Exploring the Use of Virtual Reality in Episodic Memory Research

Thesis submitted for the degree of

Doctor of Philosophy

The Bishop Grosseteste University

By

Vytautas Nastajus MSc

Department of Social Sciences

Bishop Grosseteste University

October 2020

Abstract

Episodic memory (EM) allows us to receive and retain information about events, where those events happened and when they happened. This knowledge defines humans and if it deteriorates it affects everyday functioning. As such it is important to assess it in a way that reflects our everyday experiences. Evidence suggests that an ecologically valid way of testing EM is needed. One way of achieving this is by using virtual reality. The present thesis aimed to explore how HMD-VR can be used to test EM in an ecologically valid fashion and to attempt to conceptualise and understand the nature of long-term EM as events.

Experiments 1 and 2 explored how EM for events differed to EM for non-events or static objects, the latter being stimuli often used in EM research. Additionally, due to the majority of research focusing just on encoding and retrieval, leaving out memory consolidation, the experiments explored how sleep-dependant memory consolidation affects EM. Events were better retrieved then non-events in several tasks. Additionally, results showed that EM for events might not rely on enhanced encoding but on preferential consolidation, as no difference between event- and non-event-retrieval was observed in EM accuracy immediately after encoding, yet events were significantly better retrieved than non-events after a 24-hour period. Experiments 3 and 4 explored how EM performance in HMD-VR differs to Desktop-VR, a system that is traditionally used in the field of memory research. The general prediction was that EM performance would be more accurate in HMD-VR, compared to Desktop-VR. The results were mixed, with Experiment 1 showing no differences in performance while Experiment 2 only partly supporting the prediction by showing better EM performance in some of the measures. The thesis proposes HMD-VR to be a valid, if not more accurate, tool for exploring daily-life-like EM.

Acknowledgements

I would like to start by thanking my supervisors – Dr. Caroline Horton and Dr. Thomas Dunn. Their guidance, feedback and endless support was invaluable at every step of my thesis. I am grateful for Caroline's constant positive attitude and the meetings during which we tried to get to the bottom of many ideas and theories. I thank Tom for his much-needed insight and comments that led me to think about my problems from completely different viewpoints. Also, I am grateful for his statistical knowledge which helped me when I got too overwhelmed by my data. I would also like to thank Dr. Emma Pearson whose kind words pushed me in the right direction. I am also grateful to Bishop Grosseteste University for funding my studies.

I would like to thank Martin Smith and Anthony Bloxham for all the great time spent together in the lab. Our conversations about the future, research and virtual reality provided a much-needed distraction. I am grateful for all the monthly lunches we had with the other doctorate students during which we shared our joys and sorrows about our studies.

A huge thank you goes to my parents. I would not have been able to do any of this without their support. Thank you for believing in me.

Finally, I want to thank Georgia. Without her help and support, I would not be where I am now. Thank you for being there when I was at my lowest and brightening my days. You made this possible and for that, I will be forever thankful.

Dedication

I would like to dedicate this thesis to my grandparents Pranciška and Karolis Čiegiai who sadly passed away before its completion. You were always so positive and always wished the best for me. I know you would be proud of me.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

1.1. Episodic memory

The ability to learn, store and retrieve experiences from past events is vital to one's survival (Pause et al., 2013). For example, after burning their hand on a fire a child will likely remember to be more careful around it. Remembering experiences is essential for healthy functioning. It allows one to remember where they left their car or to recollect the events of a robbery. Our sense of self is defined by episodic memory (Conway, 2005). It allows us to remember what we do and do not like and to recognise close and distant relationships. When this ability deteriorates, our everyday functioning gets affected and becomes more challenging. As such it is important to understand the inner workings of episodic memory and this begins with its definition.

Since the start of its study, the definition of episodic memory (EM) has grown in complexity. The first definition of EM was published by Tulving (1972) and since then, it acted as a base for all further definitions. Tulving's initial definition was that "episodic memory receives and stores information about temporally dated episodes or events, and temporal-spatial relations among these events" (Tulving, 1972, p. 385). Episodic memories were referred to as "personally experienced unique events". These experiences and events are temporally dated and contain both temporal (time) and spatial (location) relationships to other events and experiences from one's own past. While this basic notion of EM being memory for unique events has not changed much over the years (Moscovitch et al., 2016; Rugg et al., 2015; Schacter & Madore, 2016) the concept has been expanded.

