How Do Ageing Stereotypes Affect Performance on a Previously Unexplored False Memory Task?

A thesis

submitted in fulfilment

of the requirements for the degree

of

Master of Science in Psychology

at

The University of Waikato

by

Sukhandeep Ubha

THE UNIVERSITY OF
WAIKATO
Te Whare Wānanga o Waikato

2022

Abstract

False memories are memories that are thought to have happened but did not happen. Age-related neurological changes can cause an increase in false memories; however, researchers have started focusing on stereotype threat, to explain this increase. Stereotype threat occurs when negative beliefs about a particular group are introduced, causing them to perform poorly on tasks. Studies focused on the effects of age-based stereotype threat on false memories have been conflicted due to the methodological differences (such as the use of retrieval task and the focus of the task). Thus, Experiment 1 aimed to account for those methodological differences and investigated whether stereotype threat influenced performance on the memory conjunction task. In *Experiment 1*, 60 participants were given a memory conjunction task online and were randomly assigned to either a control or a stereotype threat group. Participants were then given a free recall and a recognition task. They were also asked to indicate their levels of education, employment and importance given to memory (moderating factors), as these factors can affect performance. The results indicated no difference in performance between the control and the stereotyped participants. This could be due to participants finding the memory conjunction task difficult, thus choosing to disengage from the task. Therefore, Experiment 2 aimed to investigate whether there were any changes in performance when an online DRM task with a different focus (prevention or promotion) was used. In *Experiment 2*, 120 participants were given a DRM task online and were randomly assigned to the stereotype threat-promotion, stereotype threat-prevention, control-promotion and control-prevention groups. This experiment did not find any significant differences between the groups. It is possible that previous studies mistakenly induced subtle and blatant stereotype threats while the present experiment only used blatant threat. Recruiting participants from an online platform could also explain why present study found discrepant findings compared to the previous research. Even though present study was

unable to replicate previous findings, it still added to the limited literature and focused on different methodological factors all at once. It is essential to keep investigating and focusing on different methodological factors to better understand a very complex topic. The threat could cause older adults to be diagnosed with dementia and receive unnecessary treatment when, in fact, they are healthy.

Keywords: DRM, false memory, memory conjunction task, stereotype threat

Acknowledgements

I would like to thank my supervisor, Aleea Devitt, for guiding me in the right direction. Thank you for being so amazing and always encouraging. Secondly, I would like to thank my co-supervisor, Sam Charlton, for helping me towards the end of my thesis when Aleea was on her leave. I appreciate the guidance; your experience really helped me and gave me confidence to finalise my thesis. Thirdly, my family and friends, for not letting me give up when I was tired of not being able to find participants for my study or was unable to make progress on my thesis. Last but not least, to all community and online participants that took part in this study. Thank you for taking the time to complete this study.

Table of Contents

List of Figures

Introduction

Long-term memories refer to memories stored for longer durations, days, years, or even a lifetime (Almaraz-Espinoza & Grider, 2021; Zlotnik & Vansintjan, 2019). This type of memory enables individuals to learn new skills such as driving, remembering facts, and different life experiences (Almaraz-Espinoza & Grider, 2021; Campos et al., 2010; Fossati, 2013). For example, a person is able to remember their wedding date and memories of their wedding day because of long-term memory (Almaraz-Espinoza & Grider, 2021; Campos et al., 2010). Importantly, long-term memories (e.g., learning from life experiences, skills, and knowledge) play an essential part in creating an identity for the individual (Almaraz-Espinoza & Grider, 2021; Bisaz et al., 2014).

Many memory-related illnesses prove that it is difficult or near impossible to function in daily life (Marder, 2017; Robertson, 2002; Schacter, 2013; Straube, 2012). For example, patients diagnosed with dementia become dependent on their family members or caregivers for showering, cooking and transporting; potentially affecting their self-esteem and confidence (Desai et al., 2004). Nevertheless, it is frequently observed that a perfectly healthy person forgets why they entered the room (Ossher et al., 2013; Schacter, 2013) or cannot remember information even though it feels like it is present in memory (Schacter, 2013). That indicates that memory is not always perfect and is prone to several different kinds of memory errors and distortions (Giovanello et al., 2010; Schacter & Slotnick, 2004; Schacter, 2013; Straube, 2012). This thesis will focus on one particular kind: False Memory (Newman & Lindsay, 2009).

False Memory

False memory refers to memories that never happened but are nonetheless believed to have happened (Newman & Lindsay, 2009; Nichols & Loftus, 2019). There has been extensive research on this topic (Brainerd et al., 2008; Howe et al., 2018; Otgaar et al., 2017)

since studies suggest that false memories can impact lives negatively. For example, falsely remembering details of a traumatic event has led to innocent individuals being questioned and punished (Henkel, 2014; Wang et al., 2018). On the contrary, some researchers favour the idea that false memories are associated with positive consequences (Otgaar et al., 2015); that is, they can enhance problem-solving skills (Howe et al., 2013). This occurs through the spread of activation caused by the false memory (Howe & Garner, 2018). There are several different mechanisms and brain regions responsible for the formation of false memories (Dodson & Schacter, 2002).

Potential Mechanisms Behind False Memories

Memory formation consists of three stages (encoding, storing and retrieving) (Hommel et al., 2011), and false memories can arise at any of these stages (Straube, 2012). During encoding, people perceive information from the environment and convert it to a mental representation (Campos et al., 2010; Stern & Alberini, 2013). Feature binding also occurs during this stage, where features of a particular memory or an event are combined to form one detailed representation of that memory (Dodson & Schacter, 2002). For example, the size, location of an object, and semantic features of an event are combined for a detailed representation (Bharti et al., 2020; Mammarella & Fairfield, 2008). Insufficient feature binding would suggest that the person only remembers some memory details. Consequently, it would imply that the person does not have all the specific information required to remember that memory accurately, causing an increased susceptibility to false memories (Bharti et al., 2020; Dodson & Schacter, 2002; Mammarella & Fairfield, 2008).

Insufficient feature binding could also lead to problematic source monitoring. This means that the person is unable to recollect the source of memory, for instance, a person who engaged in incomplete feature binding will recall a fact they studied but will be unable to recall where they studied it from (i.e., the source of that fact) (Devitt & Schacter, 2016;

Dodson & Schacter, 2002). Attention during this stage is required so that successful feature binding can occur. If the person was not attentive during encoding, they would be less likely to encode all the distinct information related to that memory. Therefore, they will not be able to form one detailed representation of the memory (Dodson & Schacter, 2002).

Furthermore, it is essential not to let the bound features of one memory combine with the features of other memories. Therefore, pattern separation is required during encoding (Dodson & Schacter, 2002). Pattern separation allows people to distinguish between two closely related events; that is, it allows people to tell apart overlapping features of different memories (Devitt, Tippett, et al., 2016; Yassa et al., 2011). Errors in pattern separation would lead to individuals relying on gist-based information (i.e., overall) rather than specific information. Consequently, this means a higher chance of producing false memories regarding that event (Dodson & Schacter, 2002).

One of the primary brain regions required for feature binding, patten separation, and source monitoring is the Medial Temporal Lobe (MTL) (Dodson & Schacter, 2002). The MTL comprises the hippocampus, entorhinal cortex, perirhinal cortex, and parahippocampal cortex (Dodson & Schacter, 2002; Mitchell & Johnson, 2009; Squire et al., 2004). The components of the MTL play an important part in memory. For example, the hippocampus is deeply involved in memory formation and learning (Bettio et al., 2017; Yassa & Stark, 2011) and is required for feature binding and pattern separation (Jeye et al., 2017). Several different brain regions send their outputs to the hippocampus, allowing it to join different inputs (e.g., vision input, audition input, and inputs from temporal and parietal lobes). The CA1 and CA3 fields within the hippocampus and the interaction between the hippocampus and its neighbouring cortex allow the information to travel and bind to different memory features (e.g., colour, semantic and other location features) (Jeye et al., 2017; Olsen et al., 2012; Ranganath, 2010). As discussed above, this process could lead to false memories if

insufficient levels of feature binding occur. Furthermore, the hippocampus also contributes to source monitoring (Mitchell & Johnson, 2009). Memory will be accurate when the hippocampus combines the memory information with the right context. However, false memories will be produced when the hippocampus binds the information to the wrong context (Jeye et al., 2017). Using the same example as above, the hippocampus lets the person remember the fact they studied and where they studied it from (Dodson & Schacter, 2002). The ventrolateral prefrontal cortex (PFC) supports the attention and encoding of specific memory information (Dodson & Schacter, 2002; Jeye et al., 2017). Furthermore, the dorsolateral PFC is also required for combining different memory features and, therefore, helping with source monitoring (Dodson & Schacter, 2002). The hippocampus and the PFC work together to allow memory formation (Preston & Eichenbaum, 2013) and therefore changes in these regions can potentially lead to false memories (Devitt & Schacter, 2016).

In the second stage, unstable memory is converted and stored as a long-term memory known as consolidation. This process consists of molecular, structural, and cellular changes (Campos et al., 2010; Squire et al., 2015); nevertheless, consolidated memories can change when they are reactivated. When consolidated memory is recalled, they become sensitive and labile before getting consolidated again, known as reconsolidation. This labile phase between consolidation and reconsolidation can give rise to false memories (Agren, 2014; Alberini & LeDoux, 2013; Lee, 2009). To elaborate, memories can be easily updated, weakened or intensified during this phase, leading to higher chances of false memories (Agren, 2014). In addition, the type of trace could also cause false memories. It has been suggested that there are two kinds of memory traces, namely, verbatim traces and gist traces. The former is associated with specific information about the event, whereas the latter is linked with general information (Brainerd & Reyna, 1990; Straube, 2012). It has been found that the verbatim trace is available for a short period in comparison to the gist trace (i.e., a person is more

likely to use the gist trace after a month of that memory) (Reyna et al., 2016). Thus, the details of a memory tend to fade over time, causing increased reliance on gist trace, therefore, increased susceptibility to false memories (Abadie & Camos, 2019; Reyna & Brainerd, 1995). Research suggests that the hippocampus is first required to consolidate memories; however, reliance on the hippocampus decreases with time. After some time, parts of the neocortex become more critical for memory storage than the hippocampus (Squire et al., 2015).

The final stage is retrieval, where memory is recalled or recognised. It requires the mnemonics to be activated to retrieve the stored information (Campos et al., 2010). Cues during retrieval could match several different memories, for example, if the cue was eating food with a friend, then there will be several different memories associated with that. In order to remember the correct memory, people need to engage in a process called focusing (Dodson & Schacter, 2002), however, if there is a poor focus, that could lead to false memories. Furthermore, the source monitoring problem needs to be solved to retrieve the memory correctly. To explain, the person must ensure that the recalled memory is not an imagination or product of daydreaming but is the correct event experienced earlier (Dodson & Schacter, 2002). Similar to encoding and storing, the use of gist trace during retrieval could also lead to the formation of false memories. That is, using overall rather than item-specific information while retrieving increases the chances of false memories (Abadie & Camos, 2019). Besides, the use of verbatim (item-specific) trace leads to correct rejections of false memories compared to the gist trace (Reyna et al., 2016). Similar to the first two stages in memory formation, the MTL and the PFC play an important part in memory retrieval (Dodson & Schacter, 2002). They are required to work together to retrieve memory accurately (Preston & Eichenbaum, 2013).

Tasks Used to Assess False Memory

There are several different tasks used to assess episodic false memories in cognitive experiments. Episodic memory allows people to remember past events (including, what, where, when, and with whom it occurred) (Devitt, Monk-Fromont, et al., 2016). Two of the most common tasks used are the Deese-Roediger-McDermott (DRM) task and the misinformation task. In The DRM task, participants are given different lists of semantically related words to study, for example, *bed, pillow*, and *awake*. These words are thematically linked with other words known as critical lures (e.g., *sleep*), which are not presented in the original word lists. However, when recognising or recalling the word lists, participants tend to report that they studied the critical lure in the original list (Deese, 1959; Roediger & McDermott, 1995). In the misinformation task, participants watch an event, for example, watching a video or slides about a thief stealing a hammer. This is followed by receiving misinformation about that event (e.g., showing a photo of the thief stealing a screwdriver instead). Participants are then tested on what happened in the original event. For example, what did the thief steal in the original event? Participants tend to say screwdriver instead of the hammer (Loftus & Hoffman, 1989; Nichols & Loftus, 2019; Zhu et al., 2013).

False memories in the DRM task could be explained through Activation Monitoring Theory (Roediger et al., 2001). The Activation Monitoring Theory states that memory has several interconnected nodes; therefore, activating one node activates the interrelated nodes. For instance, when someone encodes words such as *bed* and *pillow*, the word *sleep* will also be activated. However, this theory also suggests that when it is time to retrieve information, that information undergoes checking. Hence, false memories should be eliminated through the monitoring process (Otgaar et al., 2017). Nichols and Loftus (2019) suggested that errors in monitoring while encoding lead to confusion at retrieval, and therefore participants report having studied the critical lure. The Fuzzy Trace Theory (FTT) could also explain false

memories in this task. This theory focuses on memory traces; that is, when something is perceived, memory is stored as both verbatim and gist traces (Otgaar et al., 2017). The verbatim trace is a more detailed memory trace than the gist trace, which is more of an overall context of memory without the details (Nichols & Loftus, 2019; Otgaar et al., 2017). As per this theory, participants make false memory errors when they rely more on gist trace than the verbatim trace (Otgaar et al., 2017).

Research shows that there is no correlation between different tasks (namely, DRM, misinformation and an imagination task) (Nichols & Loftus, 2019). Participants who show false memories in the DRM task would not necessarily show false memories in the misinformation task (Nichols & Loftus, 2019; Ost et al., 2013). Even though the purpose of these tasks is to study false memories, they may both be affected by different mechanisms (Calvillo & Parong, 2016; Falzarano & Siedlecki, 2019; Nichols & Loftus, 2019). This may be because false memories in the DRM task result from internal factors (also known as spontaneous false memories). In contrast, false memories in the misinformation task are due to external factors (suggestion induced) (Nichols & Loftus, 2019). To be specific, false memories in the DRM task result from the influence of semantic memory (an internal process which refers to memory about knowledge of the world, facts, and words) (Yee et al., 2013), in comparison, false memories in the misinformation task, occur due to an external factor (namely, the misinformation given) (Nichols & Loftus, 2019). However, one commonality between the two tasks could be that they result from faulty source monitoring. That is, individuals fail to monitor the source of the initially perceived information during retrieval (Devitt & Schacter, 2016; Ost et al., 2013). Therefore, this explains why the lure word (e.g., *sleep*) was falsely remembered as a studied word, and the misinformation in the misinformation task was wrongly assigned to the initial event (Ost et al., 2013). Many other factors (e.g., lack of sleep and the amount of new information coming in) (Straube, 2012)

contribute to false memories. In addition, studies have also focused on individual factors such as age that could lead to the formation of false memories (Nichols & Loftus, 2019).

Age and Memory

Ageing is linked with physical and cognitive declines (Bettio et al., 2017). There is ample evidence indicating negative changes in memory with older age (i.e., adults aged 65 years and older). Specifically, the amount of information remembered decreases with age (Rosa & Gutchess, 2013), and memory errors such as false memories tend to increase with age (Jacoby & Rhodes, 2006; Rosa & Gutchess, 2013). The reason for this has been explained from a neurological perspective. Studies focusing on the neurological perspective have suggested changes in the MTL, PFC, and the hippocampus with age (Devitt & Schacter, 2016; Rosa & Gutchess, 2013), as a result, these declines are linked with poor performance on memory tasks (Jacoby & Rhodes, 2006; Rosa & Gutchess, 2013).

There are changes in the volume of the hippocampus with age. These changes could be associated with a decrease in neuron levels and the production of neurons with age (the neurons help the hippocampus function effectively) (Cameron et al., 1993). Furthermore, there are structural changes in the CA1 and CA3 parts of the hippocampus. These changes are also linked with the poor functioning of the hippocampus (Cameron et al., 1993; Dodson & Schacter, 2002). Additionally, age-related changes in some metabolism-related hormones also lead to adverse changes in the functioning of the hippocampus. These factors work together and lead to negative changes in cognitive abilities with age (Bettio et al., 2017). Studies have also found a decrease in the activation and volume of the PFC with age (Solbakk et al., 2008; Zanto & Gazzaley, 2019). Additionally, there is a reduction in the connectivity between the PFC and the MTL. This connection is crucial for correctly encoding, storing, and retrieving information (Andrews-Hanna et al., 2007). As such, changes

in the MTL, hippocampus, and PFC are associated with an increase in false memories with age.

As explained above, feature binding refers to blending different characteristics to generate a specific memory (Bharti et al., 2020; Dodson & Schacter, 2002; Mammarella & Fairfield, 2008). With age, the structural and functional changes to the hippocampus and the PFC could cause older adults to engage in unsuccessful or insufficient binding of memory features; thus, older adults will only remember some details associated with that memory (Devitt & Schacter, 2016; Dodson & Schacter, 2002). Consequently, they will be more likely to make false memory errors (Dodson & Schacter, 2002).

Furthermore, older adults might insufficiently bind the memory source to its context, causing problems with source monitoring. As explained prior, feature integration is required for successful source monitoring; however, older adults might engage in problematic feature binding, implying that they will not be able to bind the information with its context successfully (Devitt & Schacter, 2016; Dodson & Schacter, 2002; Wylie et al., 2014). Therefore older adults will be making more source monitoring errors. It is considered that a reduction in attention during encoding as well as retrieval also plays a role in weak source monitoring and hence increased susceptibility to false memories. Importantly, the dorsolateral PFC is associated with prolonging attention, and there is an age-related decline in this area (Castel & Craik, 2003; Tisserand et al., 2004), leading to an increased likelihood of false memories.

Another factor that could account for an age-related increase in false memories is the type of memory trace used. Older adults rely more on gist trace (i.e., the overall information) rather than on verbatim trace (i.e., the detailed information) (Brainerd & Reyna, 1990; Devitt & Schacter, 2016). That is, they tend to encode general information about the event rather than focusing on item-specific details (Brainerd & Reyna, 2002; Kensinger & Schacter,

1999). This increased reliance on gist trace causes older adults to perform poorly on tasks when they are required to remember distinct details (such as in the DRM and the misinformation task). Ergo, older adults tend to make more false memory errors than younger adults (Devitt & Schacter, 2016). Age-related change in the hippocampal activity is responsible for changes in memory traces (Paige et al., 2016; Wilson et al., 2006). To elaborate, the hippocampus might not engage in sufficient feature binding, thus, causing difficulties in weaving the details of memory together and producing a specific memory trace (Paige et al., 2016). The lack of details will result in a reliance on gist rather than a verbatim trace.

As evident from the above paragraphs, increased memory errors have been explained from a neurological perspective. However, recent research has focused on socioemotional factors that could also cause an age-related increase in false memories. There have been conflicting results; nonetheless, it is vital to consider socioemotional factors in addition to neurological factors to get a better understanding of how and why older adults tend to make more memory errors. This will also enable health practitioners to develop better solutions for older adults with memory problems (Scholl & Sabat, 2008). One of those socioemotional factors widely studied is stereotype threat.

Stereotype Threat

Stereotype threat refers to the idea that inducing negative beliefs about a certain group results in the group conforming to those negative stereotypes and performing poorly (Mazerolle et al., 2021; Steele & Aronson, 1995). Stereotype threat challenges the self-integrity of individuals. Self-integrity defines that one is an intelligent and appreciated member of the society who can adapt to changing situations and succeed. With the threat, the sense of self-integrity is affected, and the person is made to believe otherwise (Schmader et al., 2008). For example, black American students perform poorly on tests when they have to

disclose their ethnicity prior to the test (Steele & Aronson, 1995). Furthermore, women perform poorly on a maths test when they are told that men are better at maths than women (Steele, 1997; Quinn & Spencer, 2001). Stereotype threats also affect behaviour in older adults. For instance, Lambert et al. (2016) found that stereotyped older adults drove poorly compared to older adults in the control condition (see also Brelet et al., 2016; Joanisse et al., 2013).

Stereotype threat has real-world implications. Patients with Alzheimer's may perform poorly on neuropsychological tests due to the fear of conforming to negative stereotypes. The threat could lead to an increase in anxiety which affects their performance. Importantly, threat-related changes in performance influence the interventions required to help the individual in their day-to-day life (Scholl & Sabat, 2008). For example, someone with Alzheimer's may perform poorly due to the threat. However, for a health practitioner (who is just focused on the neurological side), this could imply that they need more support (Scholl & Sabat, 2008) or medical help (Riley et al., 2014). Thus, stereotype threat could lead to Alzheimer's patients getting unnecessary medical treatment (Scholl & Sabat, 2008).

Stereotype threat increases vigilance and causes individuals to take a preventative approach as they do not want to conform to those negative stereotypes (Mazerolle et al., 2021). However, they still end up conforming to the negative stereotypes (Smith et al., 2017), this could be why women performed poorly on a maths test and older adults on a driving test (Lambert et al., 2016; Quinn & Spencer, 2001). When the stereotype threat is activated, it interferes and affects the mental process. To elaborate, working memory is temporary and can be stored for a varied amount of time depending on the importance of the information (Campos et al., 2010). When the threat is activated, it reduces the working memory capacity available for the task (Schmader et al., 2008); consequently, resources that otherwise would have been involved in helping the individual perform better end up being consumed by the

threat. The Executive Resource Hypothesis and the Regulatory Focus Hypothesis can be used to explain the effects of stereotype threat on task performance (Barber & Mather, 2013; Fourquet et al., 2020; Mazerolle et al., 2021).

The Executive Resource Hypothesis highlights that individuals under threat have limited resources left for the task at hand. This is a consequence of using self-monitoring and emotion regulation strategies to fight negative thoughts caused by the threat. In addition, people also employ physiological responses such as an increase in blood pressure, stress levels, and anxiety to fight the threat (Fourquet et al., 2020; Johns et al., 2008; Mazerolle et al., 2021). The performance is further affected by the difficulty of the task (Mazerolle et al., 2021). This is because a free recall task, where participants have to remember the words they studied freely, would be more challenging and would require more resources in comparison to a recognition test, in which they are given a retrieval cue (e.g., did you study this word before?) (Freund & Baltes, 1998; Mazerolle et al., 2021).

The Regulatory Focus Hypothesis highlights that performance tends to improve when the internal regulatory state of the person matches the rewards of a task (Barber & Mather, 2013). To elaborate, participants who care about the presence or absence of rewards tend to have a promotion focus and will perform better on a task that is also focused on the presence of rewards. Whereas participants focused on the presence or absence of losses tend to have a prevention focus, and will perform better when they need to avoid being penalised in a task (Crowe & Higgins, 1997; Fourquet et al., 2020; Mazerolle et al., 2021). The internal focus (promotion or prevention) changes depending on the situation. For example, it will be considered a loss if an individual under stereotype threat conforms to the negative stereotype. Therefore, individuals adopt a prevention focus to avoid that loss (Seibt & Förster, 2004) and will perform better on a task that focuses on the absence or presence of losses (Barber & Mather, 2013). Nevertheless, false memory tasks are usually promotion-focused (Mazerolle

et al., 2021) and introducing stereotype threat changes the internal state of the participant, but not the task. Therefore, there is a mismatch in the focus of the task (promotion focus) and the participant's internal state (prevention focus), leading to an increase in false memory errors (Mazerolle et al., 2021; Seibt & Förster, 2004).

Stereotype Threat, Ageing, and False Memory

Several studies have focused on the effects of age-based stereotype threat on memory and have found that the performance of the stereotyped participants decline in comparison to the control participants (Bouazzaoui et al., 2020; Chasteen et al., 2005; Hess et al., 2003; Levy et al., 2012; Rahhal et al., 2001). However, limited research has examined how threat affects false memory, and most studies have found contradictory results. This could be attributed to the methodological differences amongst those studies. That is, it could depend on whether the study used a prevention or a promotion focus, focused on moderating factors (such as education, employment status and importance placed on memory), and whether they used a free-recall or a recognition test for retrieval. These factors are discussed in more detail below.

Prevention and Promotion Focus

Two studies focusing on how ageist stereotypes affect false memory used a DRM task with a subsequent recognition test to investigate false memories (Thomas & Dubois, 2011; Wong & Gallo, 2016). These studies had participants read a paragraph about negative changes in memory with age to elicit stereotype threat, with a paragraph about language processing as a control. This is a common way to elicit stereotype threats and has been widely used by researchers (Spencer et al., 1999; Shapiro & Neuberg, 2007; Smith et al., 2017; Thomas & Dubois, 2011; Wong & Gallo, 2016). Despite these similarities in design, these two studies found opposing results. This could be accounted for by apparently a minor difference in methodology.

Thomas and Dubois (2011) informed participants in the stereotype threat condition that the study was about memory and were given a paragraph about the age-related declines in memory. In contrast, participants in the control condition were notified that the study was about language ability and were given a paragraph on language. They found that participants in the stereotype threat condition made more memory errors compared to those in the control condition. Thomas and Dubois (2011) suggested that this was due to limited cognitive resources available to perform well on the task in the stereotype threat condition (i.e., the Executive Resource Hypothesis).

In comparison, all participants in Wong and Gallo (2016) were additionally warned about the DRM task. Participants were told that in the recognition test, there would be similar words to the studied words; however, they were not presented in any of the word lists studied earlier and were told to be careful. Wong and Gallo (2016) found that participants in the stereotype threat group made fewer memory errors compared to the control group. They supported their conclusion through the Regulatory Focus Hypothesis, suggesting that participants under stereotype threat performed better than the control group because there was a match between the internal regulatory state and the task. Thomas and Dubois (2011) found an increase in false memories for the stereotyped group, while Wong and Gallo (2016) found a decrease in false memories for the stereotyped group.

To elaborate, in Thomas and Dubois (2011), stereotyped participants were given a negative paragraph about age-related changes to read, thus inducing a prevention focus while the task remained promotion-focused. This means that the internal state of stereotyped participants did not align with the focus of the task (i.e., promotion focus) (Mazerrolle et al., 2021); whereas the internal state of the control participants (promotion focus) who read a paragraph on language did align with the focus of the task (promotion focus). Hence,

supporting why participants in the stereotyped group performed poorly compared to the control group.

Similarly, participants in the stereotyped group in Wong and Gallo (2016) read a negative paragraph to induce a prevention focus; however, they were also given a warning about the DRM task, which was enough to change the focus of the task from promotion to prevention (Mazerrolle et al., 2021). Therefore, participants in the stereotype threat group performed better as their internal regulatory state (prevention focus) matched the task's focus (prevention focus). While control participants read a neutral paragraph on language and thus remained promotion-focused, however, participated in a task that was prevention-focused (due to the warning). There was no match in internal regulatory state of the control participants and the task, which could be why they performed poorly compared to the stereotyped participants.

These studies imply that it is important to focus on not just one but both of these hypotheses when designing a task as well as when analysing the findings. It seems clear that methodological differences (e.g., warning versus not warning) change the findings and could lead to conflicting outcomes. Smith et al. (2017) wanted to combine the above two studies and investigate whether they could yield consistent results. To do that, they divided their participants into four groups. The first group received no warning about the deceiving nature of the DRM task but read a stereotype threat passage. The second group was warned about the task and were not given the stereotype threat passage to read. The third group was warned and also read a stereotype threat passage; however, the fourth group had no warning and no stereotype threat passage. Consistent with Thomas and Dubois (2011), Smith et al. (2017) found that participants under stereotype threat made more false memories compared to those who were not given the stereotype threat passage. However, inconsistent with Wong and Gallo (2016), there was no interaction between the stereotype threat and the warning.

Important to note that Smith et al. (2017) used the same stereotype threat passage as Thomas and Dubois (2011), in contrast, Wong and Gallo (2016) used a more complex paragraph. This could also affect the findings because there are cognitive declines with age that influence reading abilities (Hou et al., 2017). Therefore, reading a more complicated paragraph in Wong and Gallo (2016) could make it harder for older adults to comprehend the stereotype threat paragraph.

Education Status, Retirement, and Importance

It is also important to focus on other moderating factors, such as level of education, retirement status and importance placed on memory, in addition to the methodological differences that researchers might be missing. For example, highly educated people could experience stereotype threats differently from those who are not highly educated (Hess et al., 2009; Smith et al., 2017). This is because highly educated people value their cognitive abilities more than less educated people. Therefore, when encountering stereotype threat about age-related declines in cognitive abilities, highly educated people are more likely to internalise that threat compared to the less educated. Consequently, more cognitive resources will be used by the highly educated people to prevent conforming to the negative stereotype. This leads to the stereotyped participants with high education making more memory errors in free recall and recognition tests than the stereotyped adults with lower education (Gilet et al., 2021; Hess et al., 2009; Smith et al., 2017).

Similarly, retired adults are more likely to be affected by negative stereotypes than employed adults. Retired adults consider these negative age-based stereotypes more plausible due to changes in ways of living life. When someone retires, they adopt a different lifestyle which indicates that the individual is entering old age; therefore, they tend to believe in such negative stereotypes (Smith et al., 2017).

Finally, the importance given to memory can also make older adults susceptible to false memory errors (Hess et al., 2003; Mazerolle et al., 2021; Ryan & Campbell, 2021). Older adults are more likely to internalise threats when they place greater importance on memory for their day-to-day life (Hess et al., 2003; Ryan & Campbell, 2021). Due to an increased internalising, there is an increased worry about doing well and feeling good about one's cognitive capacities. However, this blocks cognitive resources that are available for the task. Hence, older adults who place greater importance on their memory end up performing poorly than those who do not place greater importance on memory under stereotype threat (Ryan & Campbell, 2021). Thomas and Dubois' (2011) and Wong and Gallo's (2016) studies did not account for any of the above moderating factors, potentially explaining the conflicting findings of these two studies. Furthermore, methodological difference in the retrieval task could also influence the findings.

The Difficulty of the Task

In recognition tasks, participants are given words and asked whether it is new or old (Smith et al., 2017; Thomas & Dubois, 2011; Wong & Gallo, 2016). In free-recall tasks, participants are expected to write as many words as they remember from the memory task without any cues (Kang & Chasteen, 2009). Free recall tasks are typically more challenging than recognition tasks. The Selective Engagement Theory states that with age, it becomes difficult to engage in cognitive tasks, exhausting cognitive resources and causing tiredness (Freund & Baltes, 1998). As a result, older adults become more careful in using their limited resources. In addition, there is a decrease in blood pressure as the difficulty of the task increases, indicating that older adults stop engaging in the task as the difficulty of the task increases. Since older adults stop engaging, they are likely to perform poorly on the free-recall task in comparison to the recognition task (even without the presence of a stereotype

threat) (Ryan & Campbell, 2021). This highlights the need for studies to investigate the effects of age-based stereotype threat on freely recalled false memories.

To note, none of the above studies in this literature review used free-recall for a false memory test (Smith et al., 2017; Thomas & Dubois, 2011; Wong & Gallo, 2016). Nevertheless, Kang and Chasteen (2009) found that age-based stereotype threat affects memory for the free recall but not the recognition task. This is because free recall requires increased cognitive resources, but those resources are already occupied in handling the negative stereotype, thereby affecting performance of the stereotyped participants to a greater extent compared to the control participants (Mazerolle et al., 2021).

Beyond the DRM Paradigm

Thomas et al. (2020) conducted a study using the misinformation paradigm to assess eyewitness memory. Participants saw a crime video, followed by a written synopsis in which they were either given information that matched with the original event or did not match. Participants were then given either a control paragraph or a stereotype threat paragraph about age-related changes in memory. In a low cognitive demand task, participants answered questions either from the original video or the synopsis. Here, the authors expected that the stereotype threat group should not be negatively affected because the cognitive demands of the task were low. However, participants in the stereotype threat group made more errors in comparison to participants in the control group. In the second experiment, participants answered questions based only on the original video, making the task more cognitively demanding. Here, there was no difference in performance between the two groups. If the Executive Hypothesis was correct, then in the first experiment, stereotyped participants should not have performed worse than the control participants, as the task was less cognitively demanding (Thomas et al., 2020). However, stereotyped participants performed worse, supporting the Regulatory Focus Hypothesis, which shows that participants had a

prevention focus induced due to the threat while the task was promotion-focused (Mazerolle et al., 2021). Likewise, the control participants had a promotion focus while the task was also promotion-focused, aligning the persons' internal state with the task (Barber & Mather, 2013; Thomas et al., 2020). Nevertheless, the Executive Resource Hypothesis seem to apply to a difficult task as in the second experiment. There was no difference observed between the two groups indicating that the task was difficult for all participants; thus, they might have decided to disengage from it (Ryan & Campbell, 2021).

To summarise this literature review, it is important to consider methodological differences when investigating the impacts of age-based stereotype threat on false memories. While designing a study, it is essential to consider whether the design changes the focus of the task from promotion to prevention or vice versa. One must not forget individual differences that could also contribute to these findings (e.g., level of education, employment status and importance placed on memory). Furthermore, it is crucial to keep in mind that free-recall and recognition tasks can have different effects of stereotype threat on memory performance. Therefore, when researchers are designing and analysing their findings, they must pay attention to these factors and examine whether these factors are the reason for inconsistent findings. Previous studies have differed in a number of supposedly minor ways that nevertheless drastically altered the results, which could explain why there have been so many mixed findings. However, no study has looked at all these factors at once, and this is what this study is going to do.

The stereotype threat literature has mainly focused on the DRM task to study the effects on false memories (Smith et al., 2017; Thomas & Dubois, 2011; Wong & Gallo, 2016). However, one type of memory error we encounter in our daily life is memory conjunction errors. For example, while calling someone, a person falsely dials a few digits of their home phone and a few of their work phone. Here, this individual integrated two distinct

memory traces to form one memory that does not exist (Reinitz et al., 1992). The memory conjunction task was introduced by Underwood and Zimmerman (1973). Participants usually read conjunction words in the memory conjunction task, such as *snowman* and *sandcastle*. They are then given a recognition memory test, in which they are shown new words (e.g., *passionfruit*), old words (e.g., *sandcastle*), recombined/conjunction words (e.g., *snow castle*), or feature words (e.g., *snowball*). Half of the old word is combined with a new half in feature words. Participants are expected to choose whether the word presented was new or old and often claim that conjunction words were initially presented. Researchers have also used faces (Jones, 2005), events (Odegard & Lampinen, 2004), and sentences (Reinitz et al., 1992) instead of words to deliver this task.

It is thought that false memories produced in this task could be due to lax binding during encoding (Giovanello et al., 2010; Leding & Lampinen, 2009). Insufficient feature binding is also linked with source monitoring errors (as explained above) (Dodson & Schacter, 2002). Furthermore, false memories are likely to be produced due to the similarity between the studied word and the conjunction word. That is, conjunction words seem familiar because both parts of the conjunction word have been viewed before, which causes the individual to respond that they have previously seen the conjunction word (Giovanello et al., 2010; Leding & Lampinen, 2009).

Filling in the Gap and Aim of the Present Study

To our knowledge, no prior study has used the memory conjunction task to understand the effects of age-based stereotype threat on false memories. Thus, the current study will be the first to do that. Since memory conjunction errors are likely to be found in daily life (Reinitz et al., 1992), it makes it important to illuminate this task and investigate if it could be used to understand the effects of age-based stereotype threat on performance. Since studies (Nichols & Loftus, 2019; Ost et al., 2013) found no correlation among the

previous tasks (e.g., the DRM and the misinformation task) and due to the inconsistent findings, it is important to look at as many tasks as possible, and that is what the current study aims to do. Experiment 1 was designed to explore methodological differences by using a free recall and a recognition test to understand if the type of retrieval task affects the results. Furthermore, we will also analyse if moderating factors such as education, employment status, and importance placed on memory influence the results. For Experiment 2, we aimed to test whether the focus of the task (i.e., promotion focus or prevention focus) is able to influence the performance of stereotyped participants using a DRM task. This will help explain the different findings from Thomas and Dubois' (2011) and Wong and Gallo's (2016) studies. That is, if it does, then the inconsistencies between those two studies could be accounted for by the task design.

Our overall aim is to extend the limited stereotype threat literature by examining how methodological differences (e.g., the difference in focus, retrieval task and other moderating factors) could contribute to mixed findings. Our secondary aim for Experiment 1 is to extend the literature to a new type of memory error (memory conjunction errors).

Therefore, our main research questions for both experiments are:

1) Do age-based stereotype threats influence performance on the memory conjunction error task? Moreover, do moderating factors (i.e., education, retirement and importance placed on memory) influence this relationship?
2) Is there any difference in performance between the control and the stereotyped participants for the free recall versus the recognition task?
3) Is there any difference in performance when using the memory conjunction task in comparison to the DRM task?
4) Do stereotyped participants perform better under prevention focus compared to control participants under prevention focus on a DRM task?

We hypothesise that participants in the stereotype threat group will show an increase in false memories compared to the control group (Smith et al., 2017; Thomas & Dubois, 2011). This is because our task is promotion focused (Mazerolle et al., 2021), while stereotyped participants will have a prevention focus (Barber & Mather, 2013; Mazerolle et al., 2021; Seibt & Förster, 2004). We further think that participants in the stereotype threat condition should perform poorly in the free recall task compared to the recognition task. This is because the free recall task is more cognitively demanding, and the threat usually consume those cognitive resources to fight negative emotions (Mazerolle et al., 2021). In addition, we hypothesise that participants that are highly educated, retired and place high importance on memory will show a greater influence of stereotype threat compared to less educated, still working adults that place medium importance on memory (Ryan & Campbell, 2021; Smith et al., 2017). For our second experiment, we hypothesise that when participants under threat are given a task that is also prevention-focused, they will perform better and make fewer false memory errors than control participants in the same focus. This is because there will be a match between the person's internal state and the task (Barber & Mather, 2013; Wong & Gallo, 2016).

Methods - Experiment 1

Participants

Data were collected from 60 participants aged 65 years or more (*M* = 69.67 years, *SD* = 5.00, 33 female). This sample size was based on previous studies, which used 64 and 84 older adults (Thomas & Dubois, 2011; Wong & Gallo, 2016). We decided to use 60 participants as that number was close to previous studies; furthermore, G*power, a statistical software, was used to check whether that was a sufficient sample. Ethics approval was gained from the ALPSS University of Waikato Human Ethics Committee (approval number FS2020-45). Twenty participants were recruited from the Waikato area by putting flyers on social media and from community groups around Hamilton (refer to Appendix A.1). These participants were able to enter a draw to win one of three $20 Warehouse vouchers once they submitted the experiment. Due to restrictions resulting from COVID-19, we were not able to reach as many participants in the community as we had aimed. Consequently, 40 participants were recruited from MTurk and CloudResearch, online platforms to recruit participants. These participants were from the U.S. and were paid U.S.$ 1.50 (N.Z.$ 2.10) for taking part in our study.

Materials

The questionnaire in current experiment was administered with Qualtrics^{XM}, an online platform used to design and distribute surveys. It allows participants to access the survey in their own home using a device of their choice (e.g., phone, iPad, laptop, or computer). Participants were able to complete the survey whenever they had some free time. However, they were expected to complete it in one sitting. The stimuli in this experiment were made up of two word lists consisting of 32 conjunction words each (refer to Appendix A.3). The paragraph manipulation was adapted from Wong and Gallo (2016) for control and stereotype threat groups. The stereotype threat paragraph was about negative changes in memory with

age compared to young adults, whereas the control paragraph was about language processing (refer to Appendix A.2 for the paragraphs).

Procedure

In this experiment, participants were tested on a memory conjunction task, and it took approximately 10 minutes to complete. Participants were required to have a stable internet connection, a device, and a quiet space where they could complete the task. Before the experiment commenced, participants were given a consent form, where they were informed about the study's aim, benefits, and methodology.

Once participants had agreed to the consent form, they were given instructions about the memory conjunction task. Participants were told that they would study two word lists, and each word would be shown for two seconds. Participants were asked to remember as many words as possible for a memory test. Participants were then shown 32 conjunction words one at a time in each list (64 words total). Two separate lists were provided because they helped break down the task for participants.

Of the 32 words in each studied list, eight were designated old words (i.e., the words presented in the recognition test later), 16 were conjunction parent words, and eight were feature parent words (note that participants were unaware of these categorisations). The conjunction parent words were combined to form a conjunction lure in the recognition test. For example, *watchtower* and *sheepdog* were studied in one of the lists, and in the recognition task, participants were asked if they saw the word *watchdog* in the list. Important to note that parent conjunction words in each list were presented in close proximity. For feature parent words, half of the word was presented with a new second half as a feature lure in the recognition test. For example, if *steamroller* was studied, participants were then asked if they saw the word *steamboat* in the recognition task.

After studying two lists, participants were told they would be given a paragraph to read and asked to answer a question about that paragraph. Qualtrics^{XM} randomly presented either the language paragraph (control) or the stereotype threat paragraph. Participants were given 90 seconds to read the paragraph, and they could not go to the next question until those 90 seconds were over. This ensured that all participants took the same time to read the paragraph and answer the question. Participants who read the stereotype threat paragraph were asked: "whether adults over the age of 60 report more memory problems than younger adults", whereas participants who read the language paragraph were given three statements about language and were asked to state which of the following statements were correct. The purpose of the question was to check if participants were reading the paragraph or not.

Once participants read either the control or the stereotype threat paragraph, they were given a free recall and a recognition test. In the free recall test, all participants were asked to write as many words as they remembered from the two word lists studied previously. Participants were given 90 seconds to complete this task and could not move to the next question until the time was over. This was to encourage participants to think and not just go to the next question because the current question was difficult.

Immediately following the free recall, participants completed the recognition test. Participants were instructed that they would see one word at a time and be asked if this was a *new* word (i.e., a word not studied in the two word lists) or an *old* word (i.e., it was studied earlier). This decision was self-paced. The recognition test was presented after the free recall test to ensure that the recognition test was not helping participants with free recall. Participants viewed 64 words in total, 16 of those words were old words, and the remaining 48 were new words. Out of the 48 new words, 16 were conjunction lures, 16 were feature lures, and the remaining 16 were brand new.

Question about any clinical diagnoses was then presented. Participants were asked to select (more than one if applicable) if they were clinically diagnosed with any illnesses (such as Dementia, Stroke, Brain aneurysm, Mild cognitive impairment, Multiple sclerosis, and Parkinson's disease) presented in the list.

The next question asked participants to rate the importance of memory in their daily lives. Participants were able to drag a red circle and choose a number between "1" and "10" (i.e., "1" = being not important and "10" = being very important). Participants were then asked several demographic questions such as their age, gender, highest qualification, years of education, employment status, and first language. Finally, participants were asked whether their data could be used for analysis. Participants were able to choose "yes", indicating that they did not cheat or did not have any technical issues or "no" and were given several different options (e.g., "I wrote the words down, I had technical issues or other reasons") for why they chose that answer. Important to note that participants were assured that their answer would not affect their compensation in any way. After completing all questions, community participants were able to choose if they wanted to participate in the raffle; if so, they were asked to provide their email, which was not linked with their data in any way. The same procedure was used for participants from Mturk and Cloud research, except they were not able to enter the draw to win and instead were paid U.S. $1.50 (N.Z. $2.10).

A few additional requirements were set to ensure that we recruited quality and genuine participants from online platforms. First, the human intelligence task (HIT) approval rate was set to greater than 95%, and the number of HITS approved was set to greater than 50. Similarly, in Cloud research, criteria such as age above 65 and location from the U.S. were selected.

There were questions to make sure that participants were paying attention throughout the experiment. To ensure that participants read the instructions, they were asked questions

like “how many seconds will they see each word” after receiving instructions about the memory conjunction task. Additionally, participants were asked questions about the control and stereotype threat paragraphs to make sure they read and understood them. To catch participants that were untruthful about their age, participants were given six seconds to answer what year they were born. Any participants who failed two attention checks were filtered out of the experiment.

Data Processing

As mentioned above, participants were asked a question towards the end whether their data could be used and were given a few options (e.g., yes, or no because “I was distracted, there were technical problems”, and so on). Participants who responded “no” to using their data for any reason were excluded. In addition, participants with Dementia, Stroke, Brain aneurysm, Multiple sclerosis, and Parkinson’s disease were also excluded from the sample. Through preliminary analysis, five participants were excluded from the experiment. Those participants said “no” to using their data due to distraction or technical problems. In addition, one participant was excluded because they reported having been diagnosed with brain tumour. As such, 65 participants were recruited, and the data from 60 participants were used for the analyses.

To analyse, the proportion of hit rates and false memory errors made by each participant in the free recall and recognition tests were calculated. There were 28 participants in the control group and 32 participants in the stereotype threat group. Hit rates focused solely on old words and were calculated by dividing the correct recognition score on old words by the total number of old words (i.e., 16) to give a proportion correct. False recognition errors were divided into three groups: false alarms for conjunction words, feature words and new words. Proportions were calculated by dividing the number of errors in each category by the total number of words in each category (i.e., 16). In addition, the total

number of correct, intrusions, feature and conjunction words recalled were also calculated. Important to note, to analyse the free-recall task, we did not use proportions as in the recognition task but instead used raw scores. SPSS was used to run different statistical analyses to compare the performance of the stereotype threat and the control group. The means obtained from these analyses were plotted into a bar graph using Excel.

Altogether we did three different analyses, which are listed below.

1. A series of independent samples t-tests were carried out to investigate any differences in the hit rates and false recognitions (on conjunction, feature and new words) between the stereotype threat and the control group.
2. A series of independent samples t-tests were carried out to investigate whether there were any differences in the number of correct hits and errors (intrusions, conjunction and feature) freely recalled between the stereotype threat and the control group.
3. A series of 2 × 2 between-subjects Analysis of variance (ANOVAs) were carried out to investigate the effects of moderating factors such as education, employment status, and importance of memory on conjunction errors made in the recognition and free-recall tasks by the stereotyped and the control participants.

Results - Experiment 1

Recognition and Recall Errors

First, we compared the control and stereotype threat groups' hit rates and false recognition errors. The average hit rate for the control group was similar (M = 0.68, SD = 0.19) to the average of the stereotype threat group (M = 0.67, SD = 0.19). An independent samples t-test was used to investigate the difference between groups. The difference between means was found not to be statistically significant, $t(58) = 0.09$, $p = .930$, $d = 0.02$, 95% CI [-0.10, 0.11], indicating that there was no difference in correct recognition for old words between the two groups.

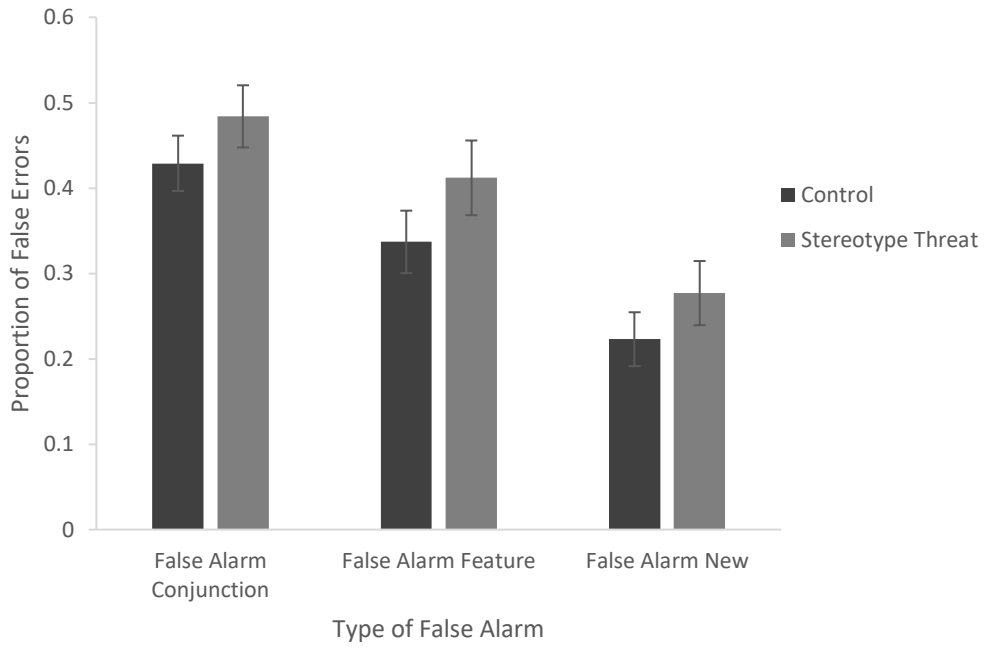
The independent samples t-test comparing false recognition errors (for all three categories) between the control and the stereotype threat group showed that participants in the two groups made a similar proportion of errors. Participants in the control group made a similar proportion of recognition errors for conjunction words (by choosing OLD instead of NEW for the conjunction words during the recognition test) (M = 0.43, SD = 0.17, i.e., an average of 6.88 conjunction errors per person) in comparison to the stereotype threat group (M = 0.48, SD = 0.21). This difference was not statistically significant, $t(58) = -1.13$, $p = .263$, $d = -0.29$, 95% CI [-0.15, 0.04].

Following this pattern, false recognition for feature words and new words were also similar for the control (M = 0.34, SD = 0.19 and M = 0.22, SD = 0.17 respectively), and the stereotype threat group (M = 0.41, SD = 0.25 and M = 0.28, SD = 0.21 respectively). These differences remain non-significant, $t(58) = -1.30$, $p = .200$, $d = -0.34$, 95% CI [-0.19, 0.04] for feature words, and $t(58) = -1.09$, $p = .282$, $d = -0.28$, 95% CI [-0.15, 0.05], for new words. The mean values are plotted in Figure 1 with standard error. One important thing to note from Figure 1 is that participants in both groups made more errors on conjunction words, then feature words and then new words. This pattern was expected and confirmed that the memory

conjunction task was working. A repeated-measures ANOVA was also conducted to investigate whether the paragraph type (control/stereotype threat) had any effect on the proportion of false memory errors. The three factors that went into the repeated measures ANOVA were the three false alarm types: conjunction, feature, and new. There was no significant main effect of paragraph on false memory errors, $F(1, 58) = 1.77$, $p = .189$, $\eta_p^2 = 0.03$.

Figure 1

Average Recognition Errors Made by Control and Stereotype Threat Group on Three Different Types of Words. Error Bars Indicate Standard Error.



We hypothesised that there should be a significant difference between free recall errors for the stereotype threat and the control group. That is, participants in the stereotype threat group should make more errors on this task than participants in the control group. Even though participants in the control group recalled numerically less correct words ($M = 4.07$ words, $SD = 3.09$) compared to the stereotype threat group ($M = 4.31$ words, $SD = 3.80$), this difference was not statistically significant, $t(58) = -0.27$, $p = .790$, $d = -0.07$, 95% CI [-2.05,

1.57]. Additionally, participants in the control group freely recalled similar number of intrusions (random words that were not present in the lists) (M = 1.04 words, SD = 2.38) compared to participants in the stereotype threat group (M = 1.03 words, SD = 2.01), $t(58)$ = 0.01, p = .994, d = 0.00, 95% CI [-1.13, 1.14]. Similarly, participants in the control group falsely recalled similar number of conjunction words (M = 0.18 words, SD = 0.39) to participants in the stereotype threat group (M = 0.22 words, SD = 0.49), $t(58)$ = -0.35, p = .729, d = -0.09, 95% CI [-0.27, 0.19]. Following the same pattern, participants in the stereotype threat group falsely recalled similar number of feature words (i.e., recalled either the first half or the second half of the word correctly) (M = 0.41 words, SD = 0.67) compared to the control group (M = 0.39 words, SD = 0.63), $t(58)$ = -0.08, p = .937, d = -0.02, 95% CI [-0.35, 0.32].

Similar to the recognition task, a repeated-measures ANOVA was conducted to examine the effect of paragraphs on false alarms made in the free-recall task. The three false alarms were conjunction, feature and intrusions. The main effect of paragraph was found to be not significant, $F(1, 58) = 0.06$, p = .937, $\eta_p^2 = 0.00$. To sum up the results thus far, we found no statistical difference in the recognition and free recall tasks between the control and the stereotype threat group. The next step was to investigate whether moderating factors influenced rates of conjunction errors.

Moderating Factors

Effects of Education on Retrieval Tasks

We predicted that highly educated participants in the stereotype threat group would make more false memory errors than highly educated participants in the control group. However, the difference in performance between the control and the stereotype threat group should be smaller when focusing on the less educated participants. Out of the 60 participants, 33 were highly educated (i.e., 15 years or above), and 26 were less educated. We chose 15

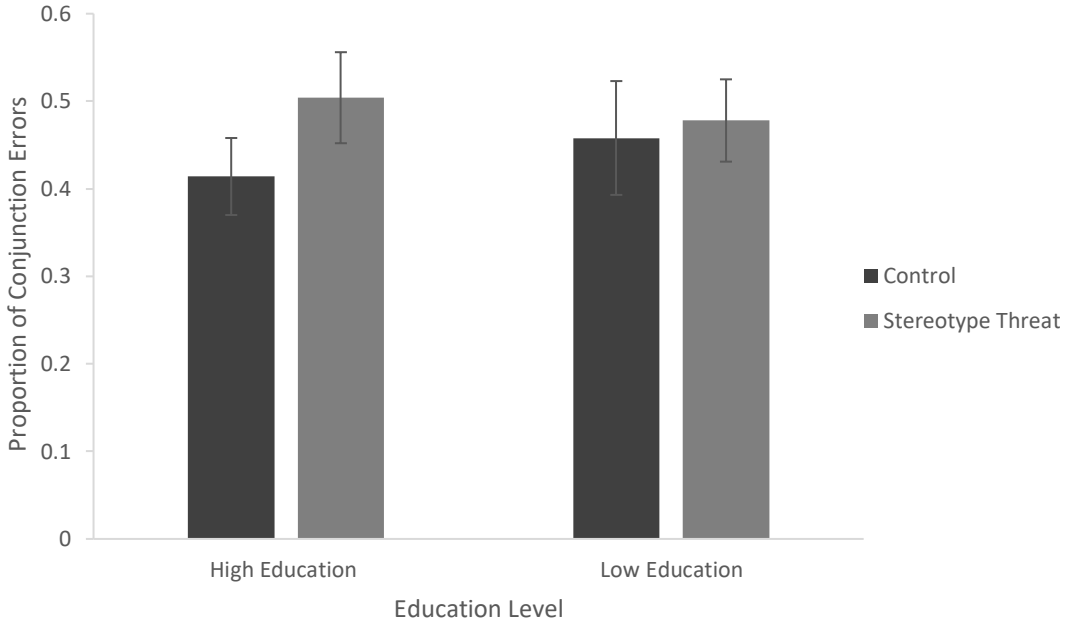
years or above as highly educated because anyone who has completed 15 years or more would have done at least a bachelor's degree. One participant did not state their years of education; therefore, they were not in any of the two groups. Figure 2 shows that the highly educated stereotyped participants made more conjunction errors compared to the control participants of the same education level. Nevertheless, the proportion of conjunction errors for the less educated participants remained similar for the control and the stereotyped participants. A 2 × 2 between-subjects ANOVA (paragraph: control/stereotype; education: high/low) was conducted to investigate the association between education and paragraph on conjunction errors. There were 19 participants in the high education-control group, 14 in the high education-stereotype threat group, nine in the low education-control group and 17 in the low education-stereotype threat group. The main effect of the paragraph was not significant for all following analyses (as mentioned above). Thus, the main effect of the paragraph will not be reported for each moderating factor. The main effect of education on conjunction errors was not-significant, $F(2, 57) = 0.36$, $p = .702$, $\eta_p^2 = 0.01$. That is, highly educated participants made a similar proportion of false memory errors on conjunction words ($M = 0.45$, $SD = 0.21$) as participants with low education ($M = 0.47$, $SD = 0.17$). For the interaction between education and paragraph, highly educated participants in the stereotype threat group made a similar rate of conjunction errors ($M = 0.50$, $SD = 0.24$) as participants in the control group ($M = 0.41$, $SD = 0.18$). Less educated participants in the stereotype threat group ($M = 0.48$, $SD = 0.18$) also produced a similar rate of conjunction errors to the control group ($M = 0.46$, $SD = 0.16$), $F(1, 55) = 0.45$, $p = .505$, $\eta_p^2 = 0.01$.

Focusing on the conjunction errors for free-recall, there was no significant main effect of education on conjunction errors. That is participants with high education made similar amount of errors on conjunction words ($M = 0.21$ words, $SD = 0.42$) as participants with low education ($M = 0.19$ words, $SD = 0.49$), $F(2, 57) = 0.11$, $p = .892$, $\eta_p^2 = 0.00$. For the

interaction, participants in the control group with high education falsely recalled a similar number of conjunction words (M = 0.16 words, SD = 0.38) compared to participants in the stereotype threat group with a same level of education (M = 0.29 words, SD = 0.47). Furthermore, participants in the control group with low education (M = 0.22 words, SD = 0.44) also falsely recalled a similar number of conjunction words compared to participants in the stereotype threat group (M = 0.18 words, SD = 0.53). Contrary to the hypothesis, this interaction was not statistically significant, $F(1, 55) = 0.49$, $p = .485$, $\eta_p^2 = 0.01$.

Figure 2

Influence of Education on Conjunction Errors Made in the Recognition Task for Participants in Stereotype Threat and Control Group. Error Bars Indicate Standard Error.



Effects of Employment Status on Retrieval Tasks

Another moderating factor analysed in this experiment was employment (retired versus not retired). Like education, the effects of employment were analysed for conjunction errors recognised and recalled. There were 39 retired participants and 21 still working (not retired) participants in this sample. It is illustrated in Figure 3 that retired participants in the control group made a fewer proportion of conjunction errors than the stereotyped

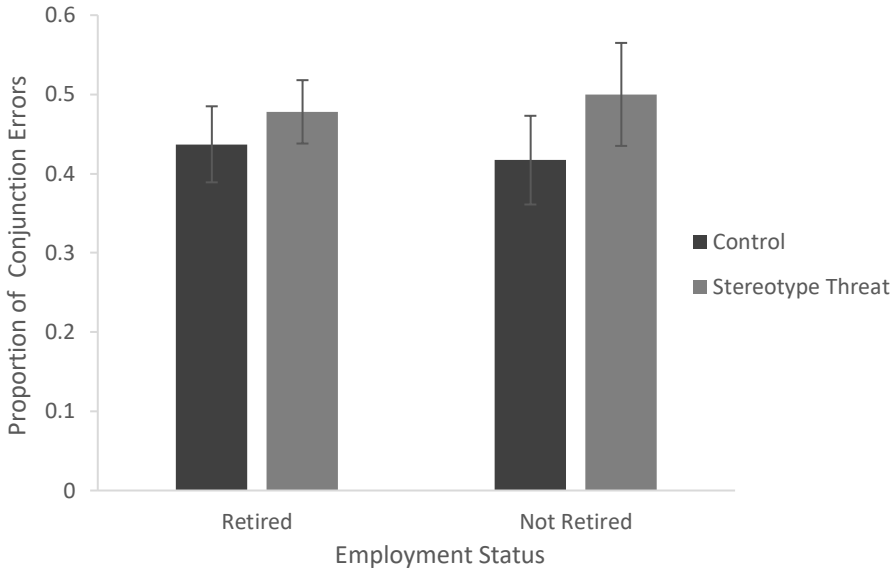
participants. Similarly, not retired participants in the control group also made a fewer proportion of conjunction errors compared to the stereotyped participants with the same employment status. These differences between the groups for both retired and not retired participants do not seem very reliable. The factors for 2 × 2 between-subjects ANOVA were paragraph type (control/stereotype threat) and employment status (retired/not retired). Out of 60 participants, 16 participants were in the retired-control group, 23 in the retired-stereotype threat, 12 in the not retired-control and nine in the not retired-stereotype threat group. Firstly, analysing the main effect of employment status illustrated that there was no significant difference in conjunction errors between the retired (M = 0.46, SD = 0.18) and the not retired participants (M = 0.45, SD = 0.21), $F(1, 58) = 0.03$, $p = .861$, $\eta_p^2 = 0.00$. Focusing on the interaction between employment status and paragraph type indicated that, retired participants in the control group made a similar number of conjunction errors (M = 0.44, SD = 0.15) to retired participants in the stereotype threat group (M = 0.48, SD = 0.20). Furthermore, not retired participants in the control group also made a similar number of conjunction errors (M = 0.42, SD = 0.20) compared to not retired participants in the stereotype threat group (M = 0.50, SD = 0.23). Contrary to the expectation, this interaction was not significant, $F(1, 56) = 0.16$, $p = .690$, $\eta_p^2 = 0.00$.

Looking at the main effect of employment status on conjunction errors for the free recall task, there was no difference in the number of conjunction errors between the retired (M = 0.18 words, SD = 0.39) and the not retired (M = 0.24 words, SD = 0.54) participants, $F(1, 58) = 0.24$, $p = .629$, $\eta_p^2 = 0.00$. When focusing on the interaction between employment status and paragraph type, retired participants in the control group freely recalled a similar number of conjunction errors (M = 0.19 words, SD = 0.40) compared to retired participants in the stereotype threat group (M = 0.17 words, SD = 0.39). Likewise, not retired participants in the control group (M = 0.17 words, SD = 0.39) freely recalled similar amount of

conjunction errors compared to not retired participants in the stereotype threat group (M = 0.33 words, SD = 0.71). This interaction was not statistically significant, $F(1, 56) = 0.53$, p = .469, $\eta_p^2 = 0.01$.

Figure 3

Effect of Employment Status on Conjunction Errors Made in the Recognition Task for Participants in Control and Stereotype Threat Group. Error Bars Denote Standard Error.



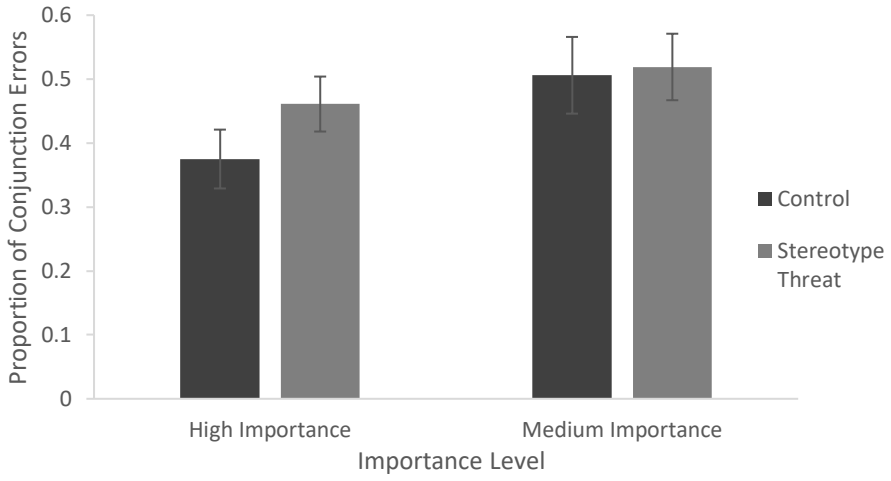
Effects of Importance Given to Memory on Retrieval Tasks

The importance given to memory was the third moderating factor accounted for in this experiment. Thirty-six participants rated the importance of memory between “9” and “10” (people who placed high importance on memory), and 23 participants rated the importance of memory between “3” and “8” (participants who placed medium importance on memory). No participant rated the importance of memory below “6” (participants who placed low importance on memory). However, one participant was unable to select a value; therefore, they were omitted. Figure 4 shows that control participants who placed high importance on memory made a fewer proportion of conjunction errors than the stereotyped participants in the same level. However, there does not seem to be any difference in the

proportion of conjunction errors between the control and the stereotyped participants that placed medium importance on memory. There were 17 participants in the high importance-control group, 19 in the high importance-stereotype threat group, 10 in the medium importance-control group, and 13 in the medium importance-stereotype threat group. For rates of falsely recognised conjunction errors, the main effect of importance was not significant, $F(2, 55) = 2.01$, $p = .144$, $\eta_p^2 = 0.07$. That is, participants that placed high importance on memory ($M = 0.42$, $SD = 0.20$) performed similarly to participants that placed medium importance on their memory ($M = 0.51$, $SD = 0.17$). For the interaction between importance and paragraph type, participants that placed high importance on memory made a similar proportion of conjunction errors in both the control ($M = 0.36$, $SD = 0.19$) and the stereotype threat group ($M = 0.46$, $SD = 0.21$). Similarly, participants in the control group that placed medium importance on memory also made equivalent conjunction errors ($M = 0.51$, $SD = 0.13$) to the stereotype threat group who also placed medium importance on memory ($M = 0.52$, $SD = 0.19$). This interaction was not statistically significant, $F(1, 55) = 0.51$, $p = .477$, $\eta_p^2 = 0.01$.

Figure 4

Effect of Importance on Average Conjunction Errors Made in the Recognition Task for Stereotype Threat and Control Group. Error Bars Show Standard Error.



For freely-recalled conjunction errors, focusing on the main effect first, participants that placed high importance on their memory freely recalled a similar number of conjunction words (M = 0.17 words, SD = 0.45) as participants that placed low importance on their memory (M = 0.25 words, SD = 0.44), $F(2, 55) = 0.32$, $p = .725$, $\eta_p^2 = 0.01$. For the interaction, control participants that placed high importance on memory made fewer conjunction errors (M = 0.18 words, SD = 0.39) compared to the ones in the stereotype threat group (M = 0.16 words, SD = 0.50). Control participants that placed medium importance on memory made a similar number of conjunction errors (M = 0.20 words, SD = 0.42) as stereotyped participants that placed medium importance on memory (M = 0.31 words, SD = 0.48), these differences were not statistically significant, $F(1, 55) = 0.27$, $p = .607$, $\eta_p^2 = 0.01$.

Discussion - Experiment 1

Experiment 1 investigated whether the type of retrieval task used (i.e., recognition task versus the free recall task) and other moderating factors (such as education, employment status, and importance placed on memory) influenced performance on false memory tasks (particularly, effects of stereotype threat on false memories). Additionally, previous research has shown no correlation between different memory tasks (Nichols & Loftus, 2019); hence, Experiment 1 used a different task (memory conjunction task) compared to the previous studies (Smith et al., 2017; Thomas & Dubois, 2011; Wong & Gallo, 2016).

Firstly, we evaluated whether there was a difference in performance between the control and the stereotype threat group for recognition and free recall tasks. There was no significant difference between the two groups; participants in the control group made a similar proportion of false memory errors as those in the stereotype threat group for recognition and retrieval tasks. Results from this experiment were inconsistent with previous literature (Smith et al., 2017; Thomas & Dubois, 2011; Wong & Gallo, 2016) as these studies found either a decrease (Smith et al., 2017; Thomas & Dubois, 2011) or an increase (Wong & Gallo, 2016) in performance due to the stereotype threat.

The use of a different task could explain why we did not observe any significant differences between the two groups. Experiment 1 used a memory conjunction task compared to the above studies (i.e., Smith et al., 2017; Thomas & Dubois, 2011; Wong & Gallo, 2016), which used a DRM task. The DRM and memory conjunction tasks are used to assess false memories (Deese, 1959; Roediger & McDermott, 1995; Underwood & Zimmerman, 1973). One similarity between the two tasks is that there is no misinformation being externally induced; instead, it is the internal processes (e.g., the similarity between the words for the memory conjunction task) that are creating those false memories (Coane et al., 2021; Deese,

1959; Jones & Jacoby, 2001; Roediger & McDermott, 1995; Underwood & Zimmerman, 1973).

Nevertheless, there are several differences between the two tasks. For example, in the DRM task, old words, critical lures, and new words are presented during the recognition task (Deese, 1959; Roediger & McDermott, 1995). While, in the memory conjunction task, old, conjunction, feature, and new words are presented (Underwood & Zimmerman, 1973). In the memory conjunction task, participants are required to be cautious of one extra category of words (feature words); therefore, it could be considered more difficult than the DRM task. Selective Engagement Theory could explain the effects of task difficulty on performance (Ryan & Campbell, 2021). As mentioned in the introduction, limited cognitive resources are available with age; thus, older adults become more cautious about employing their limited cognitive resources. When the task difficulty increases beyond one point, older adults choose to disengage from the task; consequently, this negatively impacts their performance (Ryan & Campbell, 2021). The difficulty of the memory conjunction task in comparison to the DRM task could be why there was no significant difference observed between the control and the stereotype threat group for both the recognition and the free-recall tasks. To elaborate, the threat does not make any difference as the task might be challenging to perform for both groups. This is similar to the second experiment of Thomas et al. (2020) study; that is, when participants were given a more challenging task, there was no difference observed between the control and the stereotype threat group. A study by Kroll et al. (1996) found that older adults had a lower hit rate (number of correct answers) on a memory conjunction task in comparison to young adults. This also indicates that older adults may have found the task difficult as they could not correctly remember even the initially studied words.

There could be different reasons why participants make false memory errors in the DRM and the memory conjunction task. Words in the DRM task are thematically linked;

thus, false memory errors in the DRM task occur mainly when participants focus on the overall gist of words rather than each word (Deese, 1959; Matzen & Benjamin, 2009; Roediger & McDermott, 1995). Since older adults rely more on gist-based encoding and retrieval due to changes in the hippocampus (Brainerd & Reyna, 2002; Devitt & Schacter, 2016; Kensinger & Schacter, 1999; Tun et al., 1998), then they are prone to making errors in the DRM task (Matzen & Benjamin, 2009). However, false memory errors in the memory conjunction task could arise mainly due to the familiarity and recollection processes (Duarte et al., 2010; Jones & Atchley, 2006). The recollection is usually affected with age, but the familiarity process remains unchanged (Spencer & Raz, 1995); thus, older adults rely more on familiarity with a decrease in the recollection process (Duarte et al., 2010). According to this model, false memory errors in the memory conjunction task occur because the conjunction word seems familiar (as both parts of the word were studied in the lists and in close proximity), causing the participant to think that they have studied the word earlier (Jones & Jacoby, 2001; Kroll et al., 1996; Underwood & Zimmerman, 1973). There are potentially different underlying mechanisms for false memories in the DRM and the memory conjunction task. Studies have shown no correlation between different false memory tasks, and if participants show false memories in one task, it does not mean they will show false memories in a different task (Nichols & Loftus, 2019; Ost et al., 2013). Thus, this could be why this experiment did not find any differences; however, studies using the DRM task did.

Furthermore, relying on the overall list rather than each word in the list decreases demands on cognition (Thomas & Sommers, 2005). Thus, making it less cognitively demanding for older adults to take part in the DRM task, where words are thematically linked (Matzen & Benjamin, 2009). Whereas words in the memory conjunction task are not linked (Underwood & Zimmerman, 1973), requiring older adults to attend to each word on the list,

this might add additional cognitive demands on older adults who already have limited cognitive resources (Ryan & Campbell, 2021).

However, some studies have shown false memories in older adults using a memory conjunction task (Anderson et al., 2008; Devitt, Tippett, et al., 2016; Fandakova et al., 2013; Jones & Jacoby, 2001). This indicates that the memory conjunction task has previously been successful in showing false memories. Thus, it would be incorrect to say that this task is ineffective in inducing false memories. Nevertheless, no research has focused on stereotype threat using this task, which raises an important question: Is the memory conjunction task a good task to examine the effects of age-based stereotype threats on false memories? It does not seem easy to pinpoint an exact mechanism behind false memories. Researchers focused on false memories have explored different tasks and have found different reasons for false memories (Abadie & Camos, 2019; Bharti et al., 2020; Dodson & Schacter, 2002; Giovanello et al., 2010; Leding & Lampinen, 2009; Mammarella & Fairfield, 2008; Yassa et al., 2011). Therefore, it is important to keep researching and focusing on different factors and variables to understand why we found no differences between our groups. Additionally, this experiment did not find any effects of moderating factors on performance. This could be due to using an online sample (this will be discussed in general discussion).

From our results, we were not certain whether the lack of difference in performance between the two groups was due to the memory conjunction task versus the DRM task or because we recruited an online sample. Therefore, we used a DRM task in an online sample for our second experiment. Thus, the main aim of this experiment was to compare whether the stereotyped participants were making more false memory errors compared to the control participants in the DRM task. Additionally, we explored whether the focus of the task (prevention or promotion) could influence the effects of stereotype threat on performance.

This would help account for the differences in Thomas and Dubois' (2011) and Wong and Gallo's (2016) studies.

Methods - Experiment 2

Participants

One hundred thirty-four participants aged 65 years and over took part in this experiment. Fourteen participants were excluded as one participant failed the attention check, and 13 participants noted that they were clinically diagnosed with Parkinson's disease, stroke, traumatic brain injury or brain tumour. As such, data were analysed from the remaining 120 participants (M = 69.5 years, SD = 3.40, 78 female). The number of participants was doubled compared to our first experiment. Participants were recruited from Cloud Research and were paid U.S. $1.50 (N.Z. $2.10) for participating in the experiment. The same criteria as Experiment 1 were applied to Cloud research (e.g., location and age). Additionally, participants entered a draw to win a bonus award of U.S. $2 (N.Z. $2.88).

Participants were not approved until all data were collected. Once that happened, they were entered in a draw (the number of entries depended on their performance) to win. A random draw was then carried out with all the workers' identification. One participant was paid a bonus of U.S. $2 in addition to their U.S. $1.50.

Materials

The questionnaire in the current experiment was also administered with Qualtrics^{XM} (same as Experiment 1). This experiment required different stimuli from Experiment 1, it required 12 DRM lists with 12 words in each list therefore, participants were shown 144 words in total (refer to Appendix B). The stereotype threat and control paragraphs were adapted from Wong and Gallo (2016) and were the same as our Experiment 1 (refer to Appendix A.2). To recap, the stereotype threat paragraph was about negative changes in memory with age, whereas the control paragraph was about language. In addition, this experiment also required another manipulation to change the focus of the task (promotion or prevention). This experiment used a similar method to Markman et al. (2005) to induce

promotion and prevention focus; that is, participants in the promotion focus were told "currently, your name is in the raffle once; however, if you perform better than 50% on the memory test, then your name will be added to the raffle twice" whereas participants in the prevention focus were told; "currently your name is in the raffle two times, if you perform poorer than 50% then your name will only be added once".

Procedure

This experiment involved the use of DRM lists instead of memory conjunction lists and was followed by the manipulations and retrieval tests. The experiment took approximately 12 minutes to complete. The main differences between the two experiments were the use of a new task and an additional manipulation (to change the focus). Participants in this experiment were assigned to four groups; control-promotion, control-prevention, stereotype threat-promotion, and stereotype threat-prevention, and there were 30 participants in each group.

Like Experiment 1, all participants read a consent form and were asked to provide their worker identification, as part of this online study (Experiments 1 and 2). Participants were then instructed about the DRM task. For example, "you will now see 12 word-lists, and each word will be shown for two seconds". This experiment used a similar attention check question as our previous experiment, for example, "how many seconds will you see each word". Participants then studied 12 lists, one at a time, words in each list were semantically related, such as the first list included: fast, lethargic, stop, listless, snail, cautious and so on. Each list was different; however, words within the list remained semantically related.

Once participants had viewed all the words, the following instructions were given: "you will now view a paragraph for 90 seconds and be asked a question about it". Participants were then randomly assigned to either the stereotype threat group or the control group and the same paragraphs (stereotype threat and control) and attention questions as

Experiment 1 were used. To recap, participants in the stereotype threat group were asked if adults over the age of 60 report more memory problems, while participants in the control group were given three statements about language processing to choose from.

Participants were then instructed about the free-recall task. Participants were given 90 seconds to write as many words as they remembered from the 144 words on the DRM lists. Following the free-recall task, participants were given one of two instructions to change the task's focus (prevention or promotion). Qualtrics^{XM} randomly presented one of the two instructions to each participant. Participants then had 15 seconds to read the prevention or promotion manipulation.

After this, participants undertook a recognition test where they were given 48 words in total. Out of the 48 words, 12 were old words (the words that were studied earlier in the 12 word lists, e.g., fast), 12 were critical lures (the words that were similar to the words studied earlier but were not studied, e.g., slow was the critical lure for list 1), and 24 were new words (e.g., town). Finally, participants answered the same attention check and demographic questions as our first experiment. Questions about age, education, retirement, importance placed on memory and any clinical diagnosis were presented. Similar to Experiment 1, participants were also asked if they cheated or if they had any technical problems.

Data Processing

Firstly, we determined the false recognition and free-recall scores for prevention and promotion focus groups. Hit rates were calculated by dividing correct answers on old words by the total number of old words (i.e., 12 words). The false alarms for critical lures were calculated by dividing the number of errors on critical lures by the total number of critical lures (i.e., 12 words). The false alarms on new words were calculated by dividing the number of errors on new words by the total number of new words (i.e., 24 words).

Several different ANOVA tests were performed in SPSS to investigate whether the focus type impacted the performance of control and stereotyped participants. We also compared whether there was any significant effect of stereotype threat when using the DRM task compared to the memory conjunction task. In addition, we performed a series of between-subjects ANOVAs to observe if there was a three-way interaction between focus type, paragraph, and any of the moderating factors (i.e., education, employment and importance placed on memory).

Important to note that promotion and prevention manipulation was induced after the free-recall task. Therefore, we did not expect any changes in performance due to the interaction between the different paragraphs and focus type for the free-recall task. Thus, the analyses for the free-recall task will only focus on whether there was a difference in critical lure errors between participants in the control and stereotype threat groups. We did not use proportions for the free recall task, instead used raw scores (similar to Experiment 1). An independent samples t-test was performed to investigate whether stereotype threat influenced the number of critical lure errors freely recalled. Furthermore, several different 2×2 between-subject ANOVAs (one for each moderating factor) were carried out to investigate effects of moderating factors on freely recalled critical lure errors.

Results - Experiment 2

Recognition and Recall Errors

We first looked at whether there were any differences in hit rates with a change in focus type (prevention versus promotion) for participants in the control and the stereotype threat group. There were 60 participants in the promotion focus and 60 in the prevention focus. Additionally, there were 60 participants each in the control and the stereotype threat conditions. A between-subject 2 × 2 ANOVAs (focus: prevention/promotion, paragraph: control/stereotype threat) were carried out for all following analyses to investigate if there was a change in performance (i.e., hit rates, critical lures and errors on new words) due to changes in internal and task's focus. Focusing on the main effect of focus type on hit rates, participants in the prevention focus recognised a similar proportion of correct hits (M = 0.82, SD = 0.19) as participants in the promotion focus (M = 0.84, SD = 0.15), this difference was not significant, $F(1, 118) = 0.31$, $p = .580$, $\eta_p^2 = 0.00$. Similarly, focusing on the main effect of paragraph type on performance, participants in the control group correctly recognised similar proportion of correct hits (M = 0.80, SD = 0.20) as participants in the stereotype threat group (M = 0.86, SD = 0.13), the difference was not-significant, $F(1, 118) = 3.66$, $p = .058$, $\eta_p^2 = 0.03$. Focusing on the interaction between the paragraph type (control versus stereotype threat) and focus of the task (promotion versus prevention), stereotyped participants in the prevention focus correctly recognised a higher proportion of old words (M = 0.88, SD = 0.12), compared to control participants in the prevention focus (M = 0.76, SD = 0.24). Furthermore, stereotyped participants in the promotion focus correctly recognised a similar proportion of old words (M = 0.84, SD = 0.14) as control participants in the promotion focus (M = 0.84, SD = 0.16). This interaction was not statistically significant, $F(1, 116) = 3.73$, $p = .056$, $\eta_p^2 = 0.03$.

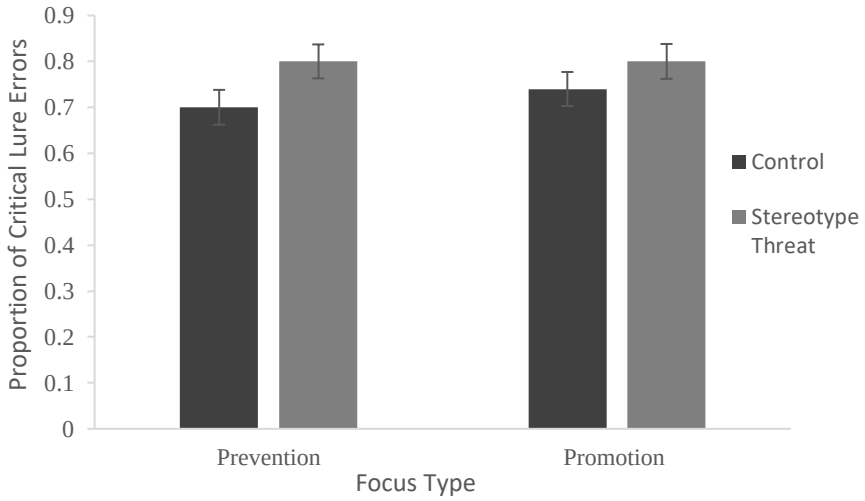
Figure 5 below shows the proportion of critical lure errors made on the recognition task. Participants in the control group made fewer errors in the prevention and promotion group when compared to the stereotyped participants, nevertheless, the difference does not seem reliable. The second analysis focused on critical lures, that is, the proportion of false memory errors made on critical lure words. Focusing on the two independent variables separately, participants in the prevention focus made a similar proportion of critical lure errors (M = 0.75, SD= 0.22) as participants in the promotion focus (M = 0.72, SD = 0.19), and the main effect of focus type on critical lures was not statistically significant, $F(1, 118) = 0.69$, $p = .409$, $\eta_p^2 = 0.01$. The same pattern was observed for the main effect of paragraph type on critical lures, that is, participants in the control (M = 0.72, SD = 0.24) and the stereotype threat group (M = 0.75, SD = 0.17) made a similar proportion of critical lure errors, $F(1, 118) = 0.63$, $p = .430$, $\eta_p^2 = 0.01$. For the interaction between paragraph type and focus type, control participants in the prevention focus made a fewer proportion of critical lure errors (M = 0.70, SD = 0.26) than stereotyped participants in the prevention focus (M = 0.80, SD = 0.16). In addition, control participants with promotion focus made similar proportion of critical lure errors (M = 0.74, SD = 0.21) compared to stereotyped participants in promotion focus (M = 0.70, SD = 0.17). This interaction was also not significant, $F(1, 116) = 3.18$, $p = .077$, $\eta_p^2 = 0.03$.

Following this, we looked at the proportion of false memory errors made on new words. Focusing on the two independent factors first; participants in the prevention focus (M = 0.22, SD = 0.18) made a similar proportion of false memory errors as participants in the promotion focus (M = 0.20, SD = 0.17), $F(1, 118) = 0.29$, $p = .589$, $\eta_p^2 = 0.00$. Likewise participants in the control group (M = 0.20, SD = 0.17) and the stereotype threat group (M = 0.21, SD = 0.19) also made similar proportions of false memory errors on new words, $F(1, 118) = 0.11$, $p = .746$, $\eta_p^2 = 0.00$. Focusing on the interaction, it was observed that

stereotyped participants in the prevention focus made a similar proportion of false memory errors on new words (M = 0.24, SD = 0.18) compared to control participants in the prevention focus (M = 0.21, SD = 0.19). Likewise, stereotyped participants in the promotion focus made a similar proportion of errors on new words (M = 0.20, SD = 0.20) as control participants in the promotion focus (M = 0.21, SD = 0.14). This difference was not statistically significant, $F(1, 116) = 0.44$, $p = .511$, $\eta_p^2 = 0.00$.

Figure 5

The Proportion of Critical Lure Errors Made by the Stereotyped and Control Participants in Prevention and Promotion Focus. Error Bars Indicate Standard Error.



As mentioned above, the tasks focus was changed after the free recall task. Thus, an independent samples t-test was carried out to investigate the difference between groups (control and stereotype threat) for the number of critical lure errors recalled. Participants in the control group (M = 1.13 words, SD = 1.24) made a similar number of critical lure errors as participants in the stereotype threat group (M = 1.17, SD = 1.33), $t(118) = -0.14$, $p = .887$, $d = -0.03$, 95% CI [-0.50, 0.43].

Moderating Factors

Effects of Education on Retrieval Tasks

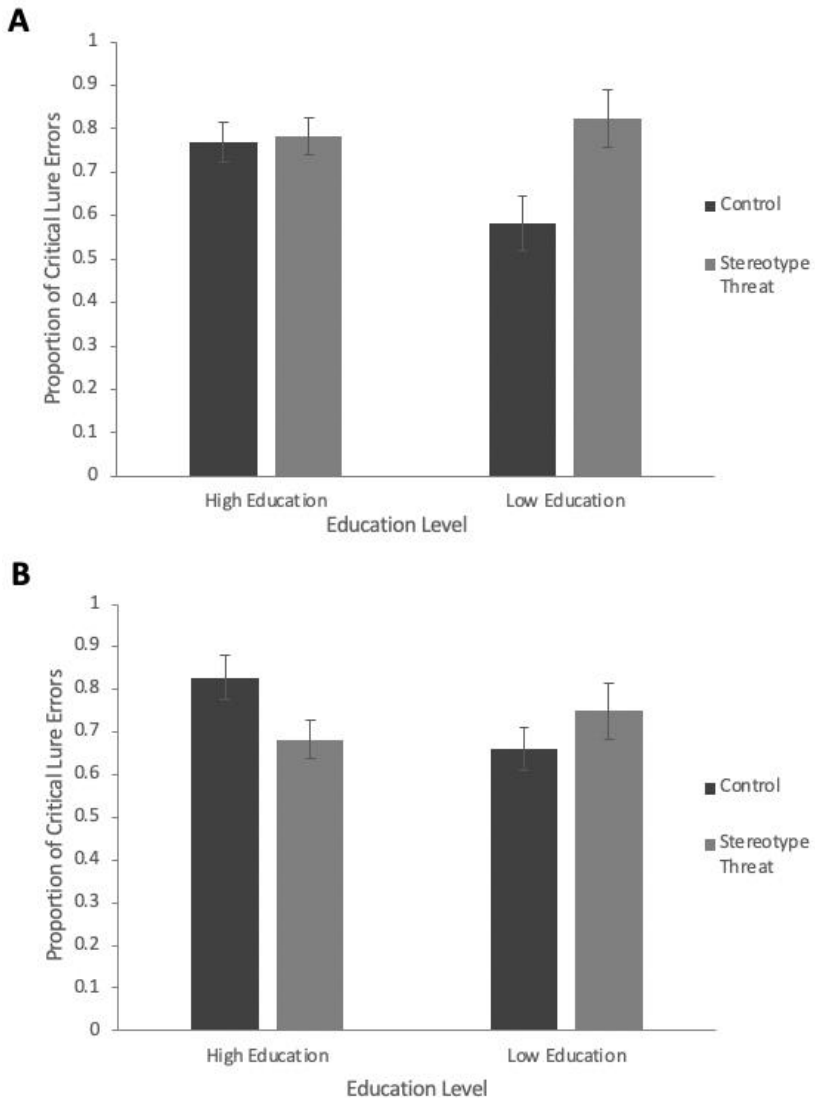
For our first experiment, we did not see any impact of education or any other moderating factor on performance. However, we expected it might change with a different task, and education or other moderating factors might influence performance. Out of 120 participants, 75 were highly educated, 44 were less educated (as with Experiment 1, participants with 15 years or more of education were categorised into the high education group), and one participant did not state their years of education. Out of the 75 highly educated participants, 15 were in the control-promotion group, 19 in the control-prevention group, 20 in the stereotype threat-promotion group and 21 in the stereotype threat-prevention group. For the less educated participants, 16 were in the control-promotion group, 10 in the control-prevention group, nine in the stereotype threat-promotion group and nine in the stereotype threat-prevention group.

Figure 6A shows that high education did not influence the proportion of critical lure errors made by the control and the stereotyped participants when given a prevention-focused recognition task. However, less educated stereotyped participants made more critical lure errors than less educated control participants in the same focus. Nevertheless, Figure 6B shows that higher education had an impact when given a promotion task; that is, control participants made a higher proportion of critical lure errors than the stereotyped participants. This trend was reversed for less educated participants in the promotion focus (i.e., control participants made a fewer proportion of errors than stereotyped participants). A between-subjects 2 × 2 × 2 ANOVA was carried out to investigate whether there was a significant interaction between education (high/low), paragraph (control/stereotype threat) and focus type (prevention/promotion). For the main effect of education on false alarms, participants with high education made a similar proportion of critical lure errors (M = 0.76, SD = 0.18) as

participants with less education (M = 0.70, SD = 0.24), $F(2, 111) = 1.68$, $p = .192$, $\eta_p^2 = 0.03$. The interaction between education level, focus type and paragraph type was observed to be not significant, $F(1, 111) = 0.00$, $p = .978$, $\eta_p^2 = 0.00$. Because the main effects of paragraphs (control versus stereotype threat) and focus type (prevention versus promotion) on critical lure errors were not significant, the main effects for paragraph and focus type will not be provided for each moderating factor.

Figure 6

The Influence of Education, Focus Type (Prevention: A, Promotion: B), and Paragraph Type on Proportion of Critical Lure Errors Recognised. Error Bars Denote Standard Error.



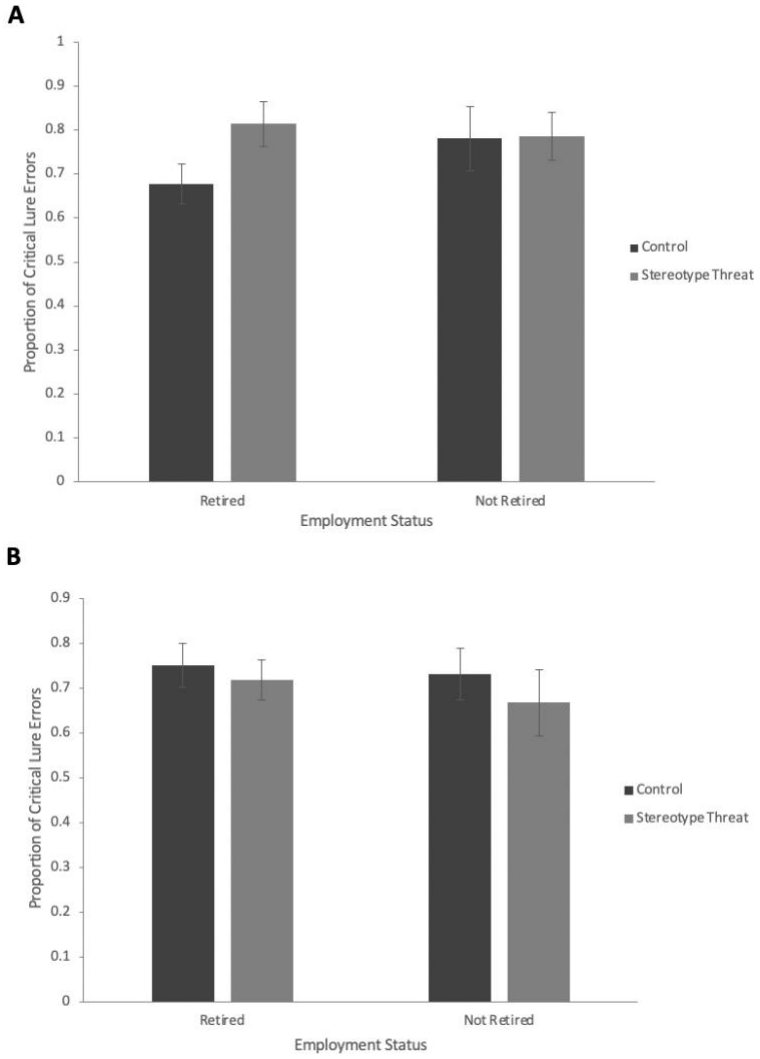
As mentioned above, prevention and promotion manipulation was induced after the free-recall task. Therefore, a 2 × 2 between-subjects ANOVA consisting education (high/low) and paragraph type (control/stereotype threat) as the factors was conducted. Main effect of education on critical lure recall was not significant, $F(2, 115) = 0.03$, $p = .970$, $\eta_p^2 = 0.00$. That is, highly educated participants ($M = 1.17$ words, $SD = 1.18$) recalled a similar number of critical lure errors as less educated participants ($M = 1.11$ words, $SD = 1.47$). Similarly, the main effect of paragraph was also not significant, $F(1, 115) = 0.03$, $p = .867$, $\eta_p^2 = 0.00$. Thus, participants in the control group ($M = 1.13$ words, $SD = 1.24$) recalled a similar number of critical lure errors as participants in the stereotype threat group ($M = 1.17$ words, $SD = 1.33$). For the interaction, highly educated participants in the control group ($M = 1.18$ words, $SD = 1.09$) made a similar number of critical recall errors as participants in the stereotype threat group ($M = 1.17$, $SD = 1.26$). Similarly, less educated participants in the control group ($M = 1.08$ words, $SD = 1.44$) and the stereotype threat group ($M = 1.17$ words, $SD = 1.54$) made a similar number of critical lure errors on a free-recall task. Thus, this interaction was not significant, $F(1, 115) = 0.04$, $p = .849$, $\eta_p^2 = 0.00$.

Effects of Employment Status on Retrieval Tasks

Figure 7A shows the performance of retired and not retired participants in the control and the stereotype threat group when given a prevention-focused recognition task. Retired participants in the control group made a fewer proportion of critical lure errors than retired participants in the stereotype threat group, but there was no apparent difference in performance for the working (not retired) participants in the control and stereotyped groups. In a promotion-focused task, retired participants in the control and the stereotype threat group made a very similar proportion of errors. Whereas, not retired participants in the control group made a higher proportion of errors than the stereotyped participants (see Figure 7B).

Figure 7

The Influence of Employment, Focus Type (Prevention: A, Promotion: B) and Paragraph Type on Proportion of Critical Lure Errors Recognised. Error Bars Show Standard Error.



To analyse whether employment status (retired/not-retired), paragraph (control/stereotype threat), and focus type (prevention/promotion) had any impacts on performance, a between-subjects 2 × 2× 2 ANOVA was conducted. There were 77 participants who were not currently working (i.e., retired) and 43 who were still working (not retired). Of the 77 retired participants, 18 were in the control-promotion group, 21 in the control-prevention group, 21 in the stereotype threat-promotion group, and 17 in the stereotype threat-prevention group. Additionally, out of the not retired participants, 13 were

in the control-promotion group, eight in the control-prevention group, eight in the stereotype threat-promotion group, and 14 in the stereotype threat-prevention group. There was no main effect of employment status on the proportion of critical lure errors. That is, retired participants (M = 0.74 , SD = 0.21) made a similar proportion of critical lure errors as not retired participants (M = 0.75 , SD = 0.19), $F(1, 112) = 0.00$, $p = .971$, $\eta_p^2 = 0.00$. For the interaction between paragraph type, focus and employment status, there was no significant difference observed in the performance, $F(1, 112) = 0.39$, $p = .535$, $\eta_p^2 = 0.00$.

As our first moderating factor, a 2 × 2 between-subjects ANOVA was carried out to investigate if there was an interaction between paragraph type and employment status (the two factors for our ANOVA) for the free recall task. The main effect of employment on the number of critical lures recalled was non-significant, $F(1, 116) = 3.44$, $p = .066$, $\eta_p^2 = 0.03$. That is retired participants (M = 1.31 words, SD = 1.35) recalled a similar number of critical lure words as not retired participants (M = 0.86 words, SD = 1.10). The interaction between paragraph type and employment status was not significant, $F(1, 116) = 0.01$, $p = .913$, $\eta_p^2 = 0.00$.

Effects of Importance Given to Memory on Retrieval Tasks

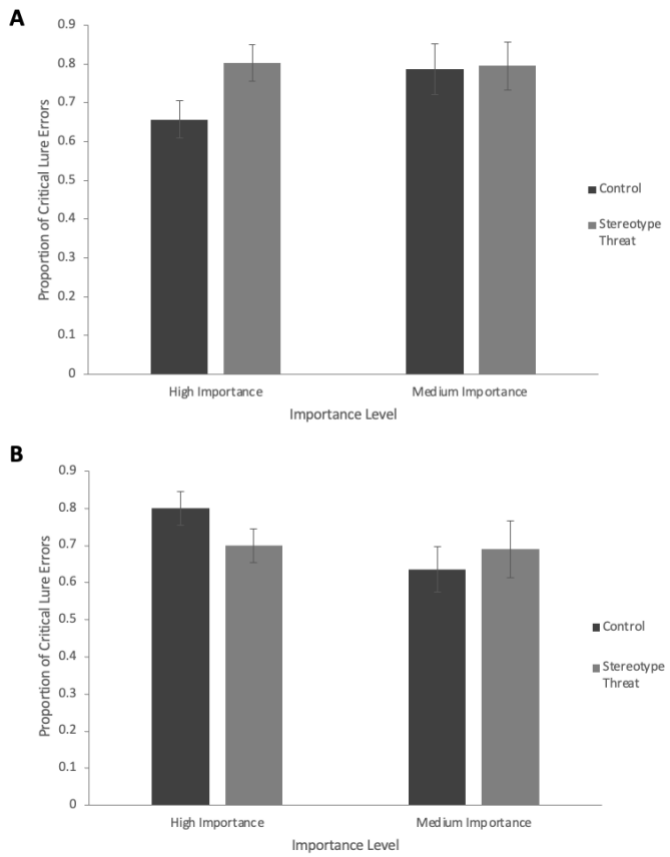
There were 77 participants that rated their memory between “9” and “10” (i.e., participants that placed high importance on memory), and 39 participants rated their memory between “3” and “8” (categorised as placing medium importance). Four participants did not select the importance of their memory; therefore, they were not included in this analysis. Out of the 77 participants that placed high importance on memory, 20 were in the control-promotion group, 18 in the control-prevention group, 20 in the stereotype threat-promotion group, and 19 in the stereotype threat-prevention group. Out of the participants that placed medium importance on memory, 11 were in the control-promotion group, 10 in the control-

prevention group, seven in the stereotype threat-promotion group and 11 in the stereotype threat-prevention group.

From Figure 8A, it seemed that control participants who placed high importance on memory made a fewer critical lure errors when given a prevention focused task compared to stereotyped participants that also placed high importance. Participants (control and stereotyped) that placed medium importance on memory made similar proportion of errors in the task with a prevention focus. Nevertheless, Figure 8B shows that control participants that placed high importance on memory made more errors in promotion focus than stereotyped participants. Control and stereotyped participants that placed medium importance on memory made a very similar proportion of errors in promotion focus.

Figure 8

The Influence of Importance, Focus (Prevention: A, Promotion: B), and Paragraph Type on Proportion of Critical Lure Errors Recognised. Error Bars Indicate Standard Error.



A between-subjects 2 × 2 × 2 ANOVA was carried out to investigate whether there was any interaction between paragraph type (control/stereotype threat), importance level (high/medium) and focus type (promotion/prevention). There was no main effect of importance on false alarms for critical lure words; participants that placed high importance (M = 0.74, SD = 0.22) performed similarly to participants that placed medium importance (M = 0.73, SD = 0.20), $F(2, 109) = 0.17$, $p = .847$, $\eta_p^2 = 0.00$. For the interaction between the three factors, there was no significant difference observed in performance, $F(1, 109) = 3.21$, $p = .076$, $\eta_p^2 = 0.03$.

For the free recall task, a 2 × 2 between-subjects ANOVA indicated no significant main effect of importance on the number of critical lures recalled, $F(2, 114) = 0.06$, $p = .940$, $\eta_p^2 = 0.00$. Participants that placed high importance (M = 1.12 words, SD = 1.23) recalled a similar number of critical lure words as participants that placed medium importance (M = 1.18 words, SD = 1.34). Main effect of paragraph is listed above. There was no significant interaction of importance and paragraph on the number of critical lures recalled, $F(2, 114) = 0.10$, $p = .907$, $\eta_p^2 = 0.00$.

Discussion – Experiment 2

This experiment aimed to investigate whether using a DRM task instead of a memory conjunction task would yield some significant differences in performance between the control and the stereotype threat group. In addition to the type of retrieval task used and the three moderating factors, we also examined the effects of promotion and prevention focus on false memory performance. We predicted that participants in the stereotype threat group with a prevention focus should perform better than control participants in the same focus (Barber & Mather, 2013). This is because there will be a match in the focus of the task (prevention focus) and the participant's internal regulatory state (stereotype threat induces prevention focus) (Barber & Mather, 2013; Mazerolle et al., 2021; Seibt & Förster, 2004). Furthermore, participants in the control group should perform better on the DRM task than those in the stereotype threat group (Smith et al., 2017; Thomas & Dubois, 2011).

We first examined whether focus type (prevention versus promotion) and paragraph type (control versus stereotype threat) influenced hit rates and critical lure errors for the recognition task. It was obtained that control participants in the prevention focus performed similarly to stereotyped participants in the prevention focus. Likewise, control and stereotyped participants in the promotion focus also performed similarly on the recognition task. The two main findings from this analysis were; that there was no difference in performance between participants in the control and the stereotype threat group for the DRM task in general, and there was no difference in performance due to a change in the focus either. Furthermore, there was no difference in the number of critical lure errors recalled between the control and stereotype threat groups. The results from our experiment were contradictory to Thomas and Dubois' (2011) and Wong and Gallo's (2016) studies. To recap, Thomas and Dubois (2011) found that stereotyped participants made more false memory errors using a promotion-focused (Mazerolle et al., 2021) DRM task. In contrast, Wong and

Gallo (2016) found that stereotyped participants performed better than control participants using a prevention-focused (Mazerolle et al., 2021) DRM task. Thus, we aimed to find something similar to Wong and Gallo's (2016) study for our stereotyped participants in the prevention group.

Two reasons could account for why this experiment could not find any significant findings (despite using a DRM task) compared to previous studies (Smith et al., 2017; Thomas & Dubois, 2011; Wong & Gallo, 2016). First, participants in Wong and Gallo (2016) were given the stereotype threat and control paragraphs twice. Furthermore, the experimenter read those paragraphs to the participants. This was similar to Thomas and Dubois's (2011) study but not Smith et al. (2017). Previous studies (Smith et al., 2017; Thomas & Dubois, 2011; Wong & Gallo, 2016) and the current experiment used blatant induction of stereotype threat (Mazerolle et al., 2021). To elaborate, blatant or explicit induction is when participants are directly told about adverse changes in performance with age. In comparison, subtle induction is less direct (Mazerolle et al., 2021; Stone & McWhinnie, 2008) and could be induced through the involvement of an outgroup experimenter (Marx & Goff, 2005). Likely, participants in Thomas and Dubois' (2011) and Wong and Gallo's (2016) studies were read a paragraph by someone not 65 years or above, therefore, an outgroup member. This could potentially lead to subtle stereotype threat induction in addition to the blatant induction that was used. Kang and Chasteen (2009) found that older adults' memory performance was affected when a younger person was present in the lab. Additionally, a meta-analysis conducted by Lamont et al. (2015) suggested that subtle stereotype induction compared to the blatant induction was more powerful in affecting older adults' working memory performance. This is because subtle inductions are associated with high levels of ambiguity (Lamont et al., 2015). Nevertheless, both induction strategies could work together and cause a greater threat impact than when they are used alone (Seibt & Förster, 2004; Stone & McWhinnie, 2008).

The current experiment was online; therefore, only blatant induction was used, and this could mean that the stereotype threat induction was not as powerful as in the previous studies (Thomas & Dubois, 2011; Wong & Gallo, 2016). This could explain why we did not observe any significant difference between the control and stereotype threat groups. It is acknowledged that Smith et al. (2017) did not read the stereotype threat paragraph to participants; however, they did state that their study was carried out by a younger person in an ageing lab. Thus, this could have worked as a subtle stereotype threat for the stereotyped and the control group.

To add to the above paragraph, subtle inductions might lead to participants internalising the threat compared to blatant induction (Strickland-Hughes & West, 2021). Internalising could lead to self-stereotyping in older participants (Barber, 2017; Seibt & Förster, 2004), and as mentioned in the introduction, stereotype threats are efficient when older adults feel part of the stereotyped group (Hess et al., 2003; Pennington et al., 2016; Steele, 1997). If older adults internalise the threat, they would feel part of the stereotyped group; thus, they are likely to be affected by it. So, in addition to a weaker stereotype threat induction, a lack of induction that causes internalising of the threat could be why there were no differences between the groups and why a change in task's focus did not change performance for the control and the stereotyped participants. The stereotype threat might have been unsuccessful in changing the internal state of the participant from promotion to prevention. Thus, changing the focus of the task did not have any significant effect on performance.

Another reason for discrepant findings between Experiment 2 and previous studies (Wong & Gallo, 2016) could be the manipulation used. That is, Wong and Gallo (2016) used a warning to induce prevention focus and this warning was given to both the control and the stereotype threat group. The warning aimed to inform participants that they cannot perform

this task based solely on gist information; instead, they need to attend to the item-specific information. This warning (asking them to be cautious) causes a match in the regulatory state of the stereotyped participants and the task. Thus, making them less likely to make false memory errors. However, our manipulation was adapted from Markman et al. (2005) and did not encourage older adults to be careful and avoid gist-based recollection. Instead, this manipulation warned them about gaining or losing an entry in the raffle. This could be problematic because, in an online sample recruited through online platforms, participants want to maximise their gains (earn as much as they can), which is usually their primary reason for taking part in studies (Turner et al., 2020). Therefore, it is possible that our manipulation just made all participants focus on winning an extra U.S. $2. To elaborate, participants in the prevention focus that were instructed about losing an entry in the raffle could have become cautious in how they responded; similarly, participants in the promotion focus that were informed about gaining another entry could have also become cautious in responding to the test (Turner et al., 2020). This could mean that all participants were in the same focus. Thus, a combination of a changed focus manipulation and a weaker stereotype threat compared to the previous studies (Smith et al., 2017; Thomas & Dubois, 2011; Wong & Gallo, 2016) could have led to the inconsistent findings in this experiment.

Similar to our Experiment 1, we did not observe any significant interactions between the paragraph type (stereotype threat and control), focus type (prevention and promotion) and the moderating factors (education, employment status and importance placed on memory). This will be explained in the general discussion. To sum up, this follow up experiment also did not yield any significant findings between groups despite using a DRM task. Furthermore, a change in the tasks focus also did not change performance in any way. Few methodological differences could have led to these conclusions.

General Discussion

Previously, many studies have observed adverse effects of age-based stereotype threat on memory (Bouazzaoui et al., 2020; Chasteen et al., 2005; Hess et al., 2003; Levy et al., 2012; Rahhal et al., 2001); however, studies focused on false memory remains limited and conflicted (Smith et al., 2017; Thomas & Dubois, 2011; Wong & Gallo, 2016). The discrepant findings could be due to the methodological differences across these studies (Mazerolle et al., 2021). Thus, this thesis aimed to account for the methodological differences through Experiments 1 and 2. Experiment 1 aimed to explore whether the type of retrieval task and moderating factors (such as education, employment status and importance placed on memory) affected performance on a previously unexplored false memory task (i.e., memory conjunction task). While Experiment 2 aimed to investigate whether the focus of the task (in addition to the type of retrieval task used and the moderating factors) affected performance on a more popular DRM task.

In our first experiment, we could not find any differences in performance between the control and the stereotype threat group for both retrieval tasks. Since this finding was at odds with previous research, it was suggested that this could have been due to using a previously unexplored task. The memory conjunction task could be considered more complex than the DRM task as there is an additional category of words (e.g., feature words) (Underwood & Zimmerman, 1973). Thus, older adults selectively engage their limited cognitive resources and stop engaging in the task (Ryan & Campbell, 2021). Furthermore, the DRM and the memory conjunction task could lead to false memories through different processes.

Our second experiment also could not yield any significant differences. It was observed that there was no change in performance between the control and the stereotyped participants when using a DRM task, additionally, a change in the focus of the task (from promotion to prevention) also did not affect the performance and this was contrary to the

expectations. This could have been due to how stereotype threat was induced in previous studies and in Experiment 2 (Thomas & Dubois, 2011; Wong & Gallo, 2016).

Nevertheless, other factors that could be noteworthy and are not mentioned above might help explain the differences in findings. Firstly, this study (Experiments 1 and 2) was an online study and recruited participants from an online platform, while prior studies were in person and recruited participants from the community (Smith et al., 2017; Thomas & Dubois, 2011; Wong & Gallo, 2016). Secondly, the reduced number of words in our tasks compared to the previous false memory literature (Smith et al., 2017; Thomas & Dubois, 2011; Wong & Gallo, 2016) could have played a role in not observing any significant differences between the control and the stereotype threat group.

Other Noteworthy Factors

Online Study Versus In-Person

There are several advantages of conducting an online study. A study by Ryan and Campbell (2021) suggested that lab settings could be stressful for older adults, thus; inducing threats even when there are no threats present in the study. This study suggested that testing older adults in their homes would make them feel more comfortable and less stressed, giving a more accurate representation of changes in performance on cognitive tasks. Moreover, online sampling lets researchers reach a more varied pool of participants in the community (Cyr et al., 2021). It also allows participants who cannot access lab settings, for example, due to lack of transportation, to take part in studies (Hultsch et al., 2002).

However, the reduced stress and anxiety could have led to discrepant findings in this study. There is some evidence (Payne et al., 2002) that lab settings can cause increased stress, increasing false memories. This could explain why previous studies (Smith et al., 2017; Thomas & Dubois, 2011; Wong & Gallo, 2016) found either an adverse or an enhancing effect of threat on false memories. Nevertheless, some researchers believe that lab-related

stress could lead to fewer or no changes in false memories (Beato et al., 2013; Zoladz et al., 2014). This could account for differences in our findings, but more research in this field is needed to confidently state this as one of the reasons for inconsistent findings. The present study also recruited participants from online platforms instead of the community.

Sample From Online Platforms

As mentioned above, age is associated with a decline in memory performance (Nyberg et al., 2012; Rosa & Gutchess, 2013); nevertheless, this decline is varied between high and low-functioning older adults (Nyberg et al., 2012). Participants on online platforms are usually high-functioning adults compared to participants in the community (Ogletree & Katz, 2021). Furthermore, they rate their memory and wellbeing higher than the community sample (Ogletree & Katz, 2021). Cabeza et al. (2002) found that low functioning older adults use only one side of the PFC for memory tasks (similar to young adults); however, high functioning older adults use both sides of the PFC to account for age-related changes in brain physiology. Thus, this will allow high functioning older adults to perform better compared to low functioning older adults (Cabeza et al., 2002; Daselaar & Cabeza, 2005; Marcom et al., 2003). To recap, stereotype threats are only effective when participants feel part of that group (Hess et al., 2003; Pennington et al., 2016; Steele, 1997). Since participants in this sample were likely to be high functioning, they may not have felt part of the stereotyped group. Thus, this could explain why this study did not find any effects of stereotype threat on false memory and why control and stereotyped participants performed similarly for both the memory conjunction and the DRM task.

Moreover, participants from an online platform could have a change in motivation (Chandler et al., 2014). As mentioned above, older adults participate in cognitive tasks because they want to learn more about their cognitive health and maintain a positive view of their cognitive abilities (Ryan & Campbell, 2021). This motivational factor can make older

adults more vulnerable to stereotype threats even in tasks with no threat induced. However, using a sample from an online platform question whether older adults are still participating in these tasks to learn about their cognitive health or are they concerned about the monetary advantages (Chandler et al., 2014; Ogletree & Katz, 2021). Some studies have suggested that there is no difference between recruiting participants from an online platform and the community (Casler et al., 2013; Mortensen & Hughes, 2018), while others have suggested that online platforms might not give an accurate representation of the community (Ogletree & Katz, 2021; Ophir et al., 2019). The generalisability of an online sample remains conflicted. Our study was unable to replicate previous findings despite accounting for several different methodological differences (Smith et al., 2017; Thomas & Dubois, 2011; Wong & Gallo, 2016). Therefore, it is suggested to conduct the same experiment with participants from the community to compare the findings and examine whether online platforms and community samples are different. We attempted to do that by initially recruiting participants from the community. However, COVID interruptions made this difficult.

Number of Words

Another factor that could account for discrepant findings in our study could be the reduced number of words studied in our tasks compared to previous studies (Smith et al., 2017; Thomas & Dubois, 2011; Wong & Gallo, 2016). The recognition test for the DRM task in Experiment 2, had 48 words, compared to 120 words in Smith et al. (2017) and 96 words in Thomas and Dubois' (2011) and Wong and Gallo's (2016) studies. Additionally, the number of words studied was less by 36 words for this experiment. Similarly, the number of words encoded and recognised in our memory conjunction task (in Experiment 1) was also less compared to these studies. It has been found that increasing the number of words studied in the DRM task could help see a clearer effect of false memory (Pardilla-Delgado & Payne, 2017). A study by Robinson and Roediger (1997) showed that there was an increase in false

memory errors with increasing the words in the DRM task. It is possible that with a greater number of words, there is a stronger reliance on gist trace than when fewer words are studied (Robinson & Roediger, 1997). Thus, this increased reliance on gist trace could explain the difference in findings between the current study (specifically Experiment 2) and previous studies (Smith et al., 2017; Thomas & Dubois, 2011; Wong & Gallo, 2016).

To conclude, there were no differences in performance between the two groups (control and stereotype threat) for both tasks. Nevertheless, we also looked at whether moderating factors (e.g., education level, retirement status and importance placed on memory) influenced performance on the memory conjunction task and the DRM task, as previous research has found an effect of these factors on performance (Hess et al., 2003; Smith et al., 2017).

Moderating Factors

It was observed that moderating factors (such as education, employment status, and importance placed on memory) did not influence performance on the memory conjunction and the DRM task for the control and the stereotype threat participants. Our sample did not have much variance in education, with 60% (for Experiments 1 and 2) of our sample completing between 13 and 17 years of education. Prior studies showing an effect of education, such as Smith et al. (2017), had a high criterion for education level; participants who had studied for at least 19 years were categorised in the "high" education category. When comparing this with Thomas and Dubois' (2011) and Wong and Gallo's (2016) studies, at least 45% and 31% of their sample met this criterion. Only 17% of our sample in Experiment 1 and approximately 13% in Experiment 2 met the high criterion of Smith et al. (2017) study. In the Smith et al. (2017) study, older adults were recruited from a pool of participants maintained by the cognitive lab at the University. However, participants in our study were randomly recruited either from the community or online platforms. Mismatch in

samples between studies could be why our study did not find any significant effects of education on performance.

There was no significant difference observed between retired and not retired older adults for the control and stereotype threat groups. Most participants in this study were recruited through an online platform. This means older adults are still active and earning money by completing different tasks (Turner et al., 2020). Furthermore, those participants are actively and comfortably using technology. Stereotype threat is usually perceived when participants can connect with that threat (Hess et al., 2003; Pennington et al., 2016; Steele, 1997). However, in this study, it could be possible that older adults do not think that their memory functioning has declined (as they are consistently participating in different experiments), therefore leading to no difference between the groups.

Similarly, for the importance placed on memory, most participants rated the importance between “7” and “10”. Generally, the community participants that place great importance on memory are affected by the threat to a greater extent as they want to perform well to feel great about their cognitive capacities (Hess et al., 2003; Mazerolle et al., 2021; Ryan & Campbell, 2021). However, as noted above, participants from an online platform might not feel part of the stereotyped group as they are high functioning adults (Ogletree & Katz, 2021) with a change in motivation; therefore, this explains why the current study was unable to find any effects of this moderating factor. Additionally, a lack of variability in the sample with 88.33% (for Experiment 1) and 86.67% (for Experiment 2) of participants rating the importance of memory higher than “7”, could also cause a lack of difference between the two groups.

Limitations and Future Research

There were several limitations in the present study that could be acknowledged in future research. Firstly, this study induced stereotype threat after encoding and before

retrieval. This was similar to previous studies (Smith et al., 2017; Thomas & Dubois, 2011; Wong & Gallo, 2016). Nevertheless, this difference in methodology could have an impact on the findings. For example, Wong and Gallo (2019) investigated whether inducing stereotype threat at encoding (i.e., before studying the test) would yield different findings from it being induced at retrieval (i.e., prior to the memory test). They used an episodic memory task rather than a DRM task and gave participants a stereotype threat paragraph before encoding or retrieval. The effects of stereotype threat on false memory were strongest when the stereotype threat was introduced prior to the encoding phase. Stereotype threat may affect the encoding phase more compared to the retrieval phase. That is, activating stereotype threat during encoding can decrease incidental thoughts regarding the studied words and allow participants to concentrate on thoughts related to the stereotype threat (Wong & Gallo, 2019). Thus, it is suggested that future researchers could induce stereotype threat at encoding and retrieval using a memory conjunction and a DRM task.

The second limitation would be using a complex paragraph to induce stereotype threat. Three studies so far have focused on the effects of age-based stereotype threat on false memory using a DRM task with few methodological changes (Smith et al., 2017; Thomas & Dubois, 2011; Wong & Gallo, 2016). As noted in our introduction, the stereotype threat paragraph in Wong and Gallo's (2016) study was more complex and longer compared to Thomas and Dubois' (2011) and Smith et al. (2017) studies. The current study adapted the stereotype threat paragraph from Wong and Gallo (2016); however, it did not use brain images. A longer and more complicated paragraph would require increased levels of attention compared to a shorter paragraph. Participants from an online platform might not attend fully and read the paragraph (Oppenheimer et al., 2009), thus using a longer paragraph might influence findings. There was an attention question to check if participants were reading the paragraphs; nevertheless, this question was not difficult, and participants could have

answered it by reading the first few lines. Thus, for future research, it is recommended that some participants read a stereotype threat passage from Wong and Gallo (2016) and some from Thomas and Dubois (2011). If there is a difference between the two groups, this will help explain why the current study could not replicate previous findings. Moreover, this would illuminate the importance of not only focusing on individual factors and methodological factors considered in this study but also focusing on what paragraph is used to induce stereotype threat.

The third limitation would be that different participants took part in Experiments 1 and 2; that is, different participants took part in the memory conjunction and the DRM task. If the same participants take part and we see an influence of the threat on one task and not on the other, this provides evidence that these tasks cause false memories in different ways and that some processes might not be affected with age? It is recommended that future studies should integrate both tasks into their methods. The same participants must engage in the memory conjunction task and the DRM task to compare the two tasks better. It will also raise questions about whether using the term false memory to explain false memories with stereotype threat is correct. If there is no correlation between different false memory tasks and different tasks create memory through different mechanisms (Calvillo & Parong, 2016; Nichols & Loftus, 2019), then is it accurate to use the word false memory in general? Instead, should researchers focus on and highlight what aspect of false memory they are studying (Gallo, 2010; Pezdek & Lam, 2007).

It is important to look at how stereotypes could impact the lives of older adults. Age-based stereotypes are present everywhere in our society. From early on in our lives, ageing has been described negatively. A study found that approximately 60% of children think that being old is scary and do not wish to be old (Levy, 2003). Thus, our society has induced stereotypes about old age from a very young age, and older adults are exposed to these

stereotypes in their daily lives. The effects of the threat could show that older adults are in poorer health when they are healthy. To elaborate, Haslam et al. (2012) showed that 70% of older adults failed the cognitive test used to diagnose dementia when in the stereotype threat group. In comparison, only 14% failed when they were not given the stereotype threat manipulation. Furthermore, older adults are regularly exposed to advertisements about dementia, implying that they are old and likely to have dementia (a subtle stereotype). This exposure to threat can cause older adults to be diagnosed with dementia or other cognitive impairments when in fact, it was the stereotype threat (Fresson et al., 2017). Instead of focusing on pharmacology to help with memory and memory errors, it is important to consider these other factors and how they could be changed as a society to help the older adults live actively and confidently in the society.

Even though the current study was unable to replicate findings from previous studies (Smith et al., 2017; Thomas & Dubois, 2011; Wong & Gallo, 2016), it still added to the limited literature on the effects of age-based stereotype threat on false memories and explored this topic with a previously unexplored task (i.e., memory conjunction task). It also raised questions about the validity of an online sample recruited from online platforms. Furthermore, this study highlighted other factors (such as the use of paragraph and the number of words in the list) that could also influence findings and future research might be interested in. Therefore, despite no significant findings, it is important to keep researching and exploring the effects of different factors on the false memory tasks so that there is a better understanding of the effects of age-based stereotype threat on false memories for older adults.

Conclusion

To conclude, this study aimed to account for the methodological differences (e.g., retrieval tasks, tasks focus and effects of moderating factors) that might have cause discrepant findings in the previous literature (focused on the effects of age-based stereotype threat on false memory). This study could not find any effects of stereotype threat on performance for both the memory conjunction and the DRM task. Furthermore, moderating factors did not influence performance for Experiments 1 and 2. Changing the focus of the task also did not affect performance of the stereotyped participants. The lack of significant findings could be due to using an online sample recruited through online platforms and using a task with fewer number of words. Participants from an online platform are usually high functioning, so they might not feel part of the stereotyped group. Therefore, the threat might not influence them and their performance. Nevertheless, despite not finding any significant findings, this study did contribute to the limited literature on the effects of stereotype threat on older adults' false memory by investigating a new task and focusing on different factors (e.g., type of retrieval task, the focus of the task, and the moderating factors) all at once. It also questioned whether using an online sample to investigate false memories is a good idea. It is recommended that future studies should do an in-person study rather than an online study with the same methodology as the present study. A community sample might provide a better understanding of the effects of stereotype threat on false memory performance. A better understanding of this topic will help improve older adults' quality of life.

References

Abadie, M., & Camos, V. (2019). False memory at short and long term. *Journal of Experimental Psychology: General*, *148*(8), 1312-1334. https://doi.org/10.1037/xge0000526

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

Alberini, C. M., & LeDoux, J. E. (2013). Memory reconsolidation. *Current Biology*, *23*(17), R746-R750. https://doi.org/10.1016/j.cub.2013.06.046

Almaraz-Espinoza, A., & Grider, M. H. (2021). Physiology, long term memory. *StatPearls*. https://www.ncbi.nlm.nih.gov/books/NBK549791/

Anderson, N. D., Ebert, P. L., Jennings, J. M., Grady, C. L., Cabeza, R., & Graham, S. J. (2008). Recollection-and familiarity-based memory in healthy aging and amnestic mild cognitive impairment. *Neuropsychology*, *22*(2), 177. https://doi.org/10.1037/0894-4105.22.2.177

Andrews-Hanna, J. R., Snyder, A. Z., Vincent, J. L., Lustig, C., Head, D., Raichle, M. E., & Buckner, R. L. (2007). Disruption of large-scale brain systems in advanced aging. *Neuron*, *56*(5), 924-935. https://doi.org/10.1016/j.neuron.2007.10.038

Barber, S. J. (2017). An examination of age-based stereotype threat about cognitive decline. *Perspectives on Psychological Science : A Journal of the Association for Psychological Science*, *12*(1), 62–90. https://doi.org/10.1177/1745691616656345

Barber, S. J., & Mather, M. (2013). Stereotype threat can both enhance and impair older adults' memory. *Psychological Science*, *24*(12), 2522–2529. https://doi.org/10.1177/0956797613497023

Beato, M. S., Cadavid, S., Pulido, R. F., & Pinho, M. S. (2013). No effect of stress on false

recognition. *Psicothema*, *25*(1), 25-30. https://doi.org/10.7334/psicothema2012.158

Bettio, L. E., Rajendran, L., & Gil-Mohapel, J. (2017). The effects of aging in the hippocampus and cognitive decline. *Neuroscience & Biobehavioral Reviews*, *79*, 66-86. https://doi.org/10.1016/j.neubiorev.2017.04.030

Bharti, A. K., Yadav, S. K., & Jaswal, S. (2020). Feature binding of sequentially presented stimuli in visual working memory. *Frontiers in Psychology*, *11*, 33. https://doi.org/10.3389/fpsyg.2020.00033

Bisaz, R., Travaglia, A., & Alberini, C. M. (2014). The neurobiological bases of memory formation: From physiological conditions to psychopathology. *Psychopathology*, *47*(6), 347–356. https://doi.org/10.1159/000363702

Bouazzaoui, B., Fay, S., Guerrero-Sastoque, L., Semaine, M., Isingrini, M., & Taconnat, L. (2020). Memory age-based stereotype threat: Role of locus of control and anxiety. *Experimental Aging Research*, *46*(1), 39-51. https://doi.org/10.1080/0361073X.2019.1693009

Brainerd, C. J., & Reyna, V. F. (1990). Gist is the grist: Fuzzy-trace theory and the new intuitionism. *Developmental Review*, *10*(1), 3-47. https://doi.org/10.1016/0273-2297(90)90003-M

Brainerd, C. J., & Reyna, V. F. (2002). Fuzzy-trace theory and false memory. *Current Directions in Psychological Science, 11*(5),164–169. https://doi.org/10.1111/1467-8721.00192

Brainerd, C. J., Reyna, V. F., & Ceci, S. J. (2008). Developmental reversals in false memory: A review of data and theory. *Psychological Bulletin*, *134*(3), 343-382. https://doi.org/10.1037/0033-2909.134.3.343

Brelet, L., Moták, L., Ginet, M., Huet, N., Izaute, M., & Gabaude, C. (2016). Enhancing

older drivers' safety: On effects induced by stereotype threat to older adults' driving performance, working memory and self-regulation. *Geriatrics*, *1*(3), 20. https://doi.org/10.3390/geriatrics1030020

Cabeza, R., Anderson, N. D., Locantore, J. K., & McIntosh, A. R. (2002). Aging gracefully: Compensatory brain activity in high-performing older adults. *Neuroimage*, *17*(3), 1394-1402. https://doi.org/10.1006/nimg.2002.1280

Calvillo, D. P., & Parong, J. A. (2016). The misinformation effect is unrelated to the DRM effect with and without a DRM warning. *Memory*, *24*(3), 324-333. https://doi.org/10.1080/09658211.2015.1005633

Cameron, H. A., Woolley, C. S., McEwen, B. S., & Gould, E. (1993). Differentiation of newly born neurons and glia in the dentate gyrus of the adult rat. *Neuroscience*, *56*(2), 337-344. https://doi.org/10.1016/0306-4522(93)90335-D

Campos, T. F., Barroso, M. T. M., & de Lara Menezes, A. A. (2010). Encoding, storage and retrieval processes of the memory and the implications for motor practice in stroke patients. *NeuroRehabilitation*, *26*(2), 135-142. https://doi.org/10.3233/NRE-2010-0545

Casler, K., Bickel, L., & Hackett, E. (2013). Separate but equal? A comparison of participants and data gathered via Amazon's MTurk, social media, and face-to-face behavioral testing. *Computers in Human Behavior*, *29*(6), 2156–2160. https://doi.org/10.1016/j.chb.2013.05.009

Castel, A. D., & Craik, F. I. (2003). The effects of aging and divided attention on memory for item and associative information. *Psychology and Aging*, *18*(4), 873-885. https://doi.org/10.1037/0882-7974.18.4.873

Chandler, J., Mueller, P., & Paolacci, G. (2014). Nonnaïveté among Amazon Mechanical

Turk workers: Consequences and solutions for behavioral researchers. *Behavior Research Methods, 46,* 112–130. https://doi.org/10.3758/s13428-013-0365-7

Chasteen, A. L., Bhattacharyya, S., Horhota, M., Tam, R., & Hasher, L. (2005). How feelings of stereotype threat influence older adults' memory performance. *Experimental Aging Research*, *31*(3), 235-260. https://doi.org/10.1080/03610730590948177

Coane, J. H., McBride, D. M., Huff, M. J., Chang, K., Marsh, E. M., & Smith, K. A. (2021). Manipulations of list type in the DRM paradigm: A review of how structural and conceptual similarity affect false memory. *Frontiers in Psychology*, *12*, 668550. https://doi.org/10.3389/fpsyg.2021.668550

Crowe, E., & Higgins, E. T. (1997). Regulatory focus and strategic inclinations: Promotion and prevention in decision-making. *Organizational Behavior and Human Decision Processes*, *69*(2), 117-132. https://doi.org/10.1006/obhd.1996.2675

Cyr, A. A., Romero, K., & Galin-Corini, L. (2021). Web-based cognitive testing of older adults in person versus at home: Within-subjects comparison study. *JMIR Aging*, *4*(1), e23384. https://doi.org/10.2196/23384

Daselaar, S., & Cabeza, R. (2005). Age-related changes in hemispheric organization. *Cognitive Neuroscience of Aging: Linking Cognitive and Cerebral Aging* (pp. 325-353). Oxford University Press.

Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology*, *58*(1), 17-22. https://doi.org/10.1037/h0046671

Desai, A. K., Grossberg, G. T., & Sheth, D. N. (2004). Activities of daily living in patients with dementia. *CNS Drugs*, *18*(13), 853-875. https://doi.org/10.2165/00023210-200418130-00003

Devitt, A. L., Monk-Fromont, E., Schacter, D. L., & Addis, D. R. (2016). Factors that

influence the generation of autobiographical memory conjunction errors. *Memory*, *24*(2), 204-222. https://doi.org/10.1080/09658211.2014.998680

Devitt, A. L., & Schacter, D. L. (2016). False memories with age: Neural and cognitive underpinnings. *Neuropsychologia*, *91*, 346–359. https://doi.org/10.1016/j.neuropsychologia.2016.08.030

Devitt, A. L., Tippett, L., Schacter, D. L., & Addis, D. R. (2016). Autobiographical memory conjunction errors in younger and older adults: Evidence for a role of inhibitory ability. *Psychology and Aging*, *31*(8), 927–942. https://doi.org/10.1037/pag0000129

Dodson, C. S., & Schacter, D. L. (2002). The cognitive neuropsychology of false memories: Theory and data. In A. D. Baddeley, M. D. Kopelman, & B. A. Wilson (Eds), *Handbook of Memory Disorders* (2, pp. 343-362).

Duarte, A., Graham, K. S., & Henson, R. N. (2010). Age-related changes in neural activity associated with familiarity, recollection and false recognition. *Neurobiology of Aging*, *31*(10), 1814-1830. https://doi.org/10.1016/j.neurobiolaging.2008.09.014

Falzarano, F., & Siedlecki, K. L. (2019). Investigating the relations among different measures of false memory. *Advances in Cognitive Psychology*, *15*(4), 290–300. https://doi.org/10.5709/acp-0276-0

Fandakova, Y., Shing, Y. L., & Lindenberger, U. (2013). High-confidence memory errors in old age: The roles of monitoring and binding processes. *Memory*, *21*(6), 732-750. https://doi.org/10.1080/09658211.2012.756038

Fossati, P. (2013). Imaging autobiographical memory. *Dialogues in Clinical Neuroscience*, *15*(4), 487–490. https://doi.org/10.31887/DCNS.2013.15.4/pfossati

Fourquet, N. Y., Patterson, T. K., Li, C., Castel, A. D., & Knowlton, B. J. (2020). Effects of age-related stereotype threat on metacognition. *Frontiers in Psychology*, 3453. https://doi.org/10.3389/fpsyg.2020.604978

Fresson, M., Dardenne, B., Geurten, M., & Meulemans, T. (2017). The effect of stereotype threat on older people's clinical cognitive outcomes: Investigating the moderating role of dementia worry. *The Clinical Neuropsychologist*, *31*(8), 1306-1328. https://doi.org/10.1080/13854046.2017.1307456

Freund, A. M., & Baltes, P. B. (1998). Selection, optimization, and compensation as strategies of life management: Correlations with subjective indicators of successful aging. *Psychology and Aging*, *13*(4), 531-543. https://doi.org/10.1037/0882-7974.13.4.531

Gallo, D. A. (2010). False memories and fantastic beliefs: 15 years of the DRM illusion. *Memory & Cognition*, *38*(7), 833-848. https://doi.org/10.3758/MC.38.7.833

Gilet, A. L., Evrard, C., Galharret, J. M., & Colombel, F. (2021). The moderating role of education on the relationship between perceived stereotype threat and false memory in aging. *Frontiers in Psychology*, *11*, 3860. https://doi.org/10.3389/fpsyg.2020.606249

Giovanello, K. S., Kensinger, E. A., Wong, A. T., & Schacter, D. L. (2010). Age-related neural changes during memory conjunction errors. *Journal of Cognitive Neuroscience*, *22*(7),1348–1361. https://doi.org/10.1162/jocn.2009.21274

Haslam, C., Morton, T. A., Haslam, S. A., Varnes, L., Graham, R., & Gamaz, L. (2012). "When the age is in, the wit is out": Age-related self-categorization and deficit expectations reduce performance on clinical tests used in dementia assessment. *Psychology and Aging*, *27*(3), 778-784. https://doi.org/10.1037/a0027754

Henkel, L. A. (2014). Do older adults change their eyewitness reports when re-questioned? *Journals of Gerontology: Series B*, *69*(3), 356-365. https://doi.org/10.1093/geronb/gbt071

Hess, T. M., Auman, C., Colcombe, S. J., & Rahhal, T. A. (2003). The impact of stereotype

threat on age differences in memory performance. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *58*(1), 3-11. https://doi.org/10.1093/geronb/58.1.P3

Hess, T. M., Hinson, J. T., & Hodges, E. A. (2009). Moderators of and mechanisms underlying stereotype threat effects on older adults' memory performance. *Experimental Aging Research*, *35*(2), 153–177. https://doi.org/10.1080/03610730802716413

Hommel, B., Kray, J., & Lindenberger, U. (2011). Feature integration across the lifespan: Stickier stimulus–response bindings in children and older adults. *Frontiers in Psychology*, *2*, 268. https://doi.org/10.3389/fpsyg.2011.00268

Hou, J., Wu, Y., & Harrell, E. (2017). Reading on paper and screen among senior adults: Cognitive map and technophobia. *Frontiers in Psychology*, *8*, 2225. https://doi.org/10.3389/fpsyg.2017.02225

Howe, M. L., & Garner, S. R. (2018). Can false memories prime alternative solutions to ambiguous problems?. *Memory*, *26*(1), 96-105. https://doi.org/10.1080/09658211.2017.1332226

Howe, M. L., Garner, S. R., & Patel, M. (2013). The positive consequences of false memories. *Behavioral Sciences and the Law*, *31*(5), 652-665. https://doi.org/10.1002/bsl.2078

Howe, M. L., Knott, L. M., & Conway, M. A. (2018). *Memory and miscarriages of justice* (1). Psychology Press. https://doi.org/10.4324/9781315752181

Hultsch, D. F., MacDonald, S. W., Hunter, M. A., Maitland, S. B., & Dixon, R. A. (2002). Sampling and generalisability in developmental research: Comparison of random and convenience samples of older adults. *International Journal of Behavioral Development*, *26*(4), 345-359. https://doi.org/10.1080/01650250143000247

Jacoby, L. L., & Rhodes, M. G. (2006). False remembering in the aged. *Current Directions in Psychological Science*, *15*(2), 49-53. https://doi.org/10.1111/j.0963-7214.2006.00405.x

Jeye, B. M., Karanian, J. M., & Slotnick, S. D. (2017). The anterior prefrontal cortex and the hippocampus are negatively correlated during false memories. *Brain Sciences*, *7*(1), 13. https://doi.org/10.3390/brainsci7010013

Joanisse, M., Gagnon, S., & Voloaca, M. (2013). The impact of stereotype threat on the simulated driving performance of older drivers. *Accident Analysis and Prevention*, *50*, 530–538. https://doi.org/10.1016/j.aap.2012.05.032

Johns, M., Inzlicht, M., & Schmader, T. (2008). Stereotype threat and executive resource depletion: Examining the influence of emotion regulation. *Journal of Experimental Psychology: General*, *137*(4), 691–705. https://doi.org/10.1037/a0013834

Jones, T. C. (2005). Study repetition and the rejection of conjunction lures. *Memory*, *13*(5), 499-515. https://doi.org/10.1080/09658210444000197

Jones, T. C., & Atchley, P. (2006). Conjunction errors, recollection-based rejections, and forgetting in a continuous recognition task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*(1), 70-78. https://doi.org/10.1037/0278-7393.32.1.70

Jones, T. C., & Jacoby, L. L. (2001). Feature and conjunction errors in recognition memory: Evidence for dual-process theory. *Journal of Memory and Language*, *45*(1), 82-102. https://doi.org/10.1006/jmla.2000.2761

Kang, S. K., & Chasteen, A. L. (2009). The moderating role of age-group identification and perceived threat on stereotype threat among older adults. *The International Journal of Aging and Human Development*, *69*(3), 201-220. https://doi.org/10.2190/AG.69.3.c

Kensinger, E. A., & Schacter, D. L. (1999). When true memories suppress false memories: Effects of ageing. *Cognitive Neuropsychology*,*16*(3-5), 399-415. https://doi.org/10.1080/026432999380852

Kroll, N. E., Knight, R. T., Metcalfe, J., Wolf, E. S., & Tulving, E. (1996). Cohesion failure as a source of memory illusions. *Journal of Memory and Language*, *35*(2), 176-196. https://doi.org/10.1006/jmla.1996.0010

Lambert, A. E., Watson, J. M., Stefanucci, J. K., Ward, N., Bakdash, J. Z., & Strayer, D. L. (2016). Stereotype threat impairs older adult driving. *Applied Cognitive Psychology*, *30*(1), 22-28. https://doi.org/10.1002/acp.3162

Lamont, R. A., Swift, H. J., & Abrams, D. (2015). A review and meta-analysis of age-based stereotype threat: Negative stereotypes, not facts, do the damage. *Psychology and Aging*, *30*(1), 180–193. https://doi.org/10.1037/a0038586

Leding, J. K., & Lampinen, J. M. (2009). Memory conjunction errors: The effects of presentation duration and study repetition, *Memory*, *17*(5), 597-607. https://doi.org/10.1080/09658210902984518

Lee, J. L. (2009). Reconsolidation: Maintaining memory relevance. *Trends in Neurosciences*, *32*(8), 413-420. https://doi.org/10.1016/j.tins.2009.05.002

Levy, B. R. (2003). Mind matters: Cognitive and physical effects of aging self-stereotypes. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *58*(4), P203-P211. https://doi.org/10.1093/geronb/58.4.P203

Levy, B. R., Zonderman, A. B., Slade, M. D., & Ferrucci, L. (2012). Memory shaped by age stereotypes over time. *The Journals of Gerontology: Series B*, *67*(4), 432-436. https://doi.org/10.1093/geronb/gbr120

Loftus, E. F., & Hoffman, H. G. (1989). Misinformation and memory: The creation of new

memories. *Journal of Experimental Psychology: General*, *118*(1), 100-104. https://doi.org/10.1037/0096-3445.118.1.100

Mammarella, N., & Fairfield, B. (2008). Source monitoring: The importance of feature binding at encoding. *European Journal of Cognitive Psychology*, *20*(1), 91-122. https://doi.org/10.1080/09541440601112522

Marder, E. (2017). The importance of remembering. *eLife*, *6*, e30599. https://doi.org/10.7554/eLife.30599

Markman, A. B., Baldwin, G. C., & Maddox, W. T. (2005). The interaction of payoff structure and regulatory focus in classification. *Psychological Science*, *16*(11), 852-855. https://doi.org/10.1111/j.1467-9280.2005.01625.x

Marx, D. M., & Goff, P. A. (2005). Clearing the air: The effect of experimenter race on target's test performance and subjective experience. *British Journal of Social Psychology*, *44*(4), 645-657. https://doi.org/10.13480/014466604X17948

Matzen, L. E., & Benjamin, A. S. (2009). Remembering words not presented in sentences: How study context changes patterns of false memories. *Memory & Cognition*, *37*(1), 52-64. https://doi.org/10.3758/MC.37.1.52

Mazerolle, M., Smith, A. M., Torrance, M., & Thomas, A. K. (2021). Understanding older adults' memory distortion in the light of stereotype threat. *Frontiers in Psychology*, *12*, 656. https://doi.org/10.3389/fpsyg.2021.628696

Mitchell, K. J., & Johnson, M. K. (2009). Source monitoring 15 years later: What have we learned from fMRI about the neural mechanisms of source memory? *Psychological Bulletin*, *135*(4), 638–677. https://doi.org/10.1037/a0015849

Morcom, A. M., Good, C. D., Frackowiak, R. S., & Rugg, M. D. (2003). Age effects on the neural correlates of successful memory encoding. *Brain*, *126*(1), 213-229. https://doi.org/10.1093/brain/awg020

Mortensen, K., & Hughes, T. L. (2018). Comparing Amazon's Mechanical Turk platform to conventional data collection methods in the health and medical research literature. *Journal of General Internal Medicine*, *33*(4), 533–538. https://doi.org/10.1007/s11606-017-4246-0

Newman, E. J., & Lindsay, D. S. (2009). False memories: What the hell are they for? *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, *23*(8), 1105-1121. https://doi.org/10.1002/acp.1613

Nichols, R. M., & Loftus, E. F. (2019). Who is susceptible in three false memory tasks? *Memory*, *27*(7), 962-984. https://doi.org/10.1080/09658211.2019.1611862

Nyberg, L., Lövdén, M., Riklund, K., Lindenberger, U., & Bäckman, L. (2012). Memory aging and brain maintenance. *Trends in Cognitive Sciences*, *16*(5), 292-305. https://doi.org/10.1016/j.tics.2012.04.005

Odegard, T. N., & Lampinen, J. M. (2004). Memory conjunction errors for autobiographical events: More than just familiarity. *Memory*, *12*(3), 288-300. https://doi.org/10.1080/09658210244000621

Ogletree, A. M., & Katz, B. (2021). How do older adults recruited using MTurk differ from those in a national probability sample? *The International Journal of Aging and Human Development*, *93*(2), 700-721. https://doi.org/10.1177/0091415020940197

Olsen, R. K., Moses, S. N., Riggs, L., & Ryan, J. D. (2012). The hippocampus supports multiple cognitive processes through relational binding and comparison. *Frontiers in Human Neuroscience*, *6*, 146. https://doi.org/10.3389/fnhum.2012.00146

Ophir, Y., Sisso, I., Asterhan, C., Tikochinski, R., & Reichart, R. (2019). The turker blues: Hidden factors behind increased depression rates in Amazon's Mechanical Turk. *Clinical Psychological Science*, *8*(1), 65–83.

https://doi.org/10.1177/2167702619865973

Oppenheimer, D. M., Meyvis, T., & Davidenko, N. (2009). Instructional manipulation checks: Detecting satisficing to increase statistical power. *Journal of Experimental Social Psychology*, *45*(4), 867-872. https://doi.org/10.1016/j.jesp.2009.03.009

Ossher, L., Flegal, K. E., & Lustig, C. (2013). Everyday memory errors in older adults. *Aging, Neuropsychology, and Cognition*, *20*(2), 220–242. https://doi.org/10.1080/13825585.2012.690365

Ost, J., Blank, H., Davies, J., Jones, G., Lambert, K., & Salmon, K. (2013). False memory ≠ false memory: DRM errors are unrelated to the misinformation effect. *PloS One*, *8*(4), e57939. https://doi.org/10.1371/journal.pone.0057939

Otgaar, H., Howe, M. L., van Beers, J., van Hoof, R., Bronzwaer, N., & Smeets, T. (2015). The positive ramifications of false memories using a perceptual closure task. *Journal of Applied Research in Memory and Cognition*, *4*(1), 43-50. https://doi.org/10.1016/j.jarmac.2014.12.001

Otgaar, H., Muris, P., Howe, M. L., & Merckelbach, H. (2017). What drives false memories in psychopathology? A case for associative activation. *Clinical Psychological Science: A Journal of the Association for Psychological Science*, *5*(6), 1048–1069. https://doi.org/10.1177/2167702617724424

Paige, L. E., Cassidy, B. S., Schacter, D. L., & Gutchess, A. H. (2016). Age differences in hippocampal activation during gist-based false recognition. *Neurobiology of Aging*, *46*, 76–83. https://doi.org/10.1016/j.neurobiolaging.2016.06.014

Pardilla-Delgado, E., & Payne, J. D. (2017). The Deese-Roediger-McDermott (DRM) task: A simple cognitive paradigm to investigate false memories in the laboratory. *Journal of Visualized Experiments*, *119*, 54793. https://doi.org/10.3791/54793

Payne, J. D., Nadel, L., Allen, J. J., Thomas, K. G., & Jacobs, W. J. (2002). The effects of

experimentally induced stress on false recognition. *Memory*, *10*(1), 1-6. https://doi.org/10.1080/09658210143000119

Pennington, C. R., Heim, D., Levy, A. R., & Larkin, D. T. (2016). Twenty years of stereotype threat research: A review of psychological mediators. *PloS One*, *11*(1), e0146487. https://doi.org/10.1371/journal.pone.0146487

Pezdek, K., & Lam, S. (2007). What research paradigms have cognitive psychologists used to study ‘false memory,' and what are the implications of these choices? *Consciousness and Cognition*, *16*(1), 2–17. https://doi.org/10.1016/j.concog.2005.06.006

Preston, A. R., & Eichenbaum, H. (2013). Interplay of hippocampus and prefrontal cortex in memory. *Current Biology*, *23*(17), R764–R773. https://doi.org/10.1016/j.cub.2013.05.041

Quinn, D. M., & Spencer, S. J. (2001). The interference of stereotype threat with women's generation of mathematical problem-solving strategies. *Journal of Social Issues*, *57*(1), 55-71. https://doi.org/10.1111/0022-4537.00201

Rahhal, T. A., Hasher, L., & Colcombe, S. J. (2001). Instructional manipulations and age differences in memory: Now you see them, now you don't. *Psychology and Aging*, *16*(4), 697-706. https://doi.org/10.1037/0882-7974.16.4.697

Ranganath, C. (2010). Binding items and contexts: The cognitive neuroscience of episodic memory. *Current Directions in Psychological Science*, *19*(3), 131-137. https://doi.org/10.1177/0963721410368805

Reinitz, M. T., Lammers, W. J., & Cochran, B. P. (1992). Memory-conjunction errors: Miscombination of stored stimulus features can produce illusions of memory. *Memory & Cognition*, *20*(1), 1-11. https://doi.org/10.3758/BF03208247

Reyna, V. F., & Brainerd, C. J. (1995). Fuzzy-trace theory: An interim synthesis. *Learning and Individual Differences*, *7*(1), 1–75.

https://doi.org/10.1016/1041-6080(95)90031-4

Reyna, V. F., Corbin, J. C., Weldon, R. B., & Brainerd, C. J. (2016). How fuzzy-trace theory predicts true and false memories for words, sentences, and narratives. *Journal of Applied Research in Memory and Cognition*, *5*(1), 1–9. https://doi.org/10.1016/j.jarmac.2015.12.003

Riley, R. J., Burgener, S., & Buckwalter, K. C. (2014). Anxiety and stigma in dementia: A threat to aging in place. *The Nursing Clinics of North America*, *49*(2), 213–231. https://doi.org/10.1016/j.cnur.2014.02.008

Robertson, L. T. (2002). Memory and the brain. *Journal of Dental Education*, *66*(1), 30-42. https://doi.org/10.1002/j.0022-0337.2002.66.1.tb03506.x

Robinson, K. J., & Roediger, H. L. (1997). Associative processes in false recall and false recognition. *Psychological Science*, *8*(3), 231-237. https://doi.org/10.1111/j.1467-9280.1997.tb00417.x

Roediger, H. L., Balota, D. A., & Watson, J. M. (2001). Spreading activation and arousal of false memories. In H. L. Roediger, J. S. Nairne, I. Neath, & A. M. Surprenant (Eds.), *The Nature of Remembering* (pp. 95-115). https://doi.org/10.1037/10394-006

Roediger, H. L., & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(4), 803-814. https://doi.org/10.1037/0278-7393.21.4.803

Rosa, M., & Gutchess, A. H. (2013). False memory in aging resulting from self-referential processing. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *68*(6), 882-892. https://doi.org/10.1093/geronb/gbt018

Ryan, A. D., & Campbell, K. L. (2021). The ironic effect of older adults' increased task motivation: Implications for neurocognitive aging. *Psychonomic Bulletin & Review*, *28*, 1743-1754. https://doi.org/10.3758/s13423-021-01963-4

Schacter, D. L. (2013). Memory: Sins and virtues. *Annals of the New York Academy of Sciences*, *1303*(1), 56–60. https://doi.org/10.1111/nyas.12168

Schacter, D. L., & Slotnick, S. D. (2004). The cognitive neuroscience of memory distortion. *Neuron*, *44*(1), 149-160. https://doi.org/10.1016/j.neuron.2004.08.017

Schmader, T., Johns, M., & Forbes, C. (2008). An integrated process model of stereotype threat effects on performance. *Psychological Review*, *115*(2), 336–356. https://doi.org/10.1037/0033-295X.115.2.336

Scholl, J. M., & Sabat, S. R. (2008). Stereotypes, stereotype threat and ageing: Implications for the understanding and treatment of people with Alzheimer's disease. *Ageing & Society*, *28*(1), 103-130. https://doi.org/10.1017/S0144686X07006241

Seibt, B., & Förster, J. (2004). Stereotype threat and performance: How self-stereotypes influence processing by inducing regulatory foci. *Journal of Personality and Social Psychology*, *87*(1), 38-56. https://doi.org/10.1037/0022-3514.87.1.38

Shapiro, J. R., & Neuberg, S. L. (2007). From stereotype threat to stereotype threats: Implications of a multi-threat framework for causes, moderators, mediators, consequences, and interventions. *Personality and Social Psychology Review*, *11*(2), 107-130. https://doi.org/10.1177/1088868306294790

Smith, A. M., Gallo, D. A., Barber, S. J., Maddox, K. B., & Thomas, A. K. (2017). Stereotypes, warnings, and identity-related variables influence older adults' susceptibility to associative false memory errors. *The Gerontologist*, *57*(suppl_2), S206-S215. https://doi.org/10.1093/geront/gnx057

Solbakk, A. K., Alpert, G. F., Furst, A. J., Hale, L. A., Oga, T., Chetty, S., Pickard, N., & Knight, R. T. (2008). Altered prefrontal function with aging: Insights into age-associated performance decline. *Brain Research*, *1232*, 30–47. https://doi.org/10.1016/j.brainres.2008.07.060

Spencer, W. D., & Raz, N. (1995). Differential effects of aging on memory for content and context: A meta-analysis. *Psychology and Aging, 10*(4), 527–539. https://doi.org/10.1037//0882-7974.10.4.527.

Spencer, S. J., Steele, C. M., & Quinn, D. M. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology, 35*(1), 4-28. https://doi.org/10.1006/jesp.1998.1373

Steele, C. M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist, 52*(6), 613-629. https://doi.org/10.1037/0003-066X.52.6.613

Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology, 69*(5), 797-811. https://doi.org/10.1037/0022-3514.69.5.797

Stern, S. A., & Alberini, C. M. (2013). Mechanisms of memory enhancement. *Wiley Interdisciplinary Reviews Systems Biology and Medicine, 5*(1), 37–53. https://doi.org/10.1002/wsbm.1196

Stone, J., & McWhinnie, C. (2008). Evidence that blatant versus subtle stereotype threat cues impact performance through dual processes. *Journal of Experimental Social Psychology*, *44*(2), 445-452. https://doi.org/10.1016/j.jesp.2007.02.006

Straube, B. (2012). An overview of the neuro-cognitive processes involved in the encoding, consolidation, and retrieval of true and false memories. *Behavioral and Brain Functions, 8*, 35. https://doi.org/10.1186/1744-9081-8-35

Strickland-Hughes, C. M., & West, R. L. (2021). The impact of naturalistic age stereotype activation. *Frontiers in Psychology*, *12*, 685448. https://doi.org/10.3389/fpsyg.2021.685448

Squire, L. R., Stark, C. E., & Clark, R. E. (2004). The medial temporal lobe. *Annual Review*

of Neuroscience, 27(1), 279-306. https://doi.org/10.1146/annurev.neuro.27.070203.144130

Squire, L. R., Genzel, L., Wixted, J. T., & Morris, R. G. (2015). Memory consolidation. *Cold Spring Harbor Perspectives in Biology*, *7*(8), a021766. https://doi.org/10.1101/cshperspect.a021766

Thomas, A. K., & Dubois, S. J. (2011). Reducing the burden of stereotype threat eliminates age differences in memory distortion. *Psychological Science*, *22*(12), 1515-1517. https://doi.org/10.1177/0956797611425932

Thomas, A. K., Smith, A. M., & Mazerolle, M. (2020). The unexpected relationship between retrieval demands and memory performance when older adults are faced with age-related stereotypes. *The Journals of Gerontology: Series B, 75*(2), 241-250. https://doi.org/10.1093/geronb/gby031

Thomas, A. K., & Sommers, M. S. (2005). Attention to item-specific processing eliminates age effects in false memories. *Journal of Memory and Language*, *52*(1), 71-86. https://doi.org/10.1016/j.jml.2004.08.001

Tisserand, D. J., van Boxtel, M. P., Pruessner, J. C., Hofman, P., Evans, A. C., & Jolles, J. (2004). A voxel-based morphometric study to determine individual differences in gray matter density associated with age and cognitive change over time. *Cerebral Cortex*, *14*(9), 966-973. https://doi.org/10.1093/cercor/bhh057

Tun, P. A., Wingfield, A., Rosen, M. J., & Blanchard, L. (1998). Response latencies for false memories: Gist-based processes in normal aging. *Psychology and Aging*, *13*(2), 230-241. https://doi.org/10.1037/0882-7974.12.2.230

Turner, A. M., Engelsma, T., Taylor, J. O., Sharma, R. K., & Demiris, G. (2020). Recruiting

older adult participants through crowdsourcing platforms: Mechanical Turk versus Prolific Academic. *AMIA Annual Symposium Proceedings Archive*, 1230-1238. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8075523/

Underwood, B. J., & Zimmerman, J. (1973). The syllable as a source of error in multisyllable word recognition. *Journal of Verbal Learning and Verbal Behavior*, *12*(6), 701-706. https://doi.org/10.1016/S0022-5371(73)80050-7

Wang, J., Otgaar, H., Smeets, T., Howe, M. L., Merckelbach, H., & Zhou, C. (2018). Consequences of false memories in eyewitness testimony: A review and implications for Chinese legal practice. *Psychological Research on Urban Society*, *1*(1), 12-25. https://doi.org/10.7454/proust.v1i1.15

Wilson, I. A., Gallagher, M., Eichenbaum, H., & Tanila, H. (2006). Neurocognitive aging: Prior memories hinder new hippocampal encoding. *Trends in Neurosciences*, *29*(12), 662-670. https://doi.org/10.1016/j.tins.2006.10.002

Wong, J. T., & Gallo, D. A. (2016). Stereotype threat reduces false recognition when older adults are forewarned. *Memory*, *24*(5), 650-658. https://doi.org/10.1080/09658211.2015.1036885

Wong, J. T., & Gallo, D. A. (2019). Activating aging stereotypes increases source recollection confusions in older adults: Effect at encoding but not retrieval. *The Journals of Gerontology: Series B*, *74*(4), 633-641. https://doi.org/10.1093/geronb/gbx103

Wylie, L. E., Patihis, L., McCuller, L. L., Davis, D., Brank, E., Loftus, E. F., & Bornstein, B. (2014). Misinformation effect in older versus younger adults: A meta-analysis and review. *The Elderly Eyewitness in Court*, 38-66.

Yassa, M. A., Lacy, J. W., Stark, S. M., Albert, M. S., Gallagher, M., & Stark, C. E. (2011).

Pattern separation deficits associated with increased hippocampal CA3 and dentate gyrus activity in nondemented older adults. *Hippocampus*, *21*(9), 968–979. https://doi.org/10.1002/hipo.20808

Yassa, M. A., & Stark, C. E. (2011). Pattern separation in the hippocampus. *Trends in Neurosciences*, *34*(10), 515–525. https://doi.org/10.1016/j.tins.2011.06.006

Yee, E., Chrysikou, E. G., & Thompson-Schill, S. L. (2013). Semantic memory. In K. Ochsner & S. Kosslyn (Eds.), *The Oxford Handbook of Cognitive Neuroscience,* (pp. 353-374).

Zanto, T. P., & Gazzaley, A. (2019). Aging of the frontal lobe. In M. D'Esposito & J. Grafman (Eds.), *Handbook of Clinical Neurology* (*163*, pp. 369-389). https://doi.org/10.1016/B978-0-12-804281-6.00020-3

Zhu, B., Chen, C., Loftus, E. F., Lin, C., & Dong, Q. (2013). The relationship between DRM and misinformation false memories. *Memory & Cognition*, *41*(6), 832-838. https://doi.org/10.3758/s13421-013-0300-2

Zlotnik, G., & Vansintjan, A. (2019). Memory: An extended definition. *Frontiers in Psychology*, *10*, 2523. https://doi.org/10.3389/fpsyg.2019.02523

Zoladz, P. R., Peters, D. M., Kalchik, A. E., Hoffman, M. M., Aufdenkampe, R. L., Woelke, S. A., Wolters, N. E., & Talbot, J. N. (2014). Brief, pre-learning stress reduces false memory production and enhances true memory selectively in females. *Physiology & Behavior*, *128*, 270-276. https://doi.org/10.1016/j.physbeh.2014.02.028

Appendices

Appendix A.1: Flyer Used for Recruitment

Participants needed
for online research study

If you are **aged 65+ years**, we invite you to complete a brief online study to help us learn more about how ageing affects our memory.

This project is being carried out by Dr Aleea Devitt and Masters student Sukhan Ubha (School of Psychology, University of Waikato).

In this study, you will view words and will be asked to remember them. Your memory will be tested for those items. This study takes approx. **10 minutes**. As a thank you, you will be able to enter in a **draw to win** one of three **$20 Warehouse vouchers.**

Interested? Please use the link below to access our study.
https://waikato.qualtrics.com/jfe/form/SV_2bfUJ4DvYnAspf0

Questions? Please contact Sukhan at waikatoaginglab@gmail.com

This research project has been approved by the Human Research Ethics Committee of the Division of Arts, Law, Psychology, and Social Sciences at the University of Waikato as FS2020-45. Any questions or concerns about the ethical conduct of this research may be sent to the Secretary of the Committee, email alpss-ethics@waikato.ac.nz.

Appendix A.2: Paragraphs

Stereotype threat paragraph: It is a common belief that memory declines with old age. Adults over age 60 report more memory problems in everyday life than do younger adults, even for individuals in good health. Older adults also report less perceived control over their own memories. Such beliefs about age-related memory declines are not groundless. Older adults perform worse than younger adults on standardised memory tests, as well as on laboratory tests of both recall and recognition. Furthermore, age-related declines in memory cannot be attributed to the artificiality of laboratory tasks, because older adults perform more poorly than younger adults on tasks that are designed to simulate "real life", such as remembering information on medicine labels, faces of people or layouts of museums they have visited (Bjork & Bjork, 1996, pp. 443–444). Neuroscience research confirms these findings. Using a brain imaging technique called MRI (magnetic resonance imaging), older adults showed more brain atrophy than younger adults, particularly in the hippocampus. The hippocampus is linked to memory, so the degeneration of this structure as well as general neural atrophy in older adults have been linked to their impairments in memory and cognition. One area of ageing impairment is memory for recent lists. Older adults are impaired in their ability to recall words presented in a list, even if these words are associated or related to each other (Butler, McDaniel, McCabe, & Dornburg, 2010). Thus, older adults show impairments in their ability to remember specific words.

Control Paragraph: Language is our primary means of communicating thought, and, it is universal: every human society has a language, and every human being of normal intelligence acquires his or her native language and uses it effortlessly. Virtually everyone can master and use an enormously complex linguistic system (Smith, Nolen-Hoeksema, Fredrickson, & Loftus, 2003, pp. 314–315). While language itself varies from culture to culture, there are commonalities that connect all languages: sentences are built from words

and words are constructed from speech sounds. Words are the building blocks for communicating ideas, and the most important aspect of a word is its meaning. Using a brain imaging technique called fMRI (functional magnetic resonance imaging), researchers have discovered that many different brain regions are associated with language processing. These findings are consistent with the idea that language is involved in a variety of cognitive processes, and this is true of all individuals. One area of cognition where language skills obviously play a large role is meaning. For example, objects have meaningful qualities to them, such as perceived likeability and pleasantness. These impressions are influenced by the words that are used to describe the objects, and the context of these descriptions that is often conveyed through language (Pilling & Davies, 2004). People all over the world form impressions of various things based on what they hear, and although there are regional differences in the surface forms of words, all languages use words to convey a core set of meanings that seem to be similar across cultures.

Appendix A.3: Experiment 1 Memory Conjunction Task Lists

List Studied

Final study list	Label
iceberg	Old
battleaxe	Old
toothpick	Old
snowplow	Old
dragonfly	Old
duct tape	Old
jailbird	Old
thumbnail	Old
watchtower	Parent
forefather	Parent
tablecloth	Parent
chickenpox	Parent
peacetime	Parent
snakeskin	Parent
tailspin	Parent
lifesaver	Parent
sheepdog	Parent
hindsight	Parent
soupspoon	Parent
lily livered	Parent
pacemaker	Parent
frostbite	Parent
floodgate	Parent
safeguard	Parent
steamroller	Feature parent A
hairbrush	Feature parent A
seaworthy	Feature parent A
sand dune	Feature parent A
whiplash	Feature parent B
towrope	Feature parent B
swordfish	Feature parent B
hailstorm	Feature parent B
ladylike	Old
childbirth	Old
plaything	Old
doorstep	Old
toadstool	Old
popcorn	Old
fanfare	Old
gemstone	Old
overlap	Parent
anywhere	Parent
stopgap	Parent
nutcracker	Parent
armpit	Parent
storekeeper	Parent
sharecropper	Parent
speed bump	Parent
topcoat	Parent
everybody	Parent
yield sign	Parent
clamshell	Parent
wheelchair	Parent
greenhouse	Parent
potholder	Parent
mousetrap	Parent
goldmine	Feature parent A
drawbridge	Feature parent A
stockyard	Feature parent A
jackhammer	Feature parent A
bedroom	Feature parent B
policeman	Feature parent B
silkworm	Feature parent B
stovepipe	Feature parent B

Recognition Test List

Final test list	Label
iceberg	Old
battleaxe	Old
toothpick	Old
snowplow	Old
dragonfly	Old
duct tape	Old
jailbird	Old
thumbnail	Old
ladylike	Old
childbirth	Old
plaything	Old
doorstep	Old
toadstool	Old
popcorn	Old
fanfare	Old
gemstone	Old
watchdog	Conjunction
foresight	Conjunction
tablespoon	Conjunction
chicken livered	Conjunction
peacemaker	Conjunction
snakebite	Conjunction
tailgate	Conjunction
lifeguard	Conjunction
overcoat	Conjunction
anybody	Conjunction
stop sign	Conjunction
nutshell	Conjunction
armchair	Conjunction
storehouse	Conjunction
shareholder	Conjunction
speed trap	Conjunction
steamboat	Feature lure A
haircut	Feature lure A
seashore	Feature lure A
sandbox	Feature lure A
goldsmith	Feature lure A
drawstring	Feature lure A
stockbroker	Feature lure A
jackrabbit	Feature lure A
eyelash	Feature lure B
tightrope	Feature lure B
starfish	Feature lure B
brainstorm	Feature lure B
bathroom	Feature lure B
fireman	Feature lure B
earthworm	Feature lure B
windpipe	Feature lure B
highland	New
nosedive	New
deadline	New
teardrop	New
pinhole	New
wallflower	New
riptide	New
shotgun	New
hedgehog	New
handbag	New
sweatshirt	New
outcast	New
masterpiece	New
flagship	New
backstroke	New
ghostwriter	New

Appendix B: Experiment 2 DRM Word Lists

List Studied

Study												
LIST 1	FAST	LETHARGIC	STOP	LISTLESS	SNAIL	CAUTIOUS	DELAY	TRAFFIC	TURTLE	HESITANT	SPEED	QUICK
LIST 2	TRUCK	BUS	TRAIN	AUTOMOBILE	VEHICLE	DRIVE	JEEP	FORD	RACE	KEYS	GARAGE	HIGHWAY
LIST 3	BED	REST	AWAKE	TIRED	DREAM	WAKE	SNOOZE	BLANKET	DOZE	SLUMBER	SNORE	NAP
LIST 4	SOUR	CANDY	SUGAR	BITTER	GOOD	TASTE	TOOTH	NICE	HONEY	SODA	CHOCOLAT	HEART
LIST 5	HILL	VALLEY	CLIMB	SUMMIT	TOP	MOLEHILL	PEAK	PLAIN	GLACIER	GOAT	BIKE	CLIMBER
LIST 6	THREAD	PIN	EYE	SEWING	SHARP	POINT	PRICK	THIMBLE	HAYSTACK	THORN	HURT	INJECTION
LIST 7	MAD	FEAR	HATE	RAGE	TEMPER	FURY	IRE	WRATH	HAPPY	FIGHT	HATRED	MEAN
LIST 8	NOSE	BREATHE	SNIFF	AROMA	HEAR	SEE	NOSTRIL	WHIFF	SCENT	REEK	STENCH	FRAGRANCE
LIST 9	WATER	STREAM	LAKE	MISSISSIPPI	BOAT	TIDE	SWIM	FLOW	RUN	BARGE	CREEK	BROOK
LIST 10	HOT	SNOW	WARM	WINTER	ICE	WET	FRIGID	CHILLY	HEAT	WEATHER	FREEZE	AIR
LIST 11	SMOOTH	BUMPY	TIRE	BALL	SANDPAPER	JAGGED	READY	COARSE	UNEVEN	RIDERS	RUGGED	SANDPAPER
LIST 12	ELASTIC	BOUNCE	GLOVES	TIRE	BALL	ERASER	SPRINGY	FOAM	GALOSHES	SOLES	LATEX	GLUE

Recognition Test List

FAST	Old
TRUCK	Old
BED	Old
SOUR	Old
HILL	Old
THREAD	Old
MAD	Old
NOSE	Old
WATER	Old
HOT	Old
SMOOTH	Old
ELASTIC	Old
SLOW	Critical lure
CAR	Critical lure
SLEEP	Critical lure
SWEET	Critical lure
MOUNTAIN	Critical lure
NEEDLE	Critical lure
ANGER	Critical lure
SMELL	Critical lure
RIVER	Critical lure
COLD	Critical lure
ROUGH	Critical lure
RUBBER	Critical lure
TOWN	New
TABLE	New
PENCIL	New
BUTTER	New
CIGARETTE	New
DOOR	New
NOTE	New
WEB	New
MUG	New
HARD	New
NURSE	New
LOW	New
CITY	New
CHAIR	New
PEN	New
BREAD	New
SMOKE	New
WINDOW	New
MUSIC	New
SPIDER	New
CUP	New
SOFT	New
DOCTOR	New
HIGH	New