Pleasure			Arousal		
	F	*p*	η^2_p	*F*	*p*	η^2_p	*F*	*p*	η^2_p	*F*	*p*	η^2_p
IT (T vs .NT)	0.01	0.91	0.00	0.42	0.52	0.01	0.01	0.92	0.00	1.15	0.29	0.03
MT (MW vs. MR)	**221.78**	**<.001**	**0.87**	**232.02**	**<.001**	**0.88**	**179.11**	**<.001**	**0.84**	0.08	0.78	0.00
ITxMT	0.08	0.78	0.00	0.11	0.74	0.00	0.55	0.46	0.02	0.42	0.52	0.01

Note. IT = Instruction Type, MT = Memory Type, T = Think, NT = No-Think, MW = Morally Wrong, MR = Morally Right. η^2_p = partial eta sq. (effect size). I-PANAS-SF = International – Positive and Negative Affect Scale – Short Form, SAM = Self-Assesssment Manikin. Significant effects ($p < .05$) are shown in bold. $N = 34$.

A.5 Multifactorial ANOVA analysis results of ERP data

The results from block 1 (first half of trials) is presented first, followed by block 2 (second half of trials), and finally the intrusion analysis is reported.

Block 1

A 3 (AP; Anterior, Central, Posterior) x 3 (H; Left, Middle, Right) x 2 (Instruction Type; Think, no-think) x 2 (Memory Type; Morally right, morally wrong) four-way repeated measures ANOVAs was conducted in each block. See Table A.4. for results of the omnibus ANOVA. Topographical maps that illustrate the raw difference in amplitudes between conditions, significant effects and effect sizes that involve the factor condition, and t-values from significant clusters (included for comparison) are illustrated in Figure A.3.

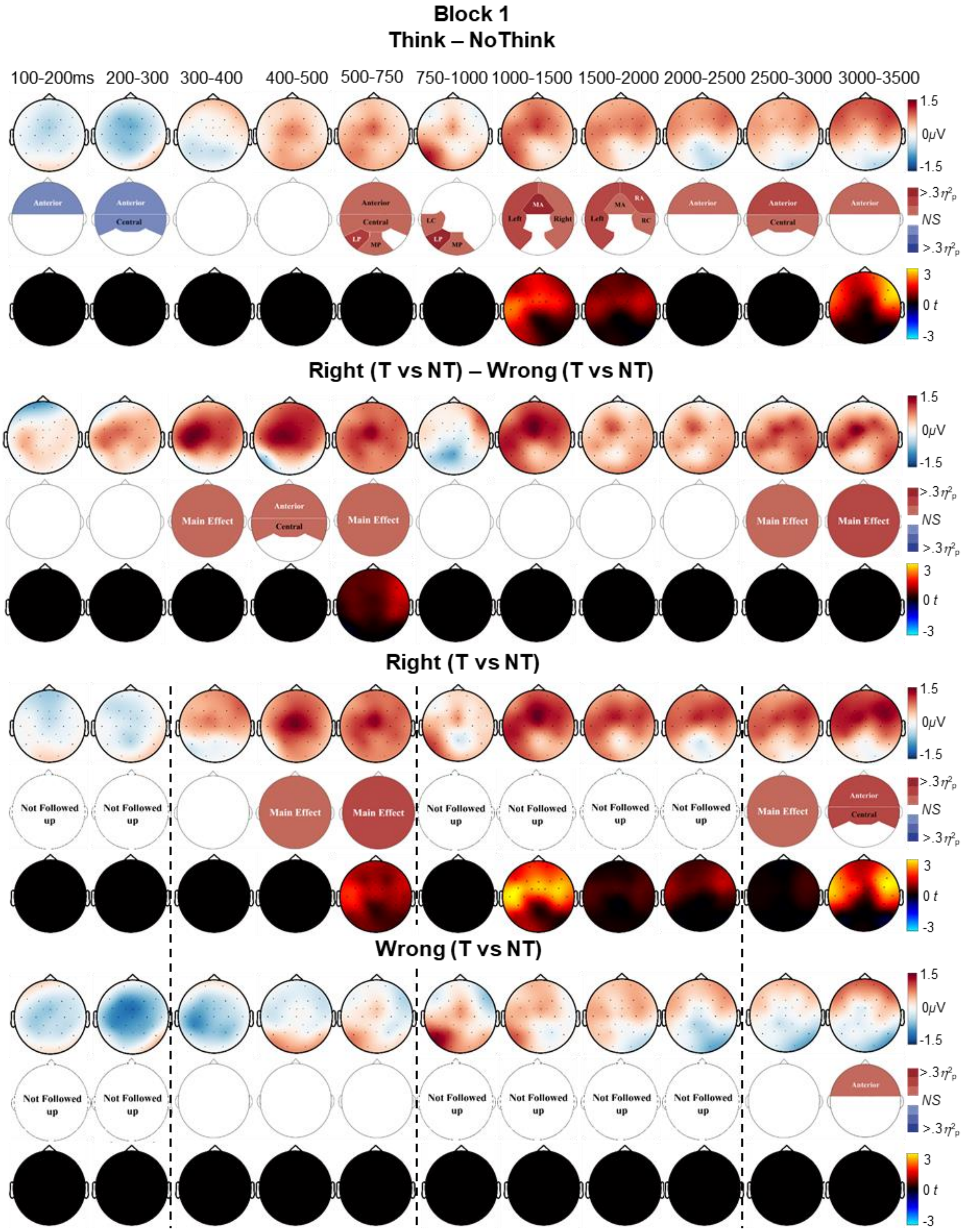


Figure A.3. ERP results for key main effects and follow-up pairwise comparisons in the first block of the TNT task using omnibus ANOVA (middle row) and cluster-based permutation testing (bottom row). Only time windows with significant Instruction Type or Instruction Type x Memory Type interactions were followed up in the ANOVA. For Cluster-based permutation tests, due to the longer time windows, all time windows were used for all comparisons. For each comparison, images on top are topographical maps of ERP amplitude differences between conditions (blue/white/red colourmap, in µV), images in the middle row illustrate statistically significant regions with effect sizes (partial η^2), images in the bottom row illustrates t-values (blue/black/red colourmap, in t) for the differences. White font represents greatest effect size. The colour scale represents magnitude and direction of the effect. Abbreviations: LA = Left Anterior, MA = Mid Anterior, RA = Right Anterior, LC = Left Central, MC = Mid Central, RC = Right Central, LP = Left Parietal, MP = Mid Parietal, RP = Right Parietal.

General think vs. no-think effects. ERPs for the think condition were more negative than the no-think condition from 100-300ms. This negativity was strongest in the mid-anterior region (100-300ms) but was also found in mid-central (200-300ms) regions. There were no significant effects of instruction type from 300-500ms.

A positive think vs. no-think difference in ERP amplitudes was found in all time windows beginning from 500ms till the end of the epoch (3500ms). In all these time windows, the instruction type difference interacted significantly with either one or both scalp factors. This effect was maximal in the left-parietal region early on (500-1000ms windows), and then maximal in mid-anterior (1000-1500ms), right-anterior (1500-2000ms), and generally anterior regions (2000-3500ms). This effect was also present in mid-parietal (500-750ms), central (500-750ms), left-central (750-2000ms), right-central (1500-2000ms), and central (2500-3000ms) regions.

Interaction between moral memory type and instruction type. A significant interaction effect between instruction type and memory type was found in ANOVA results from 300-750ms time windows. This effect did not interact with scalp factors in 300-400ms and 500-750 time windows, but it did interact with the anterior/posterior factor in the 400-500ms time window and was found to be maximal in anterior regions. This interaction was also significant in the latter time windows (2500-3500ms).

Early positivity is present for morally right but not morally wrong memories. From 400-750ms, follow up ANOVAs indicated that ERPs were significantly more positive in the think than in the no-think condition for morally right memories, but for morally wrong memories the difference between think and no-think ERPs was not significant. This pattern was spread across the whole scalp. Although the interaction was significant in the 300-400ms time window, follow-up tests in this time window did not reveal significant effects between think and no-think for either memory type. Interestingly, the topographical plots (see Figure S.6) indicate a positive think vs. no-think difference in ERPs of morally right memories, but a negative difference for morally wrong memories, indicating a polarity reversal. This explains the significant interaction but non-significant pairwise effects.

Interestingly, morally wrong memories elicited significantly larger no-think ERPs than morally right memories in the ANOVA results, but a significant difference between morally right and wrong ERPs was not found in the think condition. This pattern was spread across the whole scalp (300-400ms) but was later localised in the anterior region (400-500ms). This effect was not found in the 500-750ms time window. The cluster-based tests did not reveal any significant clusters for these comparisons.

Positive think vs. no-think ERP differences are more sustained for morally right than wrong memories. There was a think vs. no-think positivity present for morally right memories across the scalp (2500-3000ms), and in anterior and central regions (3000-3500ms), but this positivity was found only in the 3000-3500ms time window in anterior regions for morally wrong memories. The interaction contrast was not significant in the cluster-test so the follow-up cluster tests were not conducted.

Table A.4. ANOVA results from the omnibus tests in first half of TNT task.

	100 to 200ms			200 to 300ms			300 to 400ms			400 to 500ms			500 to 750ms			750 to 1000ms		
	F	*p*	$\eta 2_p$	*F*	*p*	$\eta 2_p$	*F*	*p*	$\eta 2_p$	*F*	*p*	$\eta 2_p$	*F*	*p*	$\eta 2_p$	*F*	*p*	$\eta 2_p$
IT	2.02	0.16	0.06	3.43	0.07	0.09	0.05	0.82	0.00	2.35	0.14	0.07	**6.12**	**0.019**	**0.16**	2.80	0.10	0.08
MT	0.41	0.53	0.01	1.01	0.32	0.03	1.04	0.31	0.03	0.48	0.49	0.01	0.47	0.50	0.01	0.09	0.76	0.00
APxIT	2.83	0.09	0.08	2.60	0.10	0.07	2.93	0.07	0.08	0.07	0.89	0.00	0.04	0.91	0.00	1.23	0.29	0.04
HxIT	0.13	0.84	0.00	1.56	0.22	0.05	0.64	0.53	0.02	1.01	0.37	0.03	2.12	0.13	0.06	**3.17**	**0.048**	**0.09**
APxHxIT	**4.96**	**0.005**	**0.13**	**3.23**	**0.028**	**0.09**	0.93	0.40	0.03	0.41	0.67	0.01	**2.76**	**0.03**	**0.08**	**6.78**	**0.001**	**0.17**
APxMT	0.73	0.43	0.02	0.34	0.59	0.01	1.19	0.29	0.04	1.35	0.26	0.04	2.59	0.10	0.07	0.32	0.62	0.01
HxMT	0.45	0.64	0.01	1.22	0.30	0.04	**5.38**	**0.007**	**0.14**	0.70	0.50	0.02	0.26	0.77	0.01	0.54	0.59	0.02
APxHxMT	0.82	0.51	0.02	0.21	0.93	0.01	1.19	0.32	0.04	0.39	0.82	0.01	0.47	0.76	0.01	0.66	0.62	0.02
ITxMT	0.41	0.53	0.01	1.73	0.20	0.05	**4.65**	**0.038**	**0.12**	3.77	0.06	0.10	**4.58**	**0.04**	**0.12**	0.06	0.82	0.00
APxITxMT	1.05	0.36	0.03	0.85	0.40	0.03	2.49	0.11	0.07	**3.59**	**0.05**	**0.10**	1.32	0.27	0.04	1.16	0.30	0.03
HxITxMT	0.05	0.93	0.00	0.03	0.97	0.00	0.11	0.89	0.00	0.14	0.87	0.00	0.78	0.46	0.02	0.89	0.41	0.03
APxHxITxMT	1.58	0.20	0.05	1.90	0.14	0.06	2.58	0.06	0.07	0.73	0.54	0.02	0.26	0.85	0.01	0.43	0.73	0.01

	1000 to 1500ms			1500 to 2000ms			2000 to 2500ms			2500 to 3000ms			3000 to 3500ms		
	F	*p*	$\eta 2_p$	*F*	*p*	$\eta 2_p$	*F*	*p*	$\eta 2_p$	*F*	*p*	$\eta 2_p$	*F*	*p*	$\eta 2_p$
IT	**10.26**	**0.003**	**0.24**	**5.66**	**0.023**	**0.15**	1.43	0.24	0.04	1.82	0.19	0.05	**4.82**	**0.035**	**0.13**
MT	0.01	0.93	0.00	0.20	0.66	0.01	0.15	0.71	0.00	1.21	0.28	0.04	0.00	0.98	0.00
APxIT	1.80	0.19	0.05	2.51	0.11	0.07	**8.41**	**0.004**	**0.20**	**4.15**	**0.04**	**0.11**	**11.82**	**0.001**	**0.26**
HxIT	1.95	0.15	0.06	1.23	0.30	0.04	2.06	0.14	0.06	0.97	0.38	0.03	2.77	0.07	0.08
APxHxIT	**5.56**	**0.002**	**0.14**	**3.92**	**0.013**	**0.11**	1.72	0.17	0.05	1.22	0.31	0.04	0.46	0.68	0.01
APxMT	0.30	0.63	0.01	0.19	0.70	0.01	1.40	0.25	0.04	0.75	0.43	0.02	1.82	0.19	0.05
HxMT	**3.35**	**0.041**	**0.09**	1.84	0.17	0.05	0.25	0.78	0.01	0.33	0.72	0.01	0.87	0.43	0.03
APxHxMT	1.35	0.26	0.04	0.33	0.86	0.01	0.73	0.57	0.02	1.68	0.16	0.05	0.24	0.88	0.01
ITxMT	3.52	0.07	0.10	1.95	0.17	0.06	2.27	0.14	0.06	**4.07**	**0.048**	**0.11**	3.99	0.05	0.11
APxITxMT	0.81	0.40	0.02	0.05	0.89	0.00	0.20	0.71	0.01	0.12	0.77	0.00	0.48	0.53	0.01
HxITxMT	0.67	0.52	0.02	0.06	0.94	0.00	0.46	0.63	0.01	0.49	0.61	0.02	0.00	1.00	0.00
APxHxITxMT	0.35	0.77	0.01	1.74	0.17	0.05	2.51	0.07	0.07	1.94	0.13	0.06	**4.23**	**0.003**	**0.11**

Note. AP = Anterior/Posterior, H = Hemisphere, IT = Instruction Type, MT = Memory Type, T = Think, NT = No-Think. Significant results are in bold.

Block 2

In the second block of the think/no-think task, there were no significant effects of Right vs. Wrong memory type of ERPs in any time window, but there were significant effects of think vs. no-think instruction type that interacted with the scalp factors. However, in this block there were no interactions between instruction type and memory type except in the very first time-window (100-200ms). See Table A.5. for results of the omnibus ANOVA. Topographical maps that illustrate the raw difference in amplitudes between conditions, significant effects and effect sizes that involve the factor condition, and t-values from significant clusters are illustrated in Figure A.4.

General think vs. no-think effects. Similar to block 1, the ANOVA results indicate a negative think vs. no-think difference in 200-300ms time window across all regions except the mid-posterior scalp that was strongest in the mid-anterior region. A sustained positive effect was found in all time windows from 750-1000ms to 2500-3000ms time windows. All these effects interacted with scalp factors. The positivity was maximal in left-posterior (750-1500ms) and posterior (1500-2000ms) regions and present in central and right hemispherical regions from 1500-3000ms. A negative difference was found in right anterior regions from 500-1000ms. This is likely a polarity-reversed difference reflecting the same generator as the left-parietal positivity found in these time windows. There were no significant effects from 100 to 500ms.

Interaction between moral memory type and instruction type. There was a significant instruction type and memory type interaction in the 100-200ms time window: think ERPs were more negative than no-think ERPs for morally right, but not morally wrong memories. Additionally, no-think ERPs for morally right memories were more negative than no-think ERPs for morally wrong memories in this time-window. This effect was spread across the whole scalp.

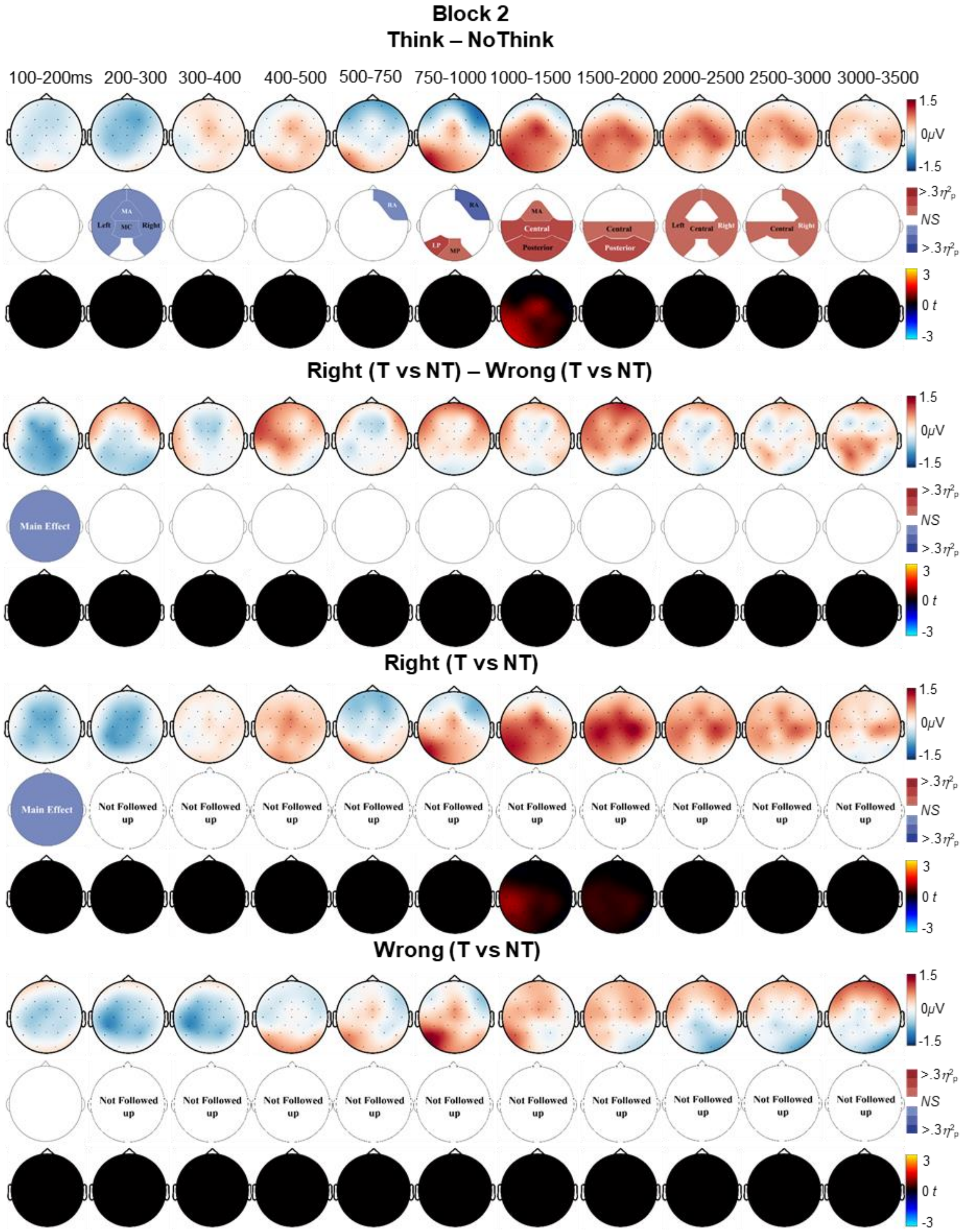


Figure A.4. Results from key main effects and follow up pairwise comparisons in the second half of the TNT task using both omnibus ANOVA and data driven cluster-based permutation testing. Only time windows with significant Instruction Type or Instruction Type x Memory Type interactions were followed up in the ANOVA. For Cluster-based permutation tests, due to the longer time windows, all time windows were used for all comparisons. For each comparison, images on top are topographical maps of ERP amplitude differences between conditions (blue/white/red colourmap, in μV), images in the middle row illustrate statistically significant regions with effect sizes (partial η2), images in the bottom row illustrates t-values (blue/black/red colourmap, in t) for the differences. White font represents greatest effect size. The colour scale represents magnitude and direction of the effect. Abbreviations: LA = Left Anterior, MA = Mid Anterior, RA = Right Anterior, LC = Left Central, MC = Mid Central, RC = Right Central, LP = Left Parietal, MP = Mid Parietal, RP = Right Parietal.

Table A.5. Omnibus ANOVA results from the second half of TNT task.

	100 to 200ms			200 to 300ms			300 to 400ms			400 to 500ms			500 to 750ms			750 to 1000ms		
	F	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$
IT	2.37	0.133	0.07	**5.26**	**0.028**	**0.14**	0.39	0.536	0.01	0.70	0.411	0.02	0.10	0.751	0.00	0.96	0.335	0.03
MT	0.85	0.362	0.03	0.11	0.742	0.00	0.55	0.465	0.02	0.07	0.793	0.00	2.56	0.119	0.07	0.63	0.435	0.02
APxIT	0.71	0.429	0.02	1.91	0.171	0.06	0.27	0.689	0.01	0.25	0.703	0.01	3.35	0.065	0.09	**5.81**	**0.014**	**0.15**
HxIT	0.26	0.719	0.01	0.41	0.625	0.01	2.00	0.144	0.06	0.81	0.425	0.02	0.53	0.593	0.02	**9.08**	**<.001**	**0.22**
APxHxIT	**3.75**	**0.015**	**0.10**	**3.21**	**0.03**	**0.09**	1.29	0.285	0.04	1.80	0.164	0.05	**4.30**	**0.01**	**0.12**	**6.57**	**0.001**	**0.17**
APxMT	1.25	0.283	0.04	0.18	0.777	0.01	2.06	0.155	0.06	0.75	0.428	0.02	1.83	0.182	0.05	0.41	0.579	0.01
HxMT	1.43	0.248	0.04	0.58	0.519	0.02	0.60	0.522	0.02	0.61	0.532	0.02	0.37	0.666	0.01	0.52	0.594	0.02
APxHxMT	1.01	0.379	0.03	1.34	0.268	0.04	0.47	0.683	0.01	0.93	0.421	0.03	0.98	0.397	0.03	1.10	0.346	0.03
ITxMT	**6.65**	**0.015**	**0.17**	0.44	0.51	0.01	0.04	0.847	0.00	1.33	0.257	0.04	0.03	0.875	0.00	0.39	0.535	0.01
APxITxMT	0.59	0.493	0.02	3.15	0.068	0.09	0.31	0.661	0.01	0.79	0.413	0.02	0.01	0.95	0.00	0.74	0.444	0.02
HxITxMT	0.72	0.468	0.02	0.69	0.478	0.02	0.76	0.465	0.02	0.82	0.445	0.02	0.37	0.695	0.01	1.02	0.366	0.03
APxHxITxMT	2.34	0.089	0.07	2.15	0.105	0.06	0.41	0.727	0.01	0.78	0.512	0.02	1.57	0.2	0.05	0.85	0.467	0.03

	1000 to 1500ms			1500 to 2000ms			2000 to 2500ms			2500 to 3000ms			3000 to 3500ms		
	F	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$
IT	**10.91**	**0.002**	**0.25**	**7.40**	**0.01**	**0.18**	**4.85**	**0.035**	**0.13**	3.57	0.068	0.10	0.2	0.691	0.01
MT	0.81	0.375	0.02	0.00	0.973	0	0.25	0.624	0.01	0	0.988	0	0.0	0.909	0
APxIT	3.48	0.066	0.10	2.34	0.114	0.07	1.17	0.302	0.03	1.14	0.308	0.03	1.2	0.289	0.04
HxIT	**8.22**	**0.001**	**0.20**	1.48	0.235	0.04	0.04	0.96	0.00	0.00	1	0	1.2	0.317	0.03
APxHxIT	**5.62**	**0.005**	**0.15**	**6.59**	**<.001**	**0.17**	**3.70**	**0.015**	**0.10**	**3.61**	**0.008**	**0.10**	**2.8**	**0.027**	**0.08**
APxMT	0.38	0.597	0.01	0.13	0.776	0.00	0.23	0.668	0.01	0.47	0.531	0.01	0.4	0.566	0.01
HxMT	1.26	0.287	0.04	0.30	0.71	0.01	1.16	0.319	0.03	0.40	0.657	0.01	1.0	0.356	0.03
APxHxMT	0.88	0.423	0.03	0.22	0.848	0.01	1.79	0.156	0.05	2.39	0.069	0.07	2.3	0.077	0.07
ITxMT	0.01	0.929	0.00	1.15	0.292	0.03	0.03	0.876	0.00	0.06	0.808	0.00	0.5	0.481	0.02
APxITxMT	0.59	0.495	0.02	0.67	0.48	0.02	0.01	0.938	0.00	0.09	0.811	0.00	0.8	0.393	0.03
HxITxMT	1.64	0.202	0.05	0.29	0.751	0.01	1.07	0.349	0.03	0.20	0.819	0.01	0.8	0.458	0.02
APxHxITxMT	0.49	0.657	0.02	1.09	0.359	0.03	0.16	0.92	0.01	0.87	0.469	0.03	1.0	0.414	0.03

Note. AP = Anterior/Posterior, H = Hemisphere, IT = Instruction Type, MT = Memory Type. Significant results are in bold.

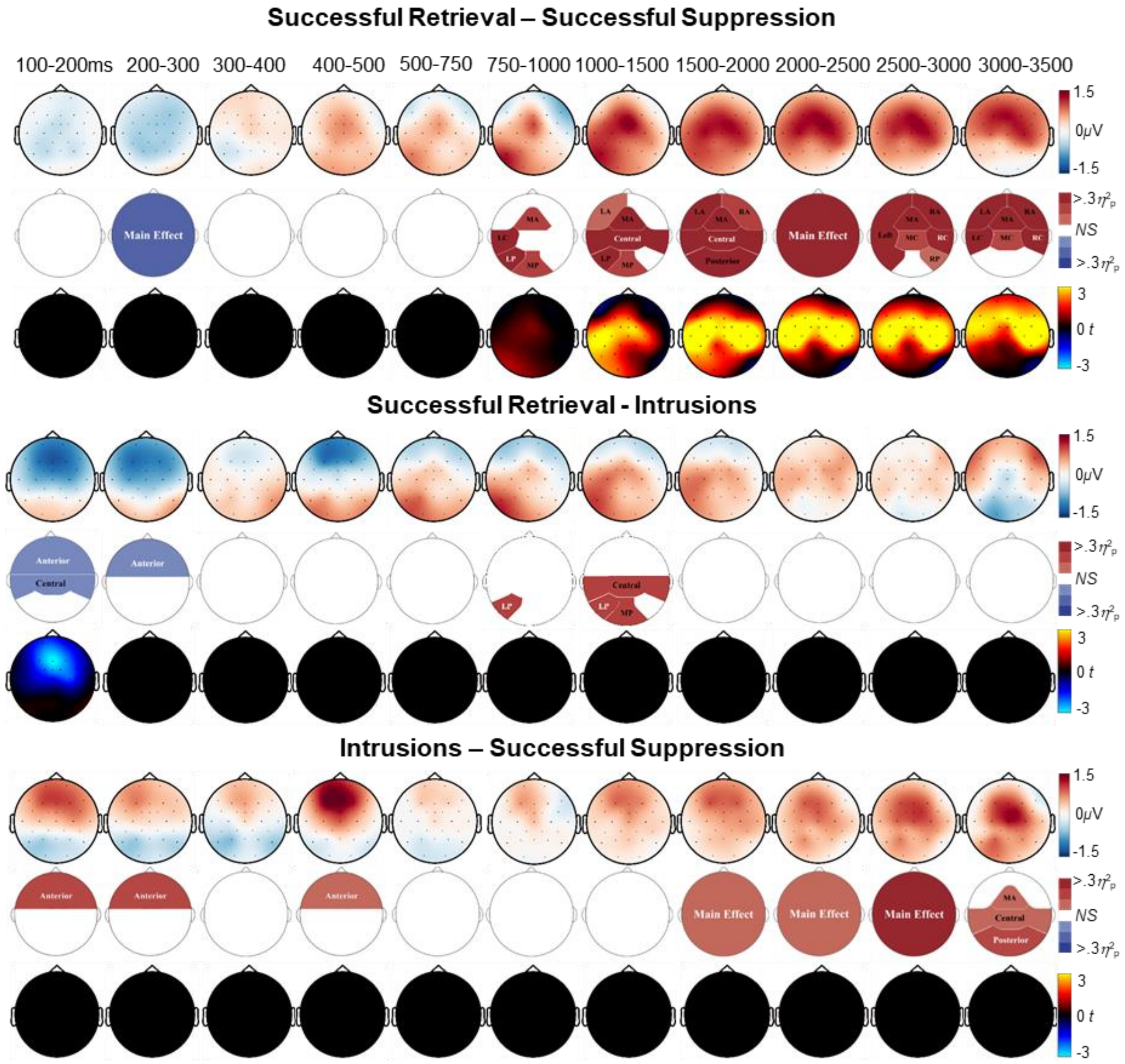
Intrusion analysis

The same scalp factors (Anterior/Posterior and Hemisphere) and time windows were used for ANOVAs of intrusion-related ERP effects as in the primary analysis. So, a 3 (Anterior, Central, Posterior) x 3 (Left, Middle, Right) x 3 (Condition: Successful Suppression, Intrusions, Successful Suppression) repeated-measures omnibus ANOVA was conducted in each time-window. See Table A.6. for results of the omnibus ANOVA. Topographical maps that illustrate the raw difference in amplitudes between conditions, significant effects and effect sizes that involve the factor condition, and t-values from significant clusters are illustrated in Figure A.5.

Successful retrieval vs. successful suppression. Similar to the general think vs. no-think effects, the ANOVAs indicated an early negative difference in the 200-300ms time window, and a sustained positive difference was found in all time windows beginning from 750 to 3500ms. The early negative effect was spread across the whole scalp, but the positive effect interacted with scalp factors across all time windows. The positivity was maximal in the left-posterior region in the 750-1000ms window, in central (1000-2000ms), spread across the whole scalp (2000-2500ms), and right-central from 2500-3500ms. In line with ANOVA results, the cluster analyses revealed significant positive clusters from 950 to 1340ms (p = .001), 1350 to 1580ms (p = .001), and 1750 to 3500ms ($p < .001$). This effect was spread across left posterior regions in the early clusters, and in central and anterior regions across all clusters.

Successful retrieval vs. intrusions. The ANOVAs indicated an early negative effect in 100-200ms and 200-300ms time windows, which was maximal in anterior and also found in central regions (100-200ms). A positive effect was also found from 750-1500ms, which was maximal in left-parietal regions, but also found in central and mid-posterior regions (1000-1500ms). The cluster-based results found one significant negative cluster from 40 to 200ms (p = .01) in anterior regions.

Intrusions vs. successful suppression. The ANOVAs indicated significant positive effects in several time-windows. The positivity was significant in anterior regions from 100-300ms and aligns well with the early negativity found when comparing intrusions with successful retrieval. The positivity was also present only in anterior regions in the 400-500ms. In latter time periods, the positive effect was spread across the whole scalp from 1500-3000ms, and maximal in posterior regions in the final time window (3000-3500ms). There were no significant clusters found in this comparison.



Successful Retrieval – Successful Suppression

Successful Retrieval - Intrusions

Intrusions – Successful Suppression

Figure A.5. Results from pairwise comparisons in the complementary intrusions task using both omnibus ANOVA and data driven cluster-based permutation testing. For each comparison, images on top are topographical maps of ERP amplitude differences between conditions (blue/white/red colourmap, in μV), images in the middle row illustrate statistically significant regions with effect sizes (partial η^2), images in the bottom row illustrates t-values (blue/black/red colourmap, in t) for the differences. White font represents greatest effect size. Abbreviations: LA = Left Anterior, MA = Mid Anterior, RA = Right Anterior, LC = Left Central, MC = Mid Central, RC = Right Central, LP = Left Parietal, MP = Mid Parietal, RP = Right Parietal.

Table A.6. ANOVA results from memory intrusions ERP analysis.

	100 to 200ms			200 to 300ms			300 to 400ms			400 to 500ms			500 to 750ms			750 to 1000ms		
	F	*p*	$\eta 2_p$	*F*	*p*	$\eta 2_p$	*F*	*p*	$\eta 2_p$	*F*	*p*	$\eta 2_p$	*F*	*p*	$\eta 2_p$	*F*	*p*	$\eta 2_p$
C	2.43	0.118	0.09	1.48	0.239	0.06	0.10	0.826	0.00	1.12	0.324	0.04	0.86	0.402	0.03	1.72	0.19	0.06
APxC	**5.16**	**0.018**	**0.17**	**5.12**	**0.013**	**0.17**	2.49	0.094	0.09	**5.37**	**0.008**	**0.18**	1.16	0.318	0.04	1.36	0.267	0.05
HxC	0.33	0.779	0.01	0.16	0.904	0.01	0.72	0.531	0.03	0.66	0.582	0.03	0.93	0.425	0.04	**4.20**	**0.006**	**0.14**
APxHxC	1.54	0.195	0.06	0.88	0.477	0.03	0.56	0.645	0.02	0.70	0.545	0.03	2.03	0.122	0.08	**4.19**	**0.004**	**0.14**

	1000 to 1500ms			1500 to 2000ms			2000 to 2500ms			2500 to 3000ms			3000 to 3500ms		
	F	*p*	$\eta 2_p$	*F*	*p*	$\eta 2_p$	*F*	*p*	$\eta 2_p$	*F*	*p*	$\eta 2_p$	*F*	*p*	$\eta 2_p$
C	**7.245**	**0.002**	**0.225**	**6.791**	**0.002**	**0.214**	**6.196**	**0.005**	**0.199**	**6.14**	**0.006**	**0.197**	**4.634**	**0.014**	**0.156**
APxC	0.724	0.5	0.028	1.005	0.378	0.039	1.307	0.279	0.05	0.723	0.466	0.028	2.603	0.083	0.094
HxC	**2.967**	**0.023**	**0.106**	0.913	0.459	0.035	0.544	0.704	0.021	0.47	0.697	0.018	2.067	0.091	0.076
APxHxC	**4.165**	**0.003**	**0.143**	**2.78**	**0.033**	**0.1**	2.349	0.057	0.086	**2.641**	**0.031**	**0.096**	**3.343**	**0.015**	**0.118**

Note. AP = Anterior/Posterior, H = Hemisphere, C = Condition. Significant results are in bold

Tables demonstrating results of follow-up ERP ANOVAs

Table A.4.1. Follow up ANOVAs of any significant experimental factor interactions (Instruction Type and Memory Type) in Table A.4.

		100 to 200ms			200 to 300ms			300 to 400ms			400 to 500ms			500 to 750ms			750 to 1000ms		
		F	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$
Right	IT	-	-	-	-	-	-	1.38	0.248	0.04	**5.16**	**0.03**	**0.14**	**8.87**	**0.005**	**0.21**	-	-	-
	APxIT	-	-	-	-	-	-	-	-	-	1.42	0.248	0.04	-	-	-	-	-	-
	HxIT	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	APxHxIT	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Wrong	IT	-	-	-	-	-	-	2.23	0.145	0.06	0.01	0.916	0.00	0.28	0.6	0.01	-	-	-
	APxIT	-	-	-	-	-	-	-	-	-	2.27	0.111	0.06	-	-	-	-	-	-
	HxIT	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	APxHxIT	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Think	MT	-	-	-	-	-	-	0.53	0.473	0.02	0.74	0.396	0.02	2.92	0.097	0.08	-	-	-
	APxMT	-	-	-	-	-	-	-	-	-	0.38	0.688	0.01	-	-	-	-	-	-
	HxMT	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	APxHxMT	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
NoThink	MT	-	-	-	-	-	-	**7.07**	**0.012**	**0.18**	3.35	0.076	0.09	1.46	0.236	0.04	-	-	-
	APxMT	-	-	-	-	-	-	-	-	-	**4.12**	**0.034**	**0.11**	-	-	-	-	-	-
	HxMT	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	APxHxMT	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

		1000 to 1500ms			1500 to 2000ms			2000 to 2500ms			2500 to 3000ms			3000 to 3500ms		
		F	*p*	$\eta 2_p$	*F*	*p*	$\eta 2_p$	*F*	*p*	$\eta 2_p$	*F*	*p*	$\eta 2_p$	*F*	*p*	$\eta 2_p$
Right	IT	-	-	-	-	-	-	-	-	-	**4.83**	**0.035**	**0.13**	**8.76**	**0.006**	**0.21**
	APxIT	-	-	-	-	-	-	-	-	-	-	-	-	**8.12**	**0.005**	**0.20**
	HxIT	-	-	-	-	-	-	-	-	-	-	-	-	2.11	0.13	0.06
	APxHxIT	-	-	-	-	-	-	-	-	-	-	-	-	1.39	0.252	0.04
Wrong	IT	-	-	-	-	-	-	-	-	-	0.07	0.799	0.00	0.02	0.883	0.00
	APxIT	-	-	-	-	-	-	-	-	-	-	-	-	**3.92**	**0.046**	**0.11**
	HxIT	-	-	-	-	-	-	-	-	-	-	-	-	1.34	0.268	0.04
	APxHxIT	-	-	-	-	-	-	-	-	-	-	-	-	**3.50**	**0.024**	**0.10**
Think	MT	-	-	-	-	-	-	-	-	-	1.12	0.299	0.03	2.55	0.118	0.07
	APxMT	-	-	-	-	-	-	-	-	-	-	-	-	2.04	0.158	0.06
	HxMT	-	-	-	-	-	-	-	-	-	-	-	-	0.41	0.665	0.01
	APxHxMT	-	-	-	-	-	-	-	-	-	-	-	-	1.75	0.148	0.05
NoThink	MT	-	-	-	-	-	-	-	-	-	**4.79**	**0.036**	**0.13**	3.41	0.074	0.09
	APxMT	-	-	-	-	-	-	-	-	-	-	-	-	0.59	0.47	0.02
	HxMT	-	-	-	-	-	-	-	-	-	-	-	-	0.34	0.713	0.01
	APxHxMT	-	-	-	-	-	-	-	-	-	-	-	-	2.49	0.075	0.07

Note. AP = Anterior/Posterior, H = Hemisphere, IT = Instruction Type, MT = Memory Type, T = Think, NT = No-Think. Significant results are in bold.

Table A.4.2. Follow up ANOVAs of any significant effects of Instruction Type for morally right memories, as reported in Table A.4.1

		100 to 200ms			200 to 300ms			300 to 400ms			400 to 500ms			500 to 750ms			750 to 1000ms		
		F	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$
Morally Right	Anterior (T vs. NT)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	Central (T vs. NT)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	Posterior (T vs. NT)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

		1000 to 1500ms			1500 to 2000ms			2000 to 2500ms			2500 to 3000ms			3000 to 3500ms		
		F	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$
Morally Right	Anterior (T vs. NT)	-	-	-	-	-	-	-	-	-	-	-	-	**18.45**	**0.001**	**0.27**
	Central (T vs. NT)	-	-	-	-	-	-	-	-	-	-	-	-	**8.89**	**0.005**	**0.21**
	Posterior (T vs. NT)	-	-	-	-	-	-	-	-	-	-	-	-	0.99	0.328	0.03

Note. AP = Anterior/Posterior, H = Hemisphere, IT = Instruction Type, MT = Memory Type, MR = Morally Right, MW = Morally Wrong. Significant results are in bold.

Table A.4.3. Follow up ANOVAs of any significant simple main effects of Instruction Type for morally wrong memories, as reported in Table A.4.1.

		100 to 200ms			200 to 300ms			300 to 400ms			400 to 500ms			500 to 750ms			750 to 1000ms		
		F	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$
Morally Wrong	IT in Anterior	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	HxIT in Anterior	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	IT in Central	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	HxIT in Central	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	IT in Posterior	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	HxIT in Posterior	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

		1000 to 1500ms			1500 to 2000ms			2000 to 2500ms			2500 to 3000ms			3000 to 3500ms		
		F	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$
Morally Wrong	IT in Anterior	-	-	-	-	-	-	-	-	-	-	-	-	**4.27**	**0.047**	**0.12**
	HxIT in Anterior	-	-	-	-	-	-	-	-	-	-	-	-	0.01	0.957	0.00
	IT in Central	-	-	-	-	-	-	-	-	-	-	-	-	0.07	0.783	0.00
	HxIT in Central	-	-	-	-	-	-	-	-	-	-	-	-	0.22	0.788	0.01
	IT in Posterior	-	-	-	-	-	-	-	-	-	-	-	-	0.49	0.489	0.02
	HxIT in Posterior	-	-	-	-	-	-	-	-	-	-	-	-	1.41	0.215	0.04

		100 to 200ms			200 to 300ms			300 to 400ms			400 to 500ms			500 to 750ms			750 to 1000ms		
		F	*p*	$\eta 2_p$	*F*	*p*	$\eta 2_p$	*F*	*p*	$\eta 2_p$	*F*	*p*	$\eta 2_p$	*F*	*p*	$\eta 2_p$	*F*	*p*	$\eta 2_p$
Morally Wrong	IT in Left	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	APxIT in Left	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	IT in Middle	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	APxIT in Middle	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	IT in Right	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	APxIT in Right	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

		1000 to 1500ms			1500 to 2000ms			2000 to 2500ms			2500 to 3000ms			3000 to 3500ms		
		F	*p*	$\eta 2_p$	*F*	*p*	$\eta 2_p$	*F*	*p*	$\eta 2_p$	*F*	*p*	$\eta 2_p$	*F*	*p*	$\eta 2_p$
Morally Wrong	IT in Left	-	-	-	-	-	-	-	-	-	-	-	-	0.57	0.455	0.02
	APxIT in Left	-	-	-	-	-	-	-	-	-	-	-	-	2.24	0.124	0.06
	IT in Middle	-	-	-	-	-	-	-	-	-	-	-	-	0.07	0.794	0.00
	APxIT in Middle	-	-	-	-	-	-	-	-	-	-	-	-	0.22	0.713	0.01
	IT in Right	-	-	-	-	-	-	-	-	-	-	-	-	0.72	0.401	0.02
	APxIT in Right	-	-	-	-	-	-	-	-	-	-	-	-	**4.79**	**0.017**	**0.13**

Note. AP = Anterior/Posterior, H = Hemisphere, IT = Instruction Type, T = Think, NT = No-Think, MW = Morally Wrong. Significant results are in bold.

Table A.4.4. Follow up of significant effect of Instruction Type for morally wrong memories interacting with scalp factors (from Table A.4.3).

		100 to 200ms			200 to 300ms			300 to 400ms			400 to 500ms			500 to 750ms			750 to 1000ms		
		F	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$
Wrong	LA (T vs. NT)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	MA (T vs. NT)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	RA (T vs. NT)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	LC (T vs. NT)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	MC (T vs. NT)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	RC (T vs. NT)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	LP (T vs. NT)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	MP (T vs. NT)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	RP (T vs. NT)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

		1000 to 1500ms			1500 to 2000ms			2000 to 2500ms			2500 to 3000ms			3000 to 3500ms		
		F	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$
Wrong	LA (T vs. NT)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	MA (T vs. NT)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	RA (T vs. NT)	-	-	-	-	-	-	-	-	-	-	-	-	**4.29**	**0.046**	**0.12**
	LC (T vs. NT)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	MC (T vs. NT)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	RC (T vs. NT)	-	-	-	-	-	-	-	-	-	-	-	-	0.50	0.481	0.02
	LP (T vs. NT)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	MP (T vs. NT)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	RP (T vs. NT)	-	-	-	-	-	-	-	-	-	-	-	-	2.04	0.163	0.06

Note. AP = Anterior/Posterior, H = Hemisphere, IT = Instruction Type, T = Think, NT = No-Think, MW = Morally Wrong, LA = Left Anterior, MA = Mid Anterior, RA = Right Anterior, LC = Left Central, MC = Mid Central, RC = Right Central, LP = Left Parietal, MP = Mid Parietal, RP = Right Parietal. Significant results are in bold.

Table A.4.5. Follow up of significant effect of Memory Type in each level of Instruction Type, interacting with scalp factors (from Table A.4.3)

		100 to 200ms			200 to 300ms			300 to 400ms			400 to 500ms			500 to 750ms			750 to 1000ms		
		F	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$
NoThink APxMT	MT in Anterior	-	-	-	-	-	-	-	-	-	**6.43**	**0.016**	**0.16**	-	-	-	-	-	-
	MT in Central	-	-	-	-	-	-	-	-	-	3.85	0.058	0.10	-	-	-	-	-	-
	MT in Posterior	-	-	-	-	-	-	-	-	-	0.18	0.679	0.01	-	-	-	-	-	-
NoThink HxMT	MT in Left	-	-	-	-	-	-	2.73	0.108	0.08	-	-	-	-	-	-	-	-	-
	MT in Middle	-	-	-	-	-	-	1.50	0.230	0.04	-	-	-	-	-	-	-	-	-
	MT in Right	-	-	-	-	-	-	0	0.995	0	-	-	-	-	-	-	-	-	-

		1000 to 1500ms			1500 to 2000ms			2000 to 2500ms			2500 to 3000ms			3000 to 3500ms		
		F	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$
NoThink APxMT	MT in Anterior	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	MT in Central	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	MT in Posterior	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
NoThink HxMT	MT in Left	0.50	0.485	0.01	-	-	-	-	-	-	-	-	-	-	-	-
	MT in Middle	0.07	0.789	0.00	-	-	-	-	-	-	-	-	-	-	-	-
	MT in Right	0.67	0.420	0.02	-	-	-	-	-	-	-	-	-	-	-	-

Note. AP = Anterior/Posterior, H = Hemisphere, T = Think, NT = No-Think, MT = Memory Type, MW = Morally Wrong, MR = Morally Right. Significant results are in bold. 6.7. Follow up of APxHxIT main effects from omnibus effects.

Table A.4.6. Follow up ANOVAs of significant Instruction Type effects in omnibus ANOVA from Table A.4.

	100 to 200ms			200 to 300ms			300 to 400ms			400 to 500ms			500 to 750ms			750 to 1000ms		
	F	p	$\eta2_p$	F	p	$\eta2_p$	F	p	$\eta2_p$	F	p	$\eta2_p$	F	p	$\eta2_p$	F	p	$\eta2_p$
IT in Anterior	**4.23**	**0.048**	**0.11**	**4.60**	**0.039**	**0.12**	-	-	-	-	-	-	**4.32**	**0.046**	**0.12**	0.70	0.408	0.02
HxIT in Anterior	2.90	0.062	0.08	2.87	0.068	0.08	-	-	-	-	-	-	2.61	0.081	0.07	3.01	0.058	0.08
IT in Central	2.65	0.113	0.07	**4.24**	**0.047**	**0.11**	-	-	-	-	-	-	**4.52**	**0.041**	**0.12**	1.50	0.229	0.04
HxIT in Central	0.31	0.698	0.01	1.31	0.278	0.04	-	-	-	-	-	-	1.20	0.307	0.04	1.71	0.189	0.05
IT in Posterior	0.07	0.797	0.00	0.78	0.383	0.02	-	-	-	-	-	-	**6.07**	**0.019**	**0.16**	**6.89**	**0.013**	**0.17**
HxIT in Posterior	1.43	0.247	0.04	1.64	0.209	0.05	-	-	-	-	-	-	3.69	0.035	0.10	**10.92**	**<.001**	**0.25**
IT in Left	1.71	0.2	0.05	3.75	0.061	0.10	-	-	-	-	-	-	**8.06**	**0.008**	**0.20**	**6.53**	**0.015**	**0.17**
APxIT in Left	0.63	0.483	0.02	0.90	0.413	0.03	-	-	-	-	-	-	1.38	0.26	0.04	**6.64**	**0.008**	**0.17**
IT in Middle	1.81	0.187	0.05	3.67	0.064	0.10	-	-	-	-	-	-	**5.60**	**0.024**	**0.15**	2.36	0.134	0.07
APxIT in Middle	7.07	0.006	0.18	5.69	0.01	0.15	-	-	-	-	-	-	0.61	0.548	0.02	0.47	0.625	0.01
IT in Right	1.74	0.196	0.05	1.85	0.183	0.05	-	-	-	-	-	-	3.13	0.086	0.09	0.49	0.488	0.02
APxIT in Right	0.96	0.387	0.03	1.37	0.255	0.04	-	-	-	-	-	-	0.07	0.937	0.00	0.42	0.656	0.01

	1000 to 1500ms			1500 to 2000ms			2000 to 2500ms			2500 to 3000ms			3000 to 3500ms		
	F	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$
IT in Anterior	**15.856**	**<.001**	**0.325**	**11.12**	**0.002**	**0.25**	**7.33**	**0.011**	**0.18**	**5.20**	**0.029**	**0.14**	**10.97**	**0.002**	**0.3**
HxIT in Anterior	1.45	0.241	0.04	0.24	0.79	0.01	-	-	-	-	-	-	-	-	-
IT in Central	**7.09**	**0.012**	**0.18**	**4.47**	**0.042**	**0.12**	1.03	0.318	0.03	1.55	0.22	0.05	**4.14**	**0.05**	**0.1**
HxIT in Central	1.57	0.216	0.05	2.21	0.117	0.06	-	-	-	-	-	-	-	-	-
IT in Posterior	3.90	0.057	0.11	1.08	0.307	0.03	0.24	0.631	0.01	0.02	0.89	0.00	0.01	0.907	0.0
HxIT in Posterior	**7.98**	**0.002**	**0.20**	**4.90**	**0.013**	**0.13**	-	-	-	-	-	-	-	-	-
IT in Left	**13.95**	**0.001**	**0.30**	**8.99**	**0.005**	**0.21**	-	-	-	-	-	-	-	-	-
APxIT in Left	0.08	0.928	0.00	0.28	0.754	0.01	-	-	-	-	-	-	-	-	-
IT in Middle	**6.94**	**0.013**	**0.17**	2.85	0.101	0.08	-	-	-	-	-	-	-	-	-
APxIT in Middle	**6.07**	**0.004**	**0.16**	3.93	0.024	0.11	-	-	-	-	-	-	-	-	-
IT in Right	**6.56**	**0.015**	**0.17**	**4.75**	**0.037**	**0.13**	-	-	-	-	-	-	-	-	-
APxIT in Right	2.77	0.094	0.08	4.30	0.037	0.12	-	-	-	-	-	-	-	-	-

Note. AP = Anterior/Posterior, H = Hemisphere, MT = Memory Type, MW = Morally Wrong, MR = Morally Right. Significant results are in bold.

Table A.4.7. Follow up ANOVAs of significant interactions between Instruction Type and scalp factors (from Table A.4.6)

	100 to 200ms			200 to 300ms			300 to 400ms			400 to 500ms			500 to 750ms			750 to 1000ms		
	F	*p*	$\eta 2_p$	*F*	*p*	$\eta 2_p$	*F*	*p*	$\eta 2_p$	*F*	*p*	$\eta 2_p$	*F*	*p*	$\eta 2_p$	*F*	*p*	$\eta 2_p$
LA (T vs. NT)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.23	0.632	0.01
MA (T vs. NT)	**5.35**	**0.027**	**0.14**	**5.74**	**0.022**	**0.15**	-	-	-	-	-	-	-	-	-	-	-	-
RA (T vs. NT)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
LC (T vs. NT)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	**4.48**	**0.042**	**0.12**
MC (T vs. NT)	2.48	0.125	0.07	**4.27**	**0.047**	**0.12**	-	-	-	-	-	-	-	-	-	-	-	-
RC (T vs. NT)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
LP (T vs. NT)	-	-	-	-	-	-	-	-	-	-	-	-						
MP (T vs. NT)	0.22	0.645	0.01	0.45	0.508	0.01	-	-	-	-	-	-	**5.14**	**0.03**	**0.14**	**4.43**	**0.043**	**0.12**
RP (T vs. NT)	-	-	-	-	-	-	-	-	-	-	-	-	2.57	0.118	0.07	1.10	0.302	0.03

	1000 to 1500ms			1500 to 2000ms			2000 to 2500ms			2500 to 3000ms			3000 to 3500ms		
	F	*p*	$\eta 2_p$	*F*	*p*	$\eta 2_p$	*F*	*p*	$\eta 2_p$	*F*	*p*	$\eta 2_p$	*F*	*p*	$\eta 2_p$
LA (T vs. NT)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
MA (T vs. NT)				**7.77**	**0.009**	**0.19**	-	-	-	-	-	-	-	-	-
RA (T vs. NT)	-	-	-	**10.44**	**0.003**	**0.24**	-	-	-	-	-	-	-	-	-
LC (T vs. NT)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
MC (T vs. NT)	3.06	0.09	0.09	1.32	0.259	0.04	-	-	-	-	-	-	-	-	-
RC (T vs. NT)	-	-	-	**5.12**	**0.03**	**0.13**	-	-	-	-	-	-	-	-	-
LP (T vs. NT)	**8.70**	**0.006**	**0.21**	-	-	-	-	-	-	-	-	-	-	-	-
MP (T vs. NT)	0.84	0.365	0.03	0.35	0.559	0.01	-	-	-	-	-	-	-	-	-
RP (T vs. NT)	0.84	0.365	0.03	0.21	0.654	0.01	-	-	-	-	-	-	-	-	-

Note. AP = Anterior/Posterior, H = Hemisphere, IT = Instruction Type, T = Think, NT = No-Think, LA = Left Anterior, MA = Mid Anterior, RA = Right Anterior, LC = Left Central, MC = Mid Central, RC = Right Central, LP = Left Parietal, MP = Mid Parietal, RP = Right Parietal. Significant results are in bold.

Table A.5.1. Follow up ANOVAs of any significant experimental factor interactions (Instruction Type and Memory Type) in Table A.5

		100 to 200ms			200 to 300ms			300 to 400ms			400 to 500ms			500 to 750ms			750 to 1000ms			
		F	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$	
Morally Right	IT	**6.62**	**0.015**	**0.17**	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	APxIT	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	HxIT	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	APxHxIT	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Morally Wrong	IT	0.10	0.759	0.00	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	APxIT	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	HxIT	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	APxHxIT	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Think	MT	**4.39**	**0.044**	**0.12**	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	APxMT	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	HxMT	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	APxHxMT	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
NoThink	MT	0.27	0.607	0.01	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	APxMT	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	HxMT	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	APxHxMT	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

		1000 to 1500ms			1500 to 2000ms			2000 to 2500ms			2500 to 3000ms			3000 to 3500ms		
		F	*p*	$\eta 2_p$	*F*	*p*	$\eta 2_p$	*F*	*p*	$\eta 2_p$	*F*	*p*	$\eta 2_p$	*F*	*p*	$\eta 2_p$
Right	IT	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	APxIT	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	HxIT	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	APxHxIT	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Wrong	IT	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	APxIT	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	HxIT	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	APxHxIT	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Think	MT	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	APxMT	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	HxMT	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	APxHxMT	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
NoThink	MT	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	APxMT	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	HxMT	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	APxHxMT	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Note. AP = Anterior/Posterior, H = Hemisphere, IT = Instruction Type, MT = Memory Type, T = Think, NT = No-Think, MW = Morally Wrong, MR = Morally Right. Significant results are in bold.

Table A.5.2. Follow up ANOVAs of significant Instruction Type effects in omnibus ANOVA from Table A.5

	100 to 200ms			200 to 300ms			300 to 400ms			400 to 500ms			500 to 750ms			750 to 1000ms		
	F	*p*	$\eta 2_p$	*F*	*p*	$\eta 2_p$	*F*	*p*	$\eta 2_p$	*F*	*p*	$\eta 2_p$	*F*	*p*	$\eta 2_p$	*F*	*p*	$\eta 2_p$
IT in Anterior	2.97	0.094	0.08	**6.48**	**0.016**	**0.16**	-	-	-	-	-	-	1.88	0.18	0.05	0.86	0.36	0.03
HxIT in Anterior	0.35	0.701	0.01	0.71	0.486	0.02	-	-	-	-	-	-	**4.32**	**0.018**	**0.12**	**74.99**	**<.001**	**0.31**
IT in Central	1.86	0.182	0.05	**4.67**	**0.038**	**0.12**	-	-	-	-	-	-	0.12	0.731	0.00	0.86	0.362	0.03
HxIT in Central	0.76	0.454	0.02	1.33	0.271	0.04	-	-	-	-	-	-	0.18	0.838	0.01	2.51	0.09	0.07
IT in Posterior	0.68	0.416	0.02	1.66	0.207	0.05	-	-	-	-	-	-	0.88	0.353	0.03	**5.33**	**0.027**	**0.14**
HxIT in Posterior	2.68	0.091	0.08	1.63	0.209	0.05	-	-	-	-	-	-	1.60	0.214	0.05	5.97	0.006	0.15
IT in Left	3.13	0.086	0.09	**5.37**	**0.027**	**0.14**	-	-	-	-	-	-	0.07	0.789	0.00	2.40	0.131	0.07
APxIT in Left	0.15	0.74	0.01	0.43	0.562	0.01	-	-	-	-	-	-	**7.33**	**0.006**	**0.18**	**9.63**	**0.002**	**0.23**
IT in Middle	1.21	0.279	0.04	3.88	0.057	0.11	-	-	-	-	-	-	0.00	0.981	0.00	3.55	0.068	0.10
APxIT in Middle	3.47	0.056	0.10	**3.94**	**0.043**	**0.11**	-	-	-	-	-	-	1.01	0.348	0.03	0.77	0.438	0.02
IT in Right	1.72	0.199	0.05	**4.56**	**0.04**	**0.12**	-	-	-	-	-	-	0.43	0.517	0.01	0.45	0.509	0.01
APxIT in Right	0.42	0.552	0.01	2.37	0.122	0.07	-	-	-	-	-	-	3.54	0.06	0.10	**7.97**	**0.005**	**0.20**

	1000 to 1500ms			1500 to 2000ms			2000 to 2500ms			2500 to 3000ms			3000 to 3500ms		
	F	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$
IT in Anterior	1.48	0.233	0.04	1.57	0.219	0.05	2.91	0.098	0.08	2.53	0.121	0.07	0.51	0.478	0.02
HxIT in Anterior	**8.96**	**0.001**	**0.21**	**3.95**	**0.038**	**0.11**	2.03	0.149	0.06	1.64	0.205	0.05	0.21	0.756	0.01
IT in Central	**8.57**	**0.006**	**0.21**	**7.72**	**0.009**	**0.19**	**6.10**	**0.019**	**0.16**	**4.46**	**0.042**	**0.12**	0.01	0.925	0.00
HxIT in Central	0.90	0.404	0.03	0.00	0.96	0.00	0.80	0.453	0.02	0.96	0.38	0.03	1.99	0.148	0.06
IT in Posterior	**8.33**	**0.007**	**0.20**	**8.60**	**0.006**	**0.21**	2.86	0.1	0.08	1.50	0.229	0.04	0.31	0.581	0.01
HxIT in Posterior	**3.96**	**0.031**	**0.11**	1.04	0.358	0.03	0.45	0.618	0.01	0.74	0.474	0.02	0.61	0.516	0.02
IT in Left	**8.20**	**0.007**	**0.20**	6.79	0.014	0.17	**4.84**	**0.035**	**0.13**	3.66	0.064	0.10	0.11	0.74	0.00
APxIT in Left	4.13	0.042	0.11	3.38	0.062	0.09	0.92	0.367	0.03	0.04	0.608	0.01	1.54	0.225	0.05
IT in Middle	**8.29**	**0.007**	**0.20**	**5.61**	**0.02**	**0.15**	3.27	0.08	0.09	2.12	0.155	0.06	0.00	0.965	0.00
APxIT in Middle	0.14	0.756	0.00	0.14	0.783	0.00	1.59	0.217	0.05	1.41	0.251	0.04	0.82	0.416	0.02
IT in Right	3.20	0.083	0.09	5.64	0.024	0.15	**5.43**	**0.026**	**0.14**	**4.33**	**0.045**	**0.12**	1.17	0.287	0.03
APxIT in Right	2.89	0.091	0.08	3.23	0.072	0.09	2.38	0.122	0.07	3.15	0.068	0.09	3.31	0.059	0.09

Note. AP = Anterior/Posterior, H = Hemisphere, IT = Instruction Type. Significant results are in bold.

Table A.5.3. Follow up ANOVAs of significant interactions between Instruction Type and scalp factors (from Table A.5.2).

	100 to 200ms			200 to 300ms			300 to 400ms			400 to 500ms			500 to 750ms			750 to 1000ms		
	F	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$
LA (T vs. NT)	-	-	-	-	-	-	-	-	-	-	-	-	3.64	0.065	0.10	1.37	0.25	0.04
MA (T vs. NT)	-	-	-	**5.09**	**0.031**	**0.13**	-	-	-	-	-	-	0.07	0.801	0.00	1.46	0.236	0.04
RA (T vs. NT)	-	-	-	-	-	-	-	-	-	-	-	-	**4.34**	**0.045**	**0.12**	**7.78**	**0.009**	**0.19**
LC (T vs. NT)	-	-	-	-	-	-	-	-	-	-	-	-	0.08	0.784	0.00	2.04	0.163	0.06
MC (T vs. NT)	-	-	-	**4.13**	**0.05**	**0.11**	-	-	-	-	-	-	-	-	-	-	-	-
RC (T vs. NT)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.00	0.964	0.00
LP (T vs. NT)	-	-	-	-	-	-	-	-	-	-	-	-	0.02	0.877	0.00	**10.73**	**0.002**	**0.25**
MP (T vs. NT)	-	-	-	0.27	0.609	0.01	-	-	-	-	-	-	-	-	-	**5.40**	**0.026**	**0.14**
RP (T vs. NT)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.93	0.343	0.03

	1000 to 1500ms			1500 to 2000ms			2000 to 2500ms			2500 to 3000ms			3000 to 3500ms		
	F	*p*	$\eta 2_p$	*F*	*p*	$\eta 2_p$	*F*	*p*	$\eta 2_p$	*F*	*p*	$\eta 2_p$	*F*	*p*	$\eta 2_p$
LA (T vs. NT)	0.64	0.431	0.02	0.93	0.343	0.03	-	-	-	-	-	-	-	-	-
MA (T vs. NT)	**5.41**	**0.026**	**0.14**	3.48	0.071	0.10	-	-	-	-	-	-	-	-	-
RA (T vs. NT)	0.00	0.968	0.00	0.20	0.66	0.01	-	-	-	-	-	-	-	-	-
LC (T vs. NT)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
MC (T vs. NT)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
RC (T vs. NT)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
LP (T vs. NT)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
MP (T vs. NT)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
RP (T vs. NT)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Note. AP = Anterior/Posterior, H = Hemisphere, IT = Instruction Type, T = Think, NT = No-Think, LA = Left Anterior, MA = Mid Anterior, RA = Right Anterior, LC = Left Central, MC = Mid Central, RC = Right Central, LP = Left Parietal, MP = Mid Parietal, RP = Right Parietal. Significant results are in bold.

Table A.6.1. Follow up ANOVAs of any significant interactions between condition and scalp factors from Table A.6

	100 to 200ms			200 to 300ms			300 to 400ms			400 to 500ms			500 to 750ms			750 to 1000ms		
	F	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$
SR_SS	2.99	0.096	0.11	**9.40**	**0.005**	**0.27**	-	-	-	2.77	0.08	0.10	-	-	-	**5.50**	**0.027**	**0.18**
APxSR_SS	0.13	0.762	0.01	0.78	0.402	0.03	-	-	-	0.10	0.828	0.00	-	-	-	2.22	0.132	0.08
HxSR_SS	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	**8.78**	**0.001**	**0.26**
APxHxSR_SS	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	**7.13**	**0.002**	**0.22**
I_SR	4.07	0.055	0.14	1.69	0.205	0.06	-	-	-	0.01	0.914	0.00	-	-	-	1.19	0.285	0.05
APxI_SR	**6.88**	**0.012**	**0.22**	**8.96**	**0.002**	**0.26**	-	-	-	**6.15**	**0.01**	**0.20**	-	-	-	1.95	0.169	0.07
HxI_SR	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	**3.70**	**0.041**	**0.13**
APxHxI_SR	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	**4.59**	**0.006**	**0.16**
I_SS	0.90	0.351	0.04	0.02	0.881	0.00	-	-	-	1.59	0.219	0.07	-	-	-	0.32	0.578	0.01
APxI_SS	**5.03**	**0.031**	**0.17**	**3.97**	**0.05**	**0.14**	-	-	-	**6.11**	**0.011**	**0.20**	-	-	-	0.22	0.683	0.01
HxI_SS	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.65	0.52	0.03
APxHxI_SS	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1.68	0.19	0.06

	1000 to 1500ms			1500 to 2000ms			2000 to 2500ms			2500 to 3000ms			3000 to 3500ms		
	F	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$
SR_SS	**16.03**	**<.001**	**0.39**	**22.94**	**<.001**	**0.48**	**16.40**	**<.001**	**0.40**	**16.83**	**<.001**	**0.40**	**14.17**	**0.001**	**0.36**
APxSR_SS	0.21	0.684	0.01	0.83	0.392	0.03	-	-	-	3.54	0.062	0.12	**8.72**	**0.003**	**0.26**
HxSR_SS	**5.16**	**0.01**	**0.17**	1.09	0.343	0.04	-	-	-	0.19	0.791	0.01	0.28	0.735	0.01
APxHxSR_SS	**7.63**	**0.001**	**0.23**	**5.61**	**0.003**	**0.18**	-	-	-	**5.56**	**0.002**	**0.18**	**4.42**	**0.004**	**0.15**
I_SR	**4.25**	**0.05**	**0.15**	1.89	0.182	0.07	1.51	0.23	0.06	0.47	0.498	0.02	0.00	0.914	0.00
APxI_SR	1.08	0.319	0.04	0.89	0.375	0.03	-	-	-	0.18	0.721	0.01	2.44	0.118	0.09
HxI_SR	3.03	0.062		1.23	0.301	0.05	-	-	-	0.23	0.79	0.01	2.96	0.062	0.11
APxHxI_SR	**4.89**	**0.006**	**0.16**	2.83	0.065	0.10	-	-	-	0.32	0.77	0.01	0.87	0.432	0.03
I_SS	2.69	0.114	0.10	**4.36**	**0.047**	**0.15**	**4.41**	**0.046**	**0.15**	**5.71**	**0.025**	**0.19**	**5.68**	**0.025**	**0.19**
APxI_SS	0.77	0.399	0.03	1.35	0.263	0.05	-	-	-	0.45	0.542	0.02	0.10	0.813	0.00
HxI_SS	0.27	0.751	0.01	0.24	0.777	0.01	-	-	-	1.30	0.282	0.05	2.26	0.124	0.08
APxHxI_SS	0.51	0.635	0.02	0.50	0.671	0.02	-	-	-	2.66	0.063	0.10	**5.26**	**0.005**	**0.17**

Note. AP = Anterior/Posterior, H = Hemisphere, C = Condition, SR = Successful Retrieval, I = Intrusions, SS = Successful Suppression. Significant results are in bold.

Table A.6.2. Follow up ANOVAs of significant Successful Retrieval vs. Successful Suppression effects in omnibus ANOVA from Table A.6

	100 to 200ms			200 to 300ms			300 to 400ms			400 to 500ms			500 to 750ms			750 to 1000ms		
	F	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$
SR vs. SS in Anterior	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.75	0.394	0.03
HxC in Anterior	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	**17.18**	**<.001**	**0.41**
SR vs. SS in Central	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	4.00	0.057	0.14
HxC in Central	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	2.74	0.086	0.10
SR vs. SS in Posterior	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	**3.91**	**0.026**	**0.14**
HxC in Posterior	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	**10.80**	**0.001**	**0.30**
SR vs. SS in Left	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	**13.42**	**0.001**	**0.35**
APxC in Left	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	**6.60**	**0.011**	**0.21**
SR vs. SS in Middle	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	**9.65**	**0.005**	**0.28**
APxC in Middle	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.22	0.801	0.01
SR vs. SS in Right	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.03	0.864	0.00
APxC in Right	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	**4.09**	**0.038**	**0.14**

	1000 to 1500ms			1500 to 2000ms			2000 to 2500ms			2500 to 3000ms			3000 to 3500ms		
	F	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$
SR vs. SS in Anterior	**6.87**	**0.015**	**0.22**	**12.99**	**0.001**	**0.34**	**-**	**-**	**-**	**14.34**	**0.001**	**0.37**	**14.91**	**0.001**	**0.37**
HxC in Anterior	**10.55**	**0.001**	**0.30**	**4.94**	**0.018**	**0.17**	**-**	**-**	**-**	**4.04**	**0.035**	**0.14**	0.89	0.403	0.03
SR vs. SS in Central	**15.49**	**0.001**	**0.38**	**23.26**	**<.001**	**0.48**	**-**	**-**	**-**	**19.46**	**<.001**	**0.44**	**15.39**	**0.001**	**0.38**
HxC in Central	1.55	0.224	0.06	0.32	0.731	0.01	-	-	-	0.28	0.703	0.01	2.14	0.14	0.08
SR vs. SS in Posterior	**9.02**	**0.006**	**0.27**	**11.47**	**0.002**	**0.31**	-	-	-	6.26	0.019	0.20	1.90	0.18	0.07
HxC in Posterior	**5.04**	**0.017**	**0.17**	2.36	0.113	0.09	-	-	-	1.76	0.183	0.07	1.61	0.21	0.06
SR vs. SS in Left	**17.75**	**<001**	**0.42**	**21.98**	**<.001**	**0.47**	-	-	-	**17.89**	**<.001**	**0.42**	**12.17**	**0.002**	**0.33**
APxC in Left	0.93	0.353	0.04	0.96	0.353	0.04	-	-	-	0.97	0.35	0.04	**5.19**	**0.023**	**0.17**
SR vs. SS in Middle	**15.19**	**0.001**	**0.38**	**15.75**	**0.001**	**0.39**	-	-	-	**10.25**	**0.004**	**0.29**	**7.34**	**0.012**	**0.23**
APxC in Middle	1.58	0.223	0.06	1.99	0.165	0.07	-	-	-	**6.12**	**0.01**	**0.20**	**9.18**	**0.001**	**0.27**
SR vs. SS in Right	7.85	0.01	0.24	**23.85**	**<.001**	**0.49**	-	-	-	**18.51**	**<.001**	**0.43**	**17.69**	**<.001**	**0.41**
APxC in Right	1.10	0.311	0.04	1.80	0.191	0.07	-	-	-	**4.15**	**0.039**	**0.14**	**8.20**	**0.003**	**0.25**

Note. AP = Anterior/Posterior, H = Hemisphere, C = Condition, SR = Successful Retrieval, SS = Successful Suppression. Significant results are in bold.

Table A.6.3. Follow up ANOVAs of significant interactions between Condition (SR vs. SS) and scalp factors (from Table A.6.2).

	100 to 200ms			200 to 300ms			300 to 400ms			400 to 500ms			500 to 750ms			750 to 1000ms		
	F	*p*	$\eta 2_p$	*F*	*p*	$\eta 2_p$	*F*	*p*	$\eta 2_p$	*F*	*p*	$\eta 2_p$	*F*	*p*	$\eta 2_p$	*F*	*p*	$\eta 2_p$
LA (SR vs. SS)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.45	0.507	0.02
MA (SR vs. SS)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	**9.78**	**0.004**	**0.28**
RA (SR vs. SS)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1.48	0.236	0.06
LC (SR vs. SS)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	**11.45**	**0.002**	**0.31**
MC (SR vs. SS)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
RC (SR vs. SS)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.44	0.514	0.02
LP (SR vs. SS)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	**20.90**	**<.001**	**0.46**
MP (SR vs. SS)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	**8.21**	**0.008**	**0.25**
RP (SR vs. SS)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1.03	0.319	0.04

	1000 to 1500ms			1500 to 2000ms			2000 to 2500ms			2500 to 3000ms			3000 to 3500ms		
	F	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$
LA (SR vs. SS)	**5.12**	**0.033**	**0.17**	**11.00**	**0.003**	**0.31**	-	-	-	**10.92**	**0.003**	**0.31**	**12.39**	**0.002**	**0.33**
MA (SR vs. SS)	**13.01**	**0.001**	**0.34**	**14.35**	**0.001**	**0.37**	-	-	-	**13.72**	**0.001**	**0.35**	**12.91**	**0.001**	**0.34**
RA (SR vs. SS)	1.52	0.229	0.06	**7.79**	**0.01**	**0.24**	-	-	-	**11.52**	**0.002**	**0.32**	**13.68**	**0.001**	**0.35**
LC (SR vs. SS)	-	-	-	-	-	-	-	-	-	-	-	-	**11.13**	**0.003**	**0.31**
MC (SR vs. SS)	-	-	-	-	-	-	-	-	-	**8.24**	**0.008**	**0.25**	**4.96**	**0.035**	**0.17**
RC (SR vs. SS)	-	-	-	-	-	-	-	-	-	**28.49**	**<.001**	**0.53**	**26.82**	**<.001**	**0.52**
LP (SR vs. SS)	12.36	0.002	0.33	-	-	-	-	-	-	-	-	-	3.52	0.072	0.12
MP (SR vs. SS)	**7.85**	**0.01**	**0.24**	-	-	-	-	-	-	2.66	0.115	0.10	0.11	0.739	0.01
RP (SR vs. SS)	3.42	0.076	0.12	-	-	-	-	-	-	**5.13**	**0.032**	**0.17**	2.15	0.155	0.08

Note. AP = Anterior/Posterior, H = Hemisphere, C = Condition, SR = Successful Retrieval, SS = Successful Suppression, LA = Left Anterior, MA = Mid Anterior, RA = Right Anterior, LC = Left Central, MC = Mid Central, RC = Right Central, LP = Left Parietal, MP = Mid Parietal, RP = Right Parietal. Significant results are in bold.

Table A.6.4. Follow up ANOVAs of significant Intrusions vs. Successful Retrieval effects in omnibus ANOVA from Table A.6

	100 to 200ms			200 to 300ms			300 to 400ms			400 to 500ms			500 to 750ms			750 to 1000ms		
	F	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$
I vs. SR in Anterior	**17.80**	**<.001**	**0.42**	**6.80**	**0.015**	**0.21**	-	-	-	1.76	0.196	0.07	-	-	-	0.03	0.874	0.00
HxC in Anterior	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	**3.95**	**0.03**	**0.14**
I vs. SR in Central	**4.26**	**0.049**	**0.15**	1.80	0.192	0.07	-	-	-	0.00	0.986	0.00	-	-	-	1.23	0.278	0.05
HxC in Central	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	2.09	0.136	0.08
I vs. SR in Posterior	0.14	0.71	0.01	0.27	0.606	0.01	-	-	-	2.74	0.11	0.10	-	-	-	**4.70**	**0.04**	**0.16**
HxC in Posterior	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	**6.38**	**0.006**	**0.20**
I vs. SR in Left	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	2.54	0.123	0.09
APxC in Left	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	**7.49**	**0.007**	**0.23**
I vs. SR in Middle	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	2.37	0.136	0.09
APxC in Middle	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.27	0.693	0.01
I vs. SR in Right	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.00	0.982	0.00
APxC in Right	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1.26	0.285	0.05

	1000 to 1500ms			1500 to 2000ms			2000 to 2500ms			2500 to 3000ms			3000 to 3500ms		
	F	*p*	$\eta 2_p$	*F*	*p*	$\eta 2_p$	*F*	*p*	$\eta 2_p$	*F*	*p*	$\eta 2_p$	*F*	*p*	$\eta 2_p$
I vs. SR in Anterior	0.27	0.61	0.01	-	-	-	-	-		-	-	-	-	-	-
HxC in Anterior	**4.93**	**0.013**	**0.17**	-	-	-	-	-	-	-	-	-	-	-	-
I vs. SR in Central	**5.61**	**0.026**	**0.18**	-	-	-	-	-	-	-	-	-	-	-	-
HxC in Central	2.53	0.11	0.09	-	-	-	-	-	-	-	-	-	-	-	-
I vs. SR in Posterior	**6.14**	**0.02**	**0.20**	-	-	-	-	-	-	-	-	-	-	-	-
HxC in Posterior	**3.47**	**0.049**	**0.12**	-	-	-	-	-	-	-	-	-	-	-	-
I vs. SR in Left	**5.07**	**0.033**	**0.17**	-	-	-	-	-	-	-	-	-	-	-	-
APxC in Left	3.79	0.05	0.13	-	-	-	-	-	-	-	-	-	-	-	-
I vs. SR in Middle	**5.07**	**0.033**	**0.17**	-	-	-	-	-	-	-	-	-	-	-	-
APxC in Middle	0.02	0.91	0.00	-	-	-	-	-	-	-	-	-	-	-	-
I vs. SR in Right	1.12	0.29	0.04	-	-	-	-	-	-	-	-	-	-	-	-
APxC in Right	1.65	0.21	0.06	-	-	-	-	-	-	-	-	-	-	-	-

Note. AP = Anterior/Posterior, H = Hemisphere, C = Condition, I = Intrusions, SR = Successful Retrieval. Significant results are in bold.

Table A.6.5. Follow up ANOVAs of significant interactions between Condition (I vs. SR) and scalp factors (from Table A.6.4).

	100 to 200ms			200 to 300ms			300 to 400ms			400 to 500ms			500 to 750ms			750 to 1000ms		
	F	*p*	$\eta 2_p$	*F*	*p*	$\eta 2_p$	*F*	*p*	$\eta 2_p$	*F*	*p*	$\eta 2_p$	*F*	*p*	$\eta 2_p$	*F*	*p*	$\eta 2_p$
LA (I vs. SR)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.47	0.499	0.02
MA (I vs. SR)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.63	0.434	0.03
RA (I vs. SR)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.49	0.491	0.02
LC (I vs. SR)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	3.42	0.076	0.12
MC (I vs. SR)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
RC (I vs. SR)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
LP (I vs. SR)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	**11.51**	**0.002**	**0.32**
MP (I vs. SR)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	4.10	0.054	0.14
RP (I vs. SR)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.40	0.532	0.02

	1000 to 1500ms			1500 to 2000ms			2000 to 2500ms			2500 to 3000ms			3000 to 3500ms		
	F	*p*	$\eta 2_p$	*F*	*p*	$\eta 2_p$	*F*	*p*	$\eta 2_p$	*F*	*p*	$\eta 2_p$	*F*	*p*	$\eta 2_p$
LA (I vs. SR)	0.06	0.815	0.00	-	-	-	-	-	-	-	-	-	-	-	-
MA (I vs. SR)	1.94	0.176	0.07	-	-	-	-	-	-	-	-	-	-	-	-
RA (I vs. SR)	0.11	0.743	0.00	-	-	-	-	-	-	-	-	-	-	-	-
LC (I vs. SR)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
MC (I vs. SR)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
RC (I vs. SR)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
LP (I vs. SR)	**9.30**	**0.005**	**0.27**	-	-	-	-	-	-	-	-	-	-	-	-
MP (I vs. SR)	**4.29**	**0.049**	**0.15**	-	-	-	-	-	-	-	-	-	-	-	-
RP (I vs. SR)	2.21	0.149	0.08	-	-	-	-	-	-	-	-	-	-	-	-

Note. AP = Anterior/Posterior, H = Hemisphere, C = Condition, I = Intrusions, SR = Successful Retrieval, LA = Left Anterior, MA = Mid Anterior, RA = Right Anterior, LC = Left Central, MC = Mid Central, RC = Right Central, LP = Left Parietal, MP = Mid Parietal, RP = Right Parietal. Significant results are in bold.

Table A.6.6. Follow up ANOVAs of significant Intrusions vs. Successful Suppression effects in omnibus ANOVA from Table A.6

	100 to 200ms			200 to 300ms			300 to 400ms			400 to 500ms			500 to 750ms			750 to 1000ms		
	F	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$
I vs. SS in Anterior	**6.71**	**0.016**	**0.21**	1.34	0.258	0.05	-	-	-	**4.393**	**0.046**	**0.15**	-	-	-	-	-	-
HxC in Anterior	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
I vs. SS in Central	0.90	0.352	0.04	0.04	0.836	0.00	-	-	-	0.963	0.336	0.04	-	-	-	-	-	-
HxC in Central	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
I vs. SS in Posterior	0.80	0.437	0.03	1.91	0.179	0.07	-	-	-	0.133	0.74	0.00	-	-	-	-	-	-
HxC in Posterior	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
I vs. SS in Left	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
APxC in Left	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
I vs. SS in Middle	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
APxC in Middle	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
I vs. SS in Right	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
APxC in Right	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

	1000 to 1500ms			1500 to 2000ms			2000 to 2500ms			2500 to 3000ms			3000 to 3500ms		
	F	*p*	$\eta 2_p$	*F*	*p*	$\eta 2_p$	*F*	*p*	$\eta 2_p$	*F*	*p*	$\eta 2_p$	*F*	*p*	$\eta 2_p$
I vs. SS in Anterior	-	-	-	-	-	-	-	-	-	-	-	-	2.83	0.105	0.10
HxC in Anterior	-	-	-	-	-	-	-	-	-	-	-	-	**5.73**	**0.007**	**0.19**
I vs. SS in Central	-	-	-	-	-	-	-	-	-	-	-	-	**5.05**	**0.034**	**0.17**
HxC in Central	-	-	-	-	-	-	-	-	-	-	-	-	1.30	0.28	0.05
I vs. SS in Posterior	-	-	-	-	-	-	-	-	-	-	-	-	**6.76**	**0.015**	**0.21**
HxC in Posterior	-	-	-	-	-	-	-	-	-	-	-	-	0.57	0.559	0.02
I vs. SS in Left	-	-	-	-	-	-	-	-	-	-	-	-	**4.38**	**0.047**	**0.15**
APxC in Left	-	-	-	-	-	-	-	-	-	-	-	-	0.30	0.645	0.01
I vs. SS in Middle	-	-	-	-	-	-	-	-	-	-	-	-	**8.19**	**0.008**	**0.25**
APxC in Middle	-	-	-	-	-	-	-	-	-	-	-	-	1.26	0.279	0.05
I vs. SS in Right	-	-	-	-	-	-	-	-	-	-	-	-	3.06	0.093	0.11
APxC in Right	-	-	-	-	-	-	-	-	-	-	-	-	1.08	0.342	0.04

Note. AP = Anterior/Posterior, H = Hemisphere, C = Condition, I = Intrusions, SS = Successful Suppression. Significant results are in bold.

Table A.6.7. Follow up ANOVAs of significant interactions between Condition (I vs. SS) and scalp factors (from Table A.6.6).

	100 to 200ms			200 to 300ms			300 to 400ms			400 to 500ms			500 to 750ms			750 to 1000ms		
	F	*p*	$\eta 2_p$	*F*	*p*	$\eta 2_p$	*F*	*p*	$\eta 2_p$	*F*	*p*	$\eta 2_p$	*F*	*p*	$\eta 2_p$	*F*	*p*	$\eta 2_p$
LA (I vs. SS)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
MA (I vs. SS)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
RA (I vs. SS)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
LC (I vs. SS)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
MC (I vs. SS)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
RC (I vs. SS)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
LP (I vs. SS)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
MP (I vs. SS)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
RP (I vs. SS)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

	1000 to 1500ms			1500 to 2000ms			2000 to 2500ms			2500 to 3000ms			3000 to 3500ms		
	F	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$	*F*	*p*	$\eta2_p$
LA (I vs. SS)	-	-	-	-	-	-	-	-	-	-	-	-	1.60	0.219	0.06
MA (I vs. SS)	-	-	-	-	-	-	-	-	-	-	-	-	**5.72**	**0.025**	**0.19**
RA (I vs. SS)	-	-	-	-	-	-	-	-	-	-	-	-	0.74	0.397	0.03
LC (I vs. SS)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
MC (I vs. SS)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
RC (I vs. SS)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
LP (I vs. SS)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
MP (I vs. SS)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
RP (I vs. SS)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Note. AP = Anterior/Posterior, H = Hemisphere, C = Condition, SR = Successful Retrieval, SS = Successful Suppression, LA = Left Anterior, MA = Mid Anterior, RA = Right Anterior, LC = Left Central, MC = Mid Central, RC = Right Central, LP = Left Parietal, MP = Mid Parietal, RP = Right Parietal. Significant results are in bold

Appendix B. Supplementary results for Experiment 2 (Chapter 4).

B.1. Phenomenological and emotional characteristics of autobiographical memories.

A 2 (MT, Morally Wrong, Morally Right) x 2 (Group, Standard-Guilty, Suppression-Guilty) mixed ANOVA was conducted for memory age, vividness, and intentionality. There were main effects of MT for all three measures: As expected, participants rated morally wrong memories as less vivid, $F(1,36) = 6.98$, $p = .012$, $\eta_p^2 = .162$, occurring more distant in the past, $F(1,36) = 14.68$, $p <.001$, $\eta_p^2 = .29$, and their actions as less intentional, $F(1,36) = 32.46$, $p <.001$, $\eta_p^2 = .34$ than morally right memories. Therefore, morally wrong memories were perceived as more distant than morally right memories, replicating previous findings. This effect was regardless of Group, as memory type did not interact with Group (all *Fs* < .86, *ps* > .357).

A 3 (Memory Type, MT: Morally Wrong vs. Morally Right vs. Mock Crime) x 2 (Instruction Group, Group: Standard-Guilty vs. Suppression-Guilty) mixed ANOVA was conducted using morality, and emotional measures of guilt, shame, pleasure, and arousal, both at time of event and at the time of the experiment. There were no significant effects for arousal (at time of memory). There were main effects of memory type for all measures except arousal (time of the experiment), for which memory type interacted with Group. Memory type did not interact with Group for other measures. Please see Table B.1. for results of the omnibus and follow-up ANOVAs.

To follow up on the main effect of MT, 3 repeated-measures ANOVAs, each comparing two memory types (e.g., Mock Crime vs. Morally Right), were conducted for morality, guilt, shame, pleasure, and arousal (both at time of experiment and memory). Participants felt more guilty, ashamed, upset, felt less pleasure, and rated their actions as more morally wrong after describing both morally wrong and mock crime memories compared to after describing morally right memories. The same directional differences were found for emotions participants reported feeling at the time of the memory. Interestingly, although both mock crime and morally wrong memories did not differ in moral *wrongness*, participants indicated feeling more guilty, ashamed, upset, and felt less pleasure after describing morally wrong than the mock crime memory.

Table B.1. Average phenomenological and emotional ratings across memory types.

	Morally Right	Morally Wrong	Mock Crime
Memory Age	1.74 (1.97)	4.27 (3.47)	-
Vividness	3.76 (.85)	3.23 (1.10)	-
Intentionality	4.36 (.67)	3.18 (1.29)	-
Morality	5.92 (1.44)	2.13 (.74)	2.42 (1.11)
Ashamed (experiment)	1.24 (.85)	3.11 (1.15)	1.97 (1.05)
Guilty (experiment)	1.29 (.89)	3.29 (1.27)	2.26 (1.17)
Pleasure (experiment)	3.79 (1.09)	2.28 (.77)	2.89 (1.08)
Ashamed (memory)	1.42 (1.05)	2.89 (1.47)	2.00 (1.11)
Guilty (memory)	1.34 (.87)	3.23 (1.49)	2.13 (1.14)
Pleasure (memory)	3.5 (1.41)	2.55 (1.31)	3.00 (.99)
Arousal (memory)	3.61 (1.05)	3.92 (1.04)	3.61 (1.12)

Note. "Experiment" refers to emotions experienced at the time of remembering and describing the memory. "Memory" refers to emotions experienced. Standard deviations are reported in brackets.

To follow up on the interaction between MT and Group for arousal at the time of the experiment, a 3 (MT) repeated-measures ANOVA was conducted for each group. There was a significant effect of MT in the Standard-, but not the Suppression-Guilty group. As a follow up 3 repeated-measures ANOVAs comparing two memory types each for the Standard-Guilty group. Results indicated that mock crime memories were rated as significantly more arousing than morally right memories, whereas the comparison between mock crime and morally wrong, and between morally wrong and right memories were not significant. See Table B.2. For descriptive statistics of follow up tests.

Table B.2. Average ratings of arousal for each memory type compared across standard and suppression guilty groups

		Arousal (experiment)
	Morally Right	2.11 (.76)
Standard-Guilty	Morally Wrong	2.56 (1.10)
	Mock Crime	2.94 (1.11)
	Morally Right	2.80 (1.11)
Suppression-Guilty	Morally Wrong	2.65 (1.23)
	Mock Crime	2.30 (.98)

Note. Standard deviations are reported in brackets.

B.2. Focal analysis of LPN and peak-to-peak measures

See figure B.1. for results from the LPN and peak-to-peak measures. See table B.3. for results for pairwise comparison of these measures for each memory type in each instruction group.

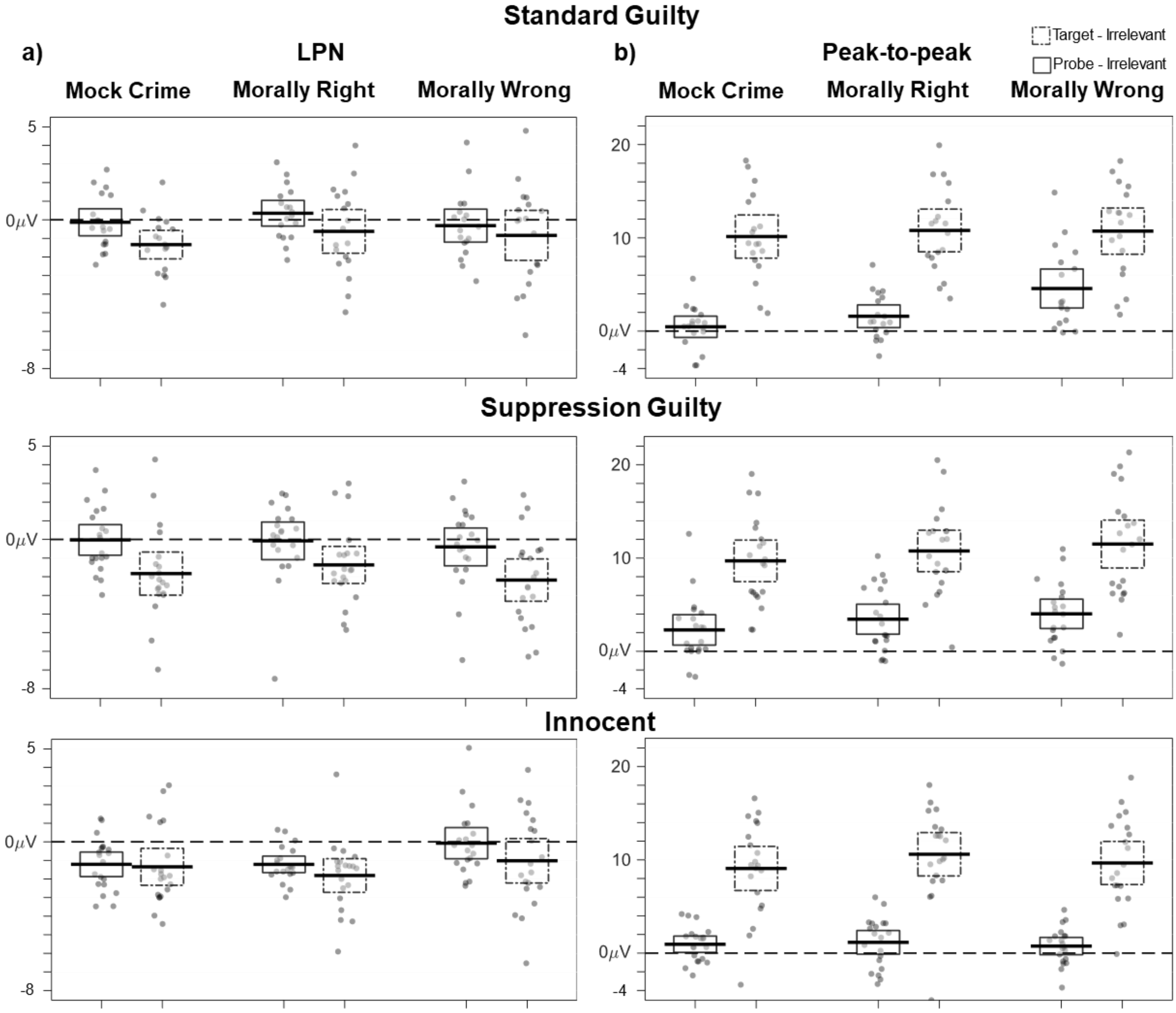


Figure B.1. Illustration of individually estimated amplitudes of probe-irrelevant and target-irrelevant differences in a) LPN and b) peak-to-peak effects at Pz. The scatter dots show P300 differences between pairwise conditions for each individual. The thick lines show the group means and the boxes depict the 95% confidence interval of the group means.

Target vs. Irrelevant. In the standard-guilty group, the LPN was significantly enhanced for targets compared to irrelevants when the CIT tested for mock crime memories, but not for the two moral memories. In the suppression-guilty group, all three blocks elicited the LPN for targets compared to irrelevants. In the innocent group, mock crime and morally right memory blocks showed a significant target-irrelevant LPN effect, but this LPN difference was not significant in the morally wrong memory block. Results for the peak-to-peak measure mirrored the P300 results, with significant target-irrelevant peak-to-peak differences for all memory types in all groups

The target-irrelevant LPN (all *Fs* < 3.01, *ps* > .09) and PtoP effects (all *Fs* < 3.02, *ps* > .098) were not affected by instruction for any target words in any memory type block.

For the LPN, there were no significant differences between any memory type in any group (all *Fs* < 2.01, *ps* > .17). For the peak-to-peak measure, mock crime memories elicited a greater LPN than morally right memories in the innocent group, *F*(1,19) = 5.41, *p* = .032, *partial* η^2 = .221. There were no other significant effects in any group for any memory type for this measure (all *Fs* < 3.02, *ps* > .098).

Probe vs. Irrelevant. The LPN was not significantly enhanced for any memory probe type compared to irrelevants in either of the guilty groups, but in the innocent group, a significant probe-irrelevant LPN difference was found for mock crimes and morally right memories, which is surprising since the innocent group did not have an episodic memory associated with those probes. The peak-to-peak measure mirrored the P300 effects.

The probe-irrelevant LPN effect was greater in the innocent group than both standard guilty and suppression guilty groups for both mock crime (Standard: *F*(1,36) = 6.09, *p* = .019, *partial* η^2 = .145; Suppression: *F*(1,38) = 5.96, *p* = .019, *partial* η^2 = .136) and morally right memories (Standard: *F*(1,36) = 14.85, *p* < .001, *partial* η^2 = .29, Suppression: *F*(1,38) = 4.07, *p* = .05, *partial* η^2 = .097). There were no significant differences in this effect between groups when comparing morally wrong memories with the other two memory types, and there were no differences between standard- and suppression-guilty groups for any memory type (all *Fs* < .53, *ps* > .47).

The peak-to-peak probe-irrelevant LPN effect was greater in both standard-guilty (*F*(1,36) = 12.89, *p* = .001, *partial* η^2 = .264) and suppression-guilty (*F*(1,38) = 13.45, *p* = .001, *partial* η^2 = .263) than innocent groups for morally wrong memories. Morally right memories showed a similar pattern but only in the suppression vs innocent comparison

($F(1,38) = 5.56$, $p = .024$, *partial* $\eta^2 = .128$). Both mock crime and morally right memories elicited greater peak-to-peak LPN effects in the suppression compared to standard groups, but this was trend-level significant (both *Fs* < 3.91, *ps* > .056). There were no other significant effects (all *Fs* < . 94, *ps* > .337).

Morally right memories elicited a greater probe-irrelevant LPN effect than morally wrong memories in the innocent group, $F(1,19) = 6.95$, $p = .016$, *partial* $\eta^2 = .268$. There were no significant differences between any other memory type in any group for the probe-irrelevant LPN (all *Fs* < 2.01, *ps* > .17). The peak-to-peak measure mirrored the P300 results: In the standard-guilty group, morally wrong memories elicited a larger peak-to-peak probe-irrelevant effect than both morally right $F(1,17) = 10.342$, $p = .005$, *partial* $\eta^2 = .37$, and mock crime ($F(1,17) = 12.79$, $p = .002$, *partial* $\eta^2 = .428$) memories. The probe-irrelevant peak-to-peak difference between morally right and mock crime memories was not significant, $F(1,17) = 2.38$, $p = .14$, *partial* $\eta^2 = .12$. There were no significant differences between memory types in the peak-to-peak measure in either the suppression-guilty or innocent groups (all *Fs* < 1.25, *p*s > .28).

See Table B.4 for a comparison of focal and global analysis results for the LPN measure.

Table B.3. Results from several pairwise t-tests comparing both target vs. irrelevant and probe vs. irrelevant stimulus types for the LPN and PtoP measures.

		Target vs. Irrelevant LPN			Probe vs. Irrelevant LPN			Target vs. Irrelevant PtoP			Probe vs. Irrelevant PtoP		
		t	*p*	*Cohen's d*	*t*	*p*	*Cohen's d*	*t*	*p*	*Cohen's d*	*t*	*p*	*Cohen's d*
Standard Guilty	Mock Crime	**-3.64**	**0.002**	**-0.39**	0.29	0.772	-0.05	**9.21**	**<.001**	**1.76**	0.73	0.473	0.12
	Morally Right	-1.17	0.26	-0.15	-1.09	0.293	0.1	**9.97**	**<.001**	**2.09**	**2.76**	**0.013**	**0.41**
	Morally Wrong	-1.31	0.206	-0.23	0.77	0.454	-0.11	**9.12**	**<.001**	**2.08**	**4.61**	**<.001**	**1.04**
Suppression Guilty	Mock Crime	**-3.32**	**0.004**	**-0.59**	0.07	0.947	-0.01	**9.12**	**<.001**	**1.61**	**2.95**	**0.008**	**0.53**
	Morally Right	**-2.89**	**0.01**	**-0.41**	0.16	0.876	-0.02	**10.14**	**<.001**	**1.76**	**4.49**	**<.001**	**0.63**
	Morally Wrong	**-4.02**	**0.001**	**-0.61**	0.83	0.417	-0.12	**9.38**	**<.001**	**1.86**	**5.35**	**<.001**	**0.85**
Innocent	Mock Crime	**-2.95**	**0.008**	**-0.41**	**3.99**	**0.001**	**-0.44**	**7.93**	**<.001**	**1.77**	**2.57**	**0.019**	**0.28**
	Morally Right	**-4.26**	**<.001**	**-0.58**	**5.17**	**<.001**	**-0.52**	**9.59**	**<.001**	**2.14**	1.79	0.089	0.36
	Morally Wrong	-1.82	0.085	-0.35	0.21	0.836	-0.03	**8.67**	**<.001**	**2**	1.72	0.102	0.2

Note. Significant results are in bold. *Standard-guilty N = 18, Suppression-guilty N = 20, Innocent N = 20*

Table B.4. Comparison of two analyses strategies for different ERP components from the intrusions analysis

		Target vs. Irrelevant LPN		Probe vs. Irrelevant LPN	
		Focal Analysis	*Global Analysis*	*Focal Analysis*	*Global Analysis*
Standard-guilty	Mock Crime	✓	✓	✗	✗
	Morally Right	✗	✗	✗	✗
	Morally Wrong	✗	✗	✗	✗
Suppression-guilty	Mock Crime	✓	✗	✗	✗
	Morally Right	✓	✓	✗	✗
	Morally Wrong	✓	✓	✗	✗
Innocent	Mock Crime	✓	✓	✓	✗
	Morally Right	✓	✓	✓	✗
	Morally Wrong	✗	✗	✗	✗

Note. A tick mark indicates that the component was statistically significant in that analysis, whereas a cross indicates that the component was not statistically significant. Please note that the results are summarised as components here only for illustrative purposes, see text and discussion for more detail.

B.3. Individual-guilt classification results

Peak-to-peak measure. In the standard-guilty group, morally wrong memories had the highest classification rates (Bootstrap: 14/18, permutation: 13/18), followed by morally right memories (Bootstrap: 8/18, permutation: 6/18), followed by mock crime memories (Bootstrap: 6/18, permutation: 1/18). A similar pattern was found in the suppression-guilty group, morally wrong memories had highest classification rates (Bootstrap: 14/20, permutation: 13/20), followed by morally right memories (Bootstrap: 12/20, permutation: 10/20), followed by mock crime memories (Bootstrap: 10/20, permutation: 9/20). The false positive classification rates were surprisingly high in the innocent group, especially from the bootstrap analysis. Both mock crime and morally wrong memories had similar classification rates (Bootstrap: 9/20, permutation: 4/20), whereas morally right memories had higher classification (Bootstrap: 11/20, permutation: 7/20).

ROC analysis was also conducted on this data (see Table B.5.). When comparing the standard guilty group to the innocent group, both bootstrap- and permutation-based classification was better than chance for both morally wrong memories, but for both mock crime and morally right memories the AUC was not significant for either bootstrap or permutation scores. Similarly, guilt classification using both techniques were better than chance for both the morally wrong memories in the suppression group, whereas for mock crime memories the AUC was not significant for either bootstrap or permutation scores. For morally right memories, bootstrap-based classification was greater than chance, but the AUC was not significant for permutation-based classification.

Table B.5. Results from the ROC analysis for the peak-to-peak measure.

	Standard vs. Innocent					
	Bootstrap			Permutation		
	AUC	*SE*	*p*	*AUC*	*SE*	*p*
Mock Crime	0.46	0.10	0.65	0.52	0.10	0.86
Morally Right	0.57	0.10	0.48	0.39	0.09	0.27
Morally Wrong	**0.78**	**0.08**	**0.003**	**0.80**	**0.07**	**0.002**
	Suppress vs. Innocent					
	Bootstrap			Permutation		
	AUC	*SE*	*p*	*AUC*	*SE*	*p*
Mock Crime	0.625	0.09	0.176	0.66	0.09	0.086
Morally Right	**0.68**	**0.09**	**0.047**	0.62	0.09	0.208
Morally Wrong	**0.75**	**0.08**	**0.008**	**0.80**	**0.07**	**0.001**

Note. Statistical significance is indicated in bold.

Classification results for each individual. The bootstrap and permutation test results for both P300 and peak-to-peak measures for each individual in the experiment is reported Table B.6.

Table B.6. Results at the participant level from the bootstrap and permutation tests from the concealed-information test.

	Standard Guilty											
	P300 peak						Peak-to-Peak					
	Mock Crime		Morally Right		Morally Wrong		Mock Crime		Morally Right		Morally Wrong	
Participant	Boot	Perm	Boot	Perm	Boot	Perm	Boot	Perm	Boot	Perm	Boot	Perm
1	17%	0.82	**91%**	**0.09**	**100%**	**<.001**	64%	0.52	40%	0.78	**100%**	**0.044**
2	13%	0.97	**100%**	**0.004**	**97%**	**0.026**	38%	0.97	**96%**	0.22	**97%**	**0.043**
3	**99%**	**0.027**	65%	0.55	73%	0.78	**96%**	0.49	**98%**	0.13	**97%**	0.75
4	**95%**	**0.041**	55%	0.62	65%	0.50	84%	0.31	58%	0.40	72%	0.34
5	34%	1.00	**99%**	**0.01**	**98%**	**0.005**	19%	1.00	87%	**0.091**	84%	0.59
6	76%	0.40	7%	0.95	**98%**	**0.043**	57%	0.71	55%	0.82	77%	0.38
7	63%	0.39	78%	0.25	**100%**	**<.001**	**90%**	0.34	86%	0.43	**100%**	**<.001**
8	64%	0.46	59%	0.48	**98%**	**0.02**	81%	0.39	58%	0.57	**100%**	**0.006**
9	**96%**	**0.07**	**93%**	**0.044**	**95%**	**<.001**	58%	0.17	**100%**	**0.011**	**100%**	**<.001**
10	78%	0.24	**96%**	0.11	**99%**	**0.026**	**99%**	0.11	**96%**	0.38	**98%**	**0.075**
11	1%	0.995	**100%**	**<.001**	**100%**	**<.001**	6%	0.99	**100%**	**<.001**	**100%**	**<.001**
12	**100%**	**0.001**	**90%**	0.108	**100%**	**<.001**	**100%**	**<.001**	53%	0.73	**100%**	**<.001**
13	**100%**	**0.006**	2%	0.982	76%	0.273	**91%**	0.17	4%	1.00	**98%**	**0.084**
14	19%	0.82	24%	0.827	**100%**	**0.001**	64%	0.41	42%	0.77	**100%**	**<.001**
15	**93%**	**0.048**	**100%**	**<.001**	86%	0.216	47%	0.59	**100%**	**0.007**	77%	0.37
16	**94%**	**0.07**	86%	0.184	**99%**	**0.012**	59%	0.66	62%	0.47	**98%**	**0.082**
17	1%	1	89%	0.131	**100%**	**<.001**	8%	1.00	**99%**	**0.031**	**100%**	**<.001**
18	**93%**	**0.07**	**100%**	**<.001**	**100%**	**<.001**	**97%**	0.45	**100%**	**0.004**	**100%**	**<.001**
Total guilty Ps	8	8	8	7	14	14	6	1	8	6	14	13

	Suppression Guilty											
	P300 peak						Peak-to-Peak					
	Mock Crime		Morally Right		Morally Wrong		Mock Crime		Morally Right		Morally Wrong	
Participant	Boot	Perm	Boot	Perm	Boot	Perm	Boot	Perm	Boot	Perm	Boot	Perm
1	13%	0.89	**94%**	**0.09**	74%	0.34	88%	0.57	**99%**	0.13	85%	0.29
2	86%	0.16	**99%**	**0.001**	**100%**	**<.001**	**99%**	**0.012**	**100%**	**<.001**	**100%**	**<.001**
3	**100%**	**<.001**	22%	0.87	**96%**	**0.033**	**100%**	**<.001**	45%	0.89	**96%**	**0.037**
4	**98%**	**0.043**	**100%**	**<.001**	**100%**	**<.001**	**100%**	**0.014**	**100%**	**<.001**	**100%**	**<.001**
5	**100%**	**<.001**	**100%**	**<.001**	47%	0.74	**100%**	**<.001**	**100%**	**<.001**	19%	0.89
6	33%	0.79	87%	0.14	**91%**	**0.09**	73%	0.56	84%	0.11	**97%**	**0.022**
7	10%	0.89	**100%**	**0.001**	**100%**	**<.001**	4%	0.98	**94%**	**0.052**	**100%**	**0.002**
8	28%	0.72	**100%**	**0.005**	**100%**	**<.001**	67%	0.64	**99%**	**0.026**	**100%**	**<.001**
9	**100%**	**0.001**	80%	0.28	**100%**	**<.001**	**90%**	0.10	86%	0.58	**100%**	**<.001**
10	**100%**	**<.001**	**100%**	**<.001**	**100%**	**<.001**	**99%**	**0.019**	**100%**	**<.001**	**100%**	**0.002**
11	76%	0.26	**98%**	**0.018**	30%	0.76	73%	0.34	79%	0.30	56%	0.56
12	**99%**	**0.006**	**100%**	**<.001**	**100%**	**0.004**	**100%**	**0.004**	**100%**	**<.001**	**100%**	**<.001**
13	83%	0.21	**100%**	**<.001**	**100%**	**<.001**	85%	0.48	**100%**	**<.001**	**100%**	**<.001**
14	24%	0.86	6%	0.91	**97%**	**0.025**	69%	0.71	**100%**	**0.007**	**98%**	0.19
15	34%	0.69	88%	0.21	67%	0.38	62%	0.55	31%	0.76	42%	0.83
16	38%	0.63	**100%**	**<.001**	86%	0.12	4%	1.00	**100%**	**<.001**	70%	0.19
17	**98%**	**0.028**	83%	0.18	**95%**	**0.097**	**100%**	**0.013**	68%	0.34	**100%**	**0.008**
18	**98%**	**0.021**	66%	0.34	82%	0.28	**98%**	**0.09**	75%	0.45	79%	**0.011**
19	70%	0.34	**92%**	0.22	**100%**	**<.001**	**99%**	**0.037**	63%	0.89	**100%**	**<.001**
20	89%	0.12	**94%**	**0.05**	82%	0.18	54%	0.58	**95%**	0.18	**96%**	0.21
Total guilty Ps	8	8	13	11	13	13	10	9	12	10	14	13

	Innocent											
	P300 peak						Peak-to-Peak					
	Mock Crime		Morally Right		Morally Wrong		Mock Crime		Morally Right		Morally Wrong	
Participant	Boot	Perm	Boot	Perm	Boot	Perm	Boot	Perm	Boot	Perm	Boot	Perm
1	73%	0.32	72%	0.25	85%	0.48	**96%**	**0.011**	**97%**	0.11	36%	0.32
2	29%	0.79	0%	1.00	12%	0.92	57%	0.55	0%	1.00	22%	0.96
3	45%	0.63	75%	0.30	**100%**	0.26	25%	0.73	**93%**	0.18	**98%**	0.51
4	21%	0.79	82%	0.15	71%	0.28	**91%**	0.15	**90%**	**0.046**	79%	0.46
5	43%	0.57	1%	1.00	**100%**	**0.001**	**93%**	**0.096**	8%	1.0	**98%**	**0.051**
6	57%	0.96	82%	0.46	48%	0.52	41%	0.98	38%	0.11	**97%**	0.31
7	81%	0.20	80%	0.31	**91%**	0.12	88%	0.20	**98%**	**0.091**	**95%**	0.17
8	80%	0.19	83%	0.21	87%	0.17	**90%**	0.19	**99%**	**0.042**	**94%**	0.17
9	18%	0.91	**99%**	**0.007**	**99%**	**0.005**	77%	0.42	**100%**	**<.001**	**95%**	**0.073**
10	87%	0.24	56%	0.52	86%	0.24	**99%**	**0.016**	70%	0.55	1%	1.00
11	49%	0.62	78%	0.17	9%	0.85	**90%**	0.37	**98%**	0.02	77%	0.25
12	70%	0.27	5%	0.98	12%	0.94	46%	0.88	6%	1.00	67%	0.79
13	49%	0.65	**100%**	**<.001**	51%	0.58	**99%**	0.16	**100%**	**<.001**	34%	0.96
14	56%	0.67	70%	0.30	87%	0.10	47%	0.18	**97%**	0.12	53%	0.82
15	87%	0.15	7%	0.92	34%	0.68	**96%**	0.17	**96%**	0.30	**99%**	0.13
16	15%	0.89	41%	0.63	**100%**	**0.001**	32%	0.94	31%	0.95	**100%**	**0.012**
17	12%	0.90	3%	0.99	**99%**	**0.012**	71%	0.59	1%	1.00	**99%**	**0.007**
18	68%	0.35	3%	0.99	6%	0.95	**100%**	**0.047**	13%	0.93	36%	0.87
19	15%	0.87	28%	0.82	3%	0.99	34%	0.81	35%	0.64	14%	0.87
20	87%	0.24	**95%**	**0.08**	49%	0.79	50%	0.65	**99%**	**0.041**	70%	0.48
Total guilty Ps	0	0	3	3	5	4	9	4	11	7	9	4

Note. A 90% threshold was used for bootstrap tests, and *p* < .1 was used as a threshold for permutation test. Results crossing these thresholds is in bold for each memory for each participant.

Appendix C. Supplementary information from Experiment 3 (Chapter 5)

C.1. Descriptive statistics of the online emotion rating scale.

Table C.1. Introspective reports of emotions from the think/no-think phases divided across moral memory type, block, instruction type, and group.

	Block 1		Block 2	
	Think	No-think	Think	No-think
	Direct Suppression			
Morally Right	.55 (.43)	-.02 (.39)	.62 (.43)	-.01 (.44)
Morally Wrong	-.39 (.61)	-.31 (.44)	-.42 (.63)	-.24 (.47)
	Thought Substitution			
Morally Right	.47 (.41)	.53 (.37)	.52 (.45)	.60 (.46)
Morally Wrong	-.35 (.64)	.15 (.59)	-.40 (.50)	.29 (.61)
	Counterfactual Thinking			
Morally Right	.56 (.39)	.60 (.46)	.54 (.44)	.65 (.45)
Morally Wrong	-.59 (.39)	.32 (.60)	-.66 (.40)	.37 (.60)

Note. Average Proportion of Responses (in %) are Divided Based on Response Type across all Conditions. Standard Deviations are denoted in Brackets.

C.2. Average ratings of autobiographical memory descriptions provided by independent coders for Experiment 3, Chapter 5

Table C.2. Average similarity ratings provided by independent coders across instruction type, memory type, and strategy group.

	Direct Suppression		Thought Substitution		Counterfactual Thinking	
	Right	Wrong	Right	Wrong	Right	Wrong
Think	3.99 (.93)	3.71 (1.17)	3.99 (.93)	3.71 (1.17)	3.94 (1.14)	3.66 (1.08)
NoThink	4.04 (.99)	3.84 (1.09)	4.04 (.99)	3.84 (1.09)	4.01 (1.00)	3.65 (1.17)

Note. Similarity was rated on a 5-point scale (1 = Not at all similar, 5 = Extremely similar). Standard deviation is reported in brackets.

Table C.3. Average ratings of specificity and valence provided by independent coders across instruction type, memory type, and strategy group.

	Direct Suppression				Thought Substitution				Counterfactual Thinking			
	Right		Wrong		Right		Wrong		Right		Wrong	
	Pre-TNT	Post-TNT	Pre-TNT	Post-TNT	Pre-TNT	Post-TNT	Pre-TNT	Post-TNT	Pre-TNT	Post-TNT	Pre-TNT	Post-TNT
	Specificity											
Think	3.00(.77)	3.05(.81)	2.82(.84)	2.85(.81)	2.76(.94)	2.99(.99)	2.69(.93)	2.77(.98)	2.73(.98)	2.74(1.12)	2.83(.88)	2.83(.93)
NoThink	2.94(.90)	2.92(.92)	2.87(.82)	3.06(.86)	2.75(.92)	2.80(1.02)	2.78(.88)	2.69(.92)	2.94(.95)	2.96(.99)	2.79(.90)	2.77(1.02)
	Valence											
Think	4.33(.82)	4.26(.82)	3.38(.86)	3.51(.82)	4.19(.70)	4.23(.75)	3.63(.90)	3.47(.91)	4.17(.67)	4.13(.92)	3.57(.73)	3.42(.82)
NoThink	4.33(.75)	4.25(.92)	3.25(.89)	3.31(.77)	4.36(.66)	4.31(.85)	3.51(.78)	3.39(.92)	4.10(.79)	4.32(.68)	3.46(.81)	3.52(.87)

Note. Specificity was rated on a 5-point scale (1=Not at all specific to 5 = Extremely specific), emotional valence was measured on a 7-point scale (1=Very Negative, 4 = Not negative or positive, 7 = Very Positive). Standard deviation is reported in brackets.

C.3. Phenomenological and emotional characteristics of autobiographical memories.

Memory characteristics and emotions successfully distinguish morally right and wrong memories. A 2 (Instruction Type, IT: Think vs. No-Think) x 2 (Memory Type, MT: Morally Wrong vs. Morally Right) x 3 (Instruction Group, Group: Direct Suppression vs. Thought Substitution vs. Counterfactual Thinking) mixed ANOVA was conducted with the chosen self-report measures from session one (see Table 1 for ANOVA results). There were no significant effects for arousal. Expectedly, the main and interactions effects of neither instruction type nor Group were significant, but there were main effects of memory type: Participants felt less ashamed, guilty, less pleasurable, and rated their actions as more wrong for morally wrong compared to morally right memories.

Replicating findings from both experiments in the thesis, morally wrong memories were rated as less vivid, less intentional, and older, than morally right memories. Please see Table C.4. for descriptive statistics of all effects presented above.

Table C.4. Results from omnibus ANOVA using self-report ratings of phenomenological and emotional characteristics of memories as measured in session one (before think/no-think task).

	Phenomenological memory characteristics											
	Memory Age			Vividness			Intention			Morality		
	F	*p*	η^2_p	*F*	*p*	η^2_p	*F*	*p*	η^2_p	*F*	*p*	η^2_p
IT	0.59	0.45	0.00	0.00	0.97	0.00	1.88	0.17	0.01	0.10	0.76	0.00
ITxGroup	0.26	0.77	0.00	1.73	0.18	0.03	0.58	0.56	0.01	0.10	0.91	0.00
MT	**121.29**	**<.001**	**0.48**	**58.77**	**<.001**	**0.31**	**233.35**	**<.001**	**0.64**	**2439.42**	**<.001**	**0.95**
MTxGroup	1.32	0.27	0.02	0.61	0.54	0.01	0.13	0.87	0.00	0.05	0.95	0.00
ITxMT	0.19	0.66	0.00	0.06	0.81	0.00	0.79	0.38	0.01	0.03	0.87	0.00
ITxMTxGroup	0.08	0.92	0.00	1.60	0.21	0.02	0.28	0.76	0.00	0.60	0.55	0.01

	I-PANAS-SF						SAM					
	Ashamed			Guilty			Pleasure			Arousal		
	F	*p*	η^2_p	*F*	*p*	η^2_p	*F*	*p*	η^2_p	*F*	*p*	η^2_p
IT	2.84	0.09	0.02	1.03	0.31	0.01	1.96	0.16	0.02	0.72	0.40	0.01
ITxGroup	0.01	0.99	0.00	0.10	0.91	0.00	0.18	0.84	0.00	0.23	0.80	0.00
MT	**443.66**	**<.001**	**0.77**	**558.88**	**<.001**	**0.81**	**403.48**	**<.001**	**0.75**	2.70	0.10	0.02
MTxGroup	0.44	0.65	0.01	0.30	0.74	0.00	0.67	0.52	0.01	0.50	0.61	0.01
ITxMT	1.20	0.28	0.01	0.01	0.94	0.00	2.60	0.11	0.02	2.22	0.14	0.02
ITxMTxGroup	0.52	0.60	0.01	0.26	0.78	0.00	1.39	0.25	0.02	0.04	0.96	0.00

Note. IT = Instruction Type, MT = Memory Type, Group = Instruction Group (between-subjects variable). η^2_p = partial eta sq. (effect size). I-PANAS-SF = International – Positive and Negative Affect Scale – Short Form, SAM = Self-Assesssment Manikin. Significant effects ($p < .05$) are shown in bold. $N = 135$.