It is important to mention that Tulving's 1972 paper was not just about EM. It was about subdividing declarative memory into two different but closely associated sub-systems: episodic and semantic memory. According to Tulving, semantic memory was "a mental thesaurus, organized knowledge a person possesses about words and other verbal symbols" (Tulving, 1972, p. 386). Semantic memory was characterised as having the capacity for "inferential reasoning and generalisation" (Tulving, 1972, p. 390) a property that was not present in EM. Moreover, Tulving suggested that semantic memory, contrary to EM, did not hold any information on how its content was acquired which led to learning being an unknown parameter (Tulving, 1972).

The original distinction between the two memory systems came from Tulving's analysis of word list memory experiments (Tulving, 1972) in which participants were provided with lists of words and later asked to recall those words. According to Tulving (1972), these experiments were quintessential EM tasks. When a participant is asked, after a retention period, whether a word appeared in the learning phase, it is not learning that is being tested but a memory for a specific event that took place at a particular time and place (Tulving, 1972). However, Tulving recognised that EM might be encoding the outcome of semantic processing if the experimental material is made up of familiar words. This led to the conclusion that semantic memory influences episodic retrieval and that memory tests are not 'process pure' meaning that both memory systems play a part in them (Jacoby, 1991; McCabe, Roediger, et al., 2011). It is important to consider as to what other qualities EM possesses that distinguishes it from semantic memory. Recent research has shown that episodic and semantic memory systems are interdependent and share many attributes (Craik, 2000; Greenberg & Verfaellie, 2010; Saive et al., 2015). This interdependence will be further discussed in a later section regarding memory consolidation.

Nevertheless, Tulving's analysis of the phenomenological components of both memory systems led him to the conclusion that EM contains an 'autonoetic' quality': a sense of subjective and lived experience that, for example, recalling semantic information (e.g. a capital city) will not have (Tulving, 1983). This autonoetic quality, or consciousness, is associated with images entering conscious awareness, attention turning inward and a strong sense of the self in the past (Conway, 2009). The autonoetic consciousness was later expanded and became one of the main attributes of EM (Klein, 2013; Tulving, 2002, 2004).

1.2. The main attributes of Episodic Memory

According to Tulving, a major difference between semantic and episodic types of memory is consciousness or "recollective experience" (Tulving, 1985, 2002, 2004). Tulving discussed that remembering is an act which requires one to be consciously aware of something that happened in the past. Tulving postulated that semantic and EM systems are characterised by different kinds of consciousness. To better explain them, Tulving adopted a term originally proposed by Husserl (1964) - “noesis” (a type of experience associated with thought and remembering).

Semantic memory was characterised by noetic consciousness. This type of consciousness allows an organism to be aware of information about the world without recollecting it. One is noetically aware when they are retrieving general information without the feeling of reliving and re-experiencing the past (Szpunar & Tulving, 2011). In other words, semantic knowledge lacks the sense that one is travelling back in time to access it. This lack of mental time travel is thought to distinguish the two memory systems (Suddendorf & Corballis, 2007; Tulving, 1985, 2001, 2002; Wheeler et al., 1997)

Episodic memory is associated with autonoetic consciousness which is defined as a sense of self in time and the mental reliving of subjective experiences. Tulving described it as “... a unique awareness of re-experiencing here and now something that happened before, at another time and in another place” (Tulving, 1993, p. 68). When one remembers an event, they are re-experiencing it and are also aware that an event is part of their own past and existence. As a result, it was proposed that to remember something episodically is to mentally time travel (Cassel et al., 2012; Markowitsch & Staniloiu, 2011; Suddendorf & Corballis, 2007; Tulving, 1985). However, such definition of the main attribute of EM is problematic. The main argument is that it is difficult to objectively verify if someone is using autonoetic consciousness (Dunn, 2004; McCabe, Geraci, et al., 2011; Wixted, 2009). For example, using the earlier mentioned word list experiments, the only way to know that a participant used EM to recall a specific word is to just believe them when they say that they remembered it.

Evidence for autonoetic consciousness has been drawn from clinical observations of amnesia (for a review of a number of cases see Hornberger & Piguet, 2012; Tulving, 1985; also see Cermak & O’Connor, 1983). For example, patient K.C could not recall any events or incidents from their past but were able to recall information such as the year their family moved into a new house or the name of the school he once attended (Tulving, 1985; also see Rosenbaum et al., 2005). Overall, K.C.’s general knowledge and language skills were relatively intact. K.C. was able to define words like “consciousness” and “tangible”, and give scripts for specific activities, like going to a restaurant or calling someone. However, when they were asked what they did before coming to the interview or what they did a day before, they could not answer and described their mind as being "blank". Tulving (1985) proposed that while the patient had noetic consciousness, being able to retrieve general information without the feeling of reliving, they lacked autonoetic consciousness, being able to mentally relive subjective experiences. He argued that this corresponded to having an intact semantic memory but damaged EM.

Tulving later modified his description of EM to include a subjective sense of time called chronesthesia (Rosenbaum et al., 2007; Szpunar, 2011; Tulving, 2004). Chronesthesia is a specific kind of consciousness which facilitates the awareness of the subjective time in which one exists. It enables one to mentally time travel to a specific point in time and using the autonoetic consciousness, re-experience a specific event from one's own subjective history. Chronesthesia and autonoetic consciousness are closely related. Although both indicate the awareness of self in time, chronesthesia focuses on the awareness of subjective time, whereas autonoetic consciousness relates to an awareness of the self (Tulving, 2002). Chronesthesia could be considered the temporal aspect of autonoetic consciousness. While the distinction is subtle it is necessary as, according to this conceptualisation, it is possible to operate with time independently of self and self can be dealt independently of time (Szpunar, 2011).

As it can be seen, EM has thus far three critical attributes: self (the rememberer), autonoetic consciousness (conscious awareness that one is remembering one's own past) and chronesthesia (mental time travel) (Tulving, 2002). These attributes, especially the latter two, are explored in a wide range of literature when defining EM (Allen & Fortin, 2013; Cassel et al., 2012; Markowitsch & Staniloiu, 2011; McCabe, Geraci, et al., 2011; Pause et al., 2013; Spreng et al., 2009). Tulving (1985) argued that while the amnesiac patient K.C. did not possess EM or autonoetic consciousness they were still able to make statements about their past. This meant that even if someone is not able to remember a specific event (autonoetic consciousness) using EM, they can still know (noetic consciousness) something about it, which relies on semantic memory (Levine, 1998; Markowitsch & Staniloiu, 2011; Rajaram, 1993). Such definitions can be useful when understanding the conceptual nature of EM, however, they pose a problem in measuring and/or capturing it if one is to ascertain if one is retrieving information episodically or semantically. One potential way to test that is to employ the Remember/Know paradigm (Tulving, 1985; Wais et al., 2008; Wixted, 2009).

1.3. Remembering and Knowing

As it has been discussed earlier, the distinction between episodic and semantic systems can be broadly characterised as the difference between 'remembering' and 'knowing'. Remembering refers to the mental recollection of personally experienced events and is supported by autonoetic consciousness, whereas knowing refers to the retrieval of decontextualized earlier learnt information (Suddendorf & Corballis, 2007; Tulving, 1972, 1985, 2001, 2002; Wheeler et al., 1997).

While the concept of remembering and knowing was introduced in the initial description of EM (Tulving, 1972), only later (Tulving, 1985) was it developed to probe the different activations of episodic and semantic memory and has become widely used today (Dewhurst et al., 2009; Dunn, 2008; Rotello & Macmillan, 2006; Wais et al., 2008; Wixted, 2009). In general, the Remember/Know paradigm requires participants to indicate if they remember (using EM) and can mentally re-experience an episode in which a specific stimulus was presented or simply know (using semantic memory) that the specific stimulus was presented in the past without re-experiencing that episode.

One problem with the Remember/Know paradigm is that it is frequently used in distinct ways by different researchers. For example, the criterion of re-experiencing is not utilised by some researchers (Yonelinas, 1997, 2002), whereby a remember response indicated recollection of qualitative details about the initial episode, and a Know response indicated the absence of recollection and only feeling of familiarity that the episode was experienced in the past (see Mandler, 2008 for a review). As an example, in a study by Yonelinas and Levy (2002), remember responses were measured as an ability to recall the colour of a studied word which is not the same as Tulving's criterion of mental re-experiencing of an episode.

While the two types of the Remember/Know paradigm overlap, the utilisation is different. For example, Yonelinas (1997) did not use the term 'episodic' and instead proposed that 'remember' responses were controlled by 'recollection' processes which were based on a threshold retrieval process - contextual information either being successfully retrieved or not and are measured by the presence of detail information regarding a studied item. On the other hand, Tulving suggests that 'remember' responses are episodic, achieved using autonoetic consciousness and are measured by the presence of re-experiencing the event. For example, in a world list experiment, remembering that a word was a certain colour would be classified as a 'remember' response by Yonelinas (1997). In contrast, using Tulving's criteria one would require the participant to report that they can re-experience it and see the word in that colour, in their 'mind's eye'. From these definitions, it is possible to see that Tulving's approach is harder to objectively test. On the other hand, while it can be argued that the recall of a perceptual detail is more episodic than the recall of just the object (Conway, 2001, 2009; Tulving, 2001), it has been shown that recall of perceptual details can be achieved using semantic knowledge (Gardiner, 1988; Gardiner et al., 2002; Knowlton & Squire, 1995; Rajaram, 1993). Thus, while asking if one can re-experience the episode might be a better approach to investigate EM, compared to the recall of detail information, it is important to remember its subjectiveness.

While it can be seen that the Remember/Know paradigm is unlikely to be process pure in that it could exclusively test the use of one specific system (Dewhurst et al., 2006; Jacoby, 1991; Jennings & Jacoby, 1993; Yonelinas, 2002), its use has revealed a number of dissociations between semantic and EM systems (Barba, 1993; Barba et al., 1997). Although the Remember/Know paradigm can be used to measure EM, there are other and arguably better ways of measuring it, which will be considered in the next section.

1.4. Measuring Episodic Memory

Adult humans are capable of reporting the content and the accompanying subjective experience of remembering. Due to this, one of the ways to investigate EM is using interviews or self-reports (Buckner & Carroll, 2007; Crovitz & Schiffman, 1974; Hashtroudi et al., 1990; Hassabis & Maguire, 2007; Kopelman et al., 1989; Levine et al., 2002; Schacter & Addis, 2007). Also, there is an agreement that measuring accuracy and vividness of events tap episodic cognition (Addis & Schacter, 2008; Barr et al., 1990; Buckner & Carroll, 2007; Hassabis & Maguire, 2007; Hokkanen et al., 1995; Kapur et al., 1997; Maguire & Mummery, 1999; Tanaka et al., 1999). However, when assessing EM in clinical or neuropsychological case studies, the heavy reliance on verbal competence is not always appropriate. Due to this, there is also a need for an EM test that is not as dependant on verbal capabilities.

As it has been discussed earlier, establishing a clear definition of EM is challenging, with different researchers highlighting different defining features. This creates a validity problem by introducing uncertainty if the test that is being used actually measures EM, rather than another related process such as semantic memory or familiarity. Nevertheless, most EM tests can be categorised into a few well-defined measurement tools, with each assessing an important aspect of EM: the What-When-Where, free recall, autobiographical and recognition tests. These tests and some rarer ones, such as the source recognition test, will be explored with the emphasis on how they link with EM definitions described so far.

1.4.1. The What-When-Where test

Clayton and Dickinson (1998) have argued that Tulving's original definition of EM as a system that "receives and stores information about temporally dated episodes or events, and temporal-spatial relations among these events" (Tulving, 1972) could be used as a behavioural criterion for EM. They have argued that to test EM the separate What, Where and When (WWW) components of EM need to form a single representation. This sum of the separate components is often referred to as an episode in EM (Martin-Ordas et al., 2017; Pause et al., 2013; Plancher et al., 2008).

The foundations of the WWW test were laid in a study by Clayton and Dickinson (1998). They examined EM by testing scrub jays in two food hiding trials separated by short (4 hour) and long (124 hour) intervals (When). During the trials, scrub jays were able to hide either worms or peanuts (What) in a storage tray filled with sand (Where). The worms used in the study expired after the long interval while the peanuts did not. The authors have found that the scrub jays recovered worms after the short (4 hour) intervals and peanuts after the long (124 hour) intervals. This was interpreted as an indication that the scrub jays had an integrated WWW-like knowledge.

A similar use of the WWW test has been used with a variety of species (Babb & Crystal, 2005; Bird, Roberts, Abroms, Kit, & Crupi, 2003; Hampton, Hampstead, &

Murray, 2005) establishing that some animals do possess episodic or episodic-like memory behaviour. However, it has not been possible to evidence that they used autonoetic consciousness for the recollection of their experiences. Tulving argued that only humans have this conscious aptitude (Tulving 2001), although it is difficult to make such an inference about animal's conscious abilities on the basis of their overt behaviour. Indeed, a number of studies (Hayne & Imuta, 2011; Holland & Smulders, 2011; Martin-Ordas & Atance, 2019) concluded that humans are able to use EM to perform the WWW tasks used in animal research and that those tasks do contain the required mental time travel and autonoetic consciousness essential for Tulving's definition of EM (Tulving, 2005).

Pause and colleagues (Pause et al., 2010) assessed the WWW components of EM in an integrated manner by using verbal ratings and non-verbal motor responses. Visual stimuli were hidden behind four out of eight quadrants presented on a computer screen. Each of the quadrants could be highlighted by using a corresponding key on a keyboard which would reveal a stimulus or remain black. The task for participants was to remember on which occasion (when), at which position (where) and which specific picture (what) had been encountered. The results showed that the mean number of button presses for each of the four quadrants with the visual stimuli positively correlated with verbal EM performance. The Authors concluded that EMs can be experimentally induced and that non-verbal behaviour of button pressing can be correlated with EM performance.

While the method used in the Pause et al. (2010) did assess the three WWW components, some researchers argued that it assessed 'episodic-like' and not fully EM (Martin-Ordas et al., 2017; Martin-Ordas & Atance, 2019). The reason for this is that it did not test chronesthesia. As discussed previously (see section 1.2), chronesthesia is thought to be a critical feature of EM. As such, testing the 'when' component with a "how long ago" question does not necessarily test chronesthesia as opposed to the question "in which order" did something happen? This is due to the order in which events happen being a better representation of EM than the time elapsed since the particular episode (Roberts et al., 2008). It is possible that the question "how long ago" might rely more on the strength of the memory than the recollection of an episode (Easton & Eacott, 2008). However, Clayton, Griffiths, Emery, & Dickinson (2001) argued against this hypothesis by showing that what-where memories were comparably good after one week as after a few hours (Clayton & Dickinson, 1999). It is important to mention that these hypotheses and arguments were formulated using data from animal models and one needs to be cautious applying them to human EM.

An example of such testing can be seen in a WWW study by Holland and Smulders (2011), which reassembled the food hiding experiments by Clayton and Dickinson (2001; 1998). Participants had to hide two types of coins (what), in different locations in a living room (where) on two subsequent days (when). On day 3 (testing), all participants had to recall the WWW information. In addition to the WWW questions, participants were also given unexpected questions about the context of the hiding episodes such as “Were there letters and fliers on the floor by the front door?” Results have shown that the performance on the unexpected questions significantly predicted the WWW task performance and that participants reported using mental time travel for recall. Authors concluded that participants used EM to solve the WWW task used in their study.

The WWW task is now widely used to explore EM (Cheke, 2016; Chrastil & Warren, 2013; Plancher et al., 2012, 2013; Smulders et al., 2017a; Wang et al., 2018) due to the evidence demonstrating that the task provides the mental time travel and autonoetic consciousness that are inherent in Tulving’s definition of EM (Tulving, 2002). However, it is important to look at other ways of assessing EM also, and how these relate to the critical components of EM.

1.4.2. Clinical tests

Clinical tests are used in clinical situations, underpinning the field of neuropsychology. Using these tests it is possible to infer cognitive structures from modular impairments to particular behaviours (and associated brain regions) via dissociations and double dissociations (Machery, 2012; Rich & Troyer, 2008). Some examples of EM tests in a clinical setting include the ´Montreal cognitive assessment´ (Nasreddine et al., 2005), the ´Blessed dementia information– memory–concentration test (Blessed et al., 1968), the ´Mini-Cog´ (Borson et al., 2000), Mini-Mental Status Examination Test (Carcaillon et al., 2009; Folstein et al., 1975), the 7-min screen (Solomon et al., 1998), the ´Three words-three shapes´ memory test (Weintraub et al., 2000), the ´Word list memory test´ of the Consortium to Establish a Registry for

Alzheimer's Disease (Welsh et al., 1992) and the Wechsler Memory Scale (Humphreys et al., 2010). However, clinical tests usually use very narrow definitions of EM that does not encompass all of the critical components.

For example, the Mini-Mental Status Examination Test (Carcaillon et al., 2009; Folstein et al., 1975) is comprised of items exploring semantic memory and executive functions and is focused on detection and measurement of cognitive impairment. The memory is tested by asking questions such as – “What is the street address of the building we are in?” or asking to remember and recall 3 words. On the other hand, the subtests of the Wechsler Memory Scale (Humphreys et al., 2010) only test verbal and visual memory by asking to recall a number of words or drawing a picture from memory, as in The Rey-Osterrieth Complex Figure Test (Shin et al., 2006). Even the more current tests that claim to target EM such as the California Verbal Learning Test (Delis et al., 2000) rely on learning and reproducing a list of items. Using the research discussed earlier in the thesis, it is possible to argue that such task falls more under semantic than EM as it does not involve any of the main aspects of EM.

Indeed, the problem with such tasks is that they do not explicitly measure the spatial (where) and temporal (when) components of EM as one integrated unit. What is more, it is not possible to assume that the individual is able to remember where and when the recognised item was presented. As has been discussed previously, the temporal component is closely linked to mental time travel, an integral part of EM.

1.4.3. The Autobiographical interview

Typically, memories for personal episodes, that occur incidentally or naturalistically, have been investigated as autobiographical memories rather than EMs. Models of autobiographical memory acknowledge both the centrality of the self at the points of encoding and retrieval, as well as the possible interplay between EMs and semantic knowledge (e.g. Conway & Pleydell-Pearce, 2000). Indeed Conway (2003) placed EMs within the autobiographical memory system, as being highly specific examples of personal episodes, but that other, more semanticised and general information such as schematic memories for a walk to work, or about living in a particular location within one’s life, also formed autobiographical memories. It is important to point-out the consensus in research that the personal reference and self-involvement are the key elements of autobiographical memory (Conway, 2005; Marsh & Roediger, 2012). As such one could easily mistake EMs for being highly detailed and accurate accounts of experience, with other autobiographical information being vaguer or even changed over time, as the episodic details are lost. The assumption that of EM is accurate is a misleading one, as research has demonstrated the ease with which false EMs can be facilitated (Gallo, 2010; Zhu et al., 2013).

Nevertheless, EM has been investigated using interviews for autobiographical or important life events (Bartsch et al., 2011; Kopelman et al., 1989; Levine et al., 2002). For example, a semi-structured interview was developed by Kopelman et al. (1989) to assess autobiographical events and “personal semantic” memory within clinical populations. This interview assessed memory across childhood, early adulthood and recent time periods by asking to recall autobiographical incidents and “personal semantic” memory from those periods in life. The term “personal semantic” memory was used for factual information about a person’s past such as names of friends or addresses where they lived or worked. Participants or patients were asked to describe full episodes that happened at a particular place and time. The verbal or written reports were then scored depending on the frequency of episodic and non-episodic details. The episodic details in the autobiographical interview refer to the usual what-where-when, perceptual, thought and emotional information Kopelman et al. (1989). In a clinical setting, a low number of episodic recalls meant an impairment in EM functioning.

A problem with this approach is that it assesses both episodic and “personal semantic” memory. Due to the nature of autobiographical memories, they are often retrieved and recounted a number of times which in turn increases the possibility of changing the memory’s content due to reconsolidation and interference (Schwabe et al., 2014; Winters et al., 2009). As a consequence, the memory that is being measured is more likely to be semantic than episodic. However, this problem can potentially be addressed by designing tests that distinguish between semantic and episodic components of autobiographical memory by way of investigating the spatio-temporal information of those events (Levine et al., 2002).

Another problem with the autobiographical interview is that EM performance is assessed by relying on memories that can be retrieved. For example, autobiographical memories for life events in patients with early stages of mild cognitive dementia or

Alzheimer's disease can still be accessible if memories for those events have been established years before the onset of memory complaints (Squire & Alvarez, 1995). As discussed earlier, memories with time become less specific and lose details becoming less episodic and more 'gist-like' (Donix et al., 2010; Martinelli et al., 2013; Murphy et al., 2008). Moreover, if we assume that EM impairments affect encoding and consolidation of new episodic information, more recent EMs could potentially provide are a better measure of EM function.

Lastly, it is difficult to verify the accuracy of the recalled autobiographical information. Research has shown that people susceptible to creating false memories (Devitt et al., 2016; Hyman & Loftus, 1998; Schmidt, 2004). Two of the main memory error types are gist and intrusions (Barclay, 1986, 1988; Bywaters et al., 2004; Hauer et al., 2006). The gist errors are related to semantisation of EMs and loss of details which results in recall of only the main gist of the memory. Intrusion errors happen when recalling specific details of events but those specifics being constructed based on general knowledge. In other words, instead of recalling actual specific details about an event, people might instead add that information using their general knowledge. Both of these errors show that to actually examine EM it is better to use controlled stimuli which would allow to objectively assess memory performance.

1.4.4. The Free recall test

Tulving argued that autonoetic consciousness can be assessed by asking participants if they remember an item from a previously provided list or whether they simply know that that item was on the list (Tulving, 1985). He proposed that remembering refers to the mental recollection of events and is based on autonoetic consciousness, whereas knowing refers to the retrieval of earlier learnt information without re-experiencing that event (Tulving, 1972). The remember/know paradigm helps to distinguish between memories that are more episodic than semantic. Tulving observed that participants were more likely to report remembering items from a word list during a free recall task compared to a cued recall task. The free recall (FR) task, as the name implies, asks a participant/rememberer to freely recall (without any cues) any information they can about a specific event. Tulving concluded that during a FR task participants use internal cues which lead to items being remembered more episodically compared to using external cues such as the first letter of the to-be-remembered word. This led to a conclusion that FR tasks explore EM to a greater extent compared to the cued recall tasks. It was hypothesised that when the support for retrieval is low, such as in the FR task (no external cues), the strength of episodic information needs to be high to lead to retrieval. When retrieval does happen it is accompanied by autonoetic consciousness (Tulving, 1985). Due to this, FR tests are very common in EM literature (Bäckman et al., 2001; Herlitz et al., 1997; Howard & Kahana, 2002; Tulving, 1985; Tulving et al., 1995) and are also included in a number of clinical assessments (Cognitive Drug Research Battery (Simpson et al., 1991), Consortium to Establish a Registry for Alzheimer's Disease Neuropsychological Test Battery (Fillenbaum et al., 2008).

In the majority of EM tests, there are main elements such as words in a list that the subjects are informed they will be tested on (e.g. van der Helm et al., 2011). This kind of setting differs from how EMs are encoded and retrieved in our everyday lives. In our everyday lives, information is often encoded without knowing that that knowledge will be later tested. What is more, some of the information that might need retrieving might not have been the focal point of the episode from which it is being retrieved. Some researchers believe that this is a defining feature of EM as in it catches everything about an event even if something is not in the central focus of attention (Morris & Frey, 1997).

1.4.5. Source memory test

It has been shown that memory for focal elements and source memory for contexts differ and are independent (Johnson & Raye, 1981; Shimamura & Squire, 1987), suggesting that these two types of memories might be representing semantic and episodic knowledge, respectively. For example, studies have shown that patients with source amnesia (usually related to frontal lobe damage) can demonstrate relatively intact memory for facts, but are impaired for memory of how and when those facts were learned (Schacter et al., 1984; Shimamura & Squire, 1987). It is possible to argue that the difference between focal and contextual information stems from the deliberateness of encoding. Focal factual information is more important, being the main point of information, and thus more likely to be encoded than source information. Tasks that investigate source memory are widely used in the EM literature (Davachi et al., 2003;

Drummey & Newcombe, 2002; Johnson & Raye, 1981; Lundstrom et al., 2005; Shimamura & Squire, 1987; Simons, 2002; Whitcombe & Robinson, 2000). While the contextual information is important for EM, tests such as the WWW represent EM. This is mainly due to the fact the WWW test already incorporates all the needed information regarding the episode.

1.4.6. Recognition tests

Another test that explores EM is based on Tulving's (1985) remember/know paradigm is the recognition test. While a lot of information regarding recognition and the remember/know paradigm was discussed earlier (section 1.3), here will focus more towards the test itself. In recognition tasks subjects are presented stimuli (usually words or pictures of objects) and are asked if they think the stimulus was previously presented (responding – yes) or if the stimulus was not previously presented (responding – no). If the response is yes – the subject is then asked to indicate if they can recollect seeing that stimuli (a slow and more deliberate process based on episodic recall) or is the stimulus just feels familiar (fast and automatic process-based more on semantic recall). This recollection/familiarity judgement is based on the dual-process theory of recognition memory – memory that indicates whether an event has been previously experienced (Aggleton & Brown, 2006). The dual-process theory assumes that item recognition is based on recollection but only if its recollected strength exceeds a threshold. If this does not happen item recognition is based on familiarity (for the review of these models see Wixted, 2007; Yonelinas, 2002). This theory directly maps onto the remember/know judgements (Rugg & Yonelinas, 2003; Wixted, 2007; Yonelinas, 2002).

Usually, recognition tests fall into one of the two categories: task-dissociation methods such as response-speed methods and item/associative recognition comparisons (Ratcliff & Murdock, 1976) or process-estimation methods such as process-dissociation (e.g. Jacoby, 1991), receiver operating characteristic (ROC) procedures (e.g. Yonelinas, 1997) and remember/know judgement tasks (e.g. Dewhurst et al., 2009). The former methods try to identify a task or a condition that isolates one of the processes while the later try to develop a set of model equations that attempt to obtain parameter estimates representing the contribution of recollection and familiarity to overall performance

(Yonelinas, 2002). However, while a brief description will be given for each of these kinds of tasks the main focus will be the Remember/Know test. This is due to the fact that it better relates to Tulving's definition of EM and the crucial element of autonoetic consciousness.

1.4.6.1. Task-Dissociation Methods

The goal of the task-dissociation method is to find a task or an experimental condition which will help to dissociate recollection from familiarity (associated respectively with episodic and semantic memory). Such a task should be able to provide results that are different from standard recognition tests in which recollection and familiarity are both involved. This kind of disassociation should allow to better understand the different contributions of familiarity and recollection to the specific task or condition.

Item/associative recognition methods are based on the assumption that recollection reflects retrieval of qualitative information about an event (Jacoby, 1991; Mandler, 1980; Tulving, 1985; Yonelinas, 1994, 2002). Recollection is said to reflect information about paired items such as knowing if the items were paired together. On the other hand, familiarity should reflect information about single items such as knowing if the item is old or new (Yonelinas, 2002). In studies that use this method (e.g. Kilb & Naveh-Benjamin, 2007; Naveh-Benjamin, 2000; Old & Naveh-Benjamin, 2008), participants study items in pairs (e.g. apple-car, table-water) and then are tested using both item and associative tests. In the associative tests, participants are given studied items in either the original (e.g. apple-car) or altered (apple-dog) pairs and are asked to provide 'old' responses to the original pairs. The rationale is that familiarity should not be helpful in this type of test as all of the items (no matter how they are paired) have been seen before, and as a result, participants have to use recollection.

1.4.6.2. Process-Estimation Methods

As it has been discussed, the task-dissociation methods focus on what kind of inferences can be made on familiarity and recollection contributions on different tasks. Contrary to that, process-estimate methods focus on providing quantitative estimates of the contribution of recollection and familiarity to the overall recognition performance on a single task. Process-estimate methods can be grouped into three main types: the process-dissociation procedure, the receiver operating characteristic procedure and the Remember/Know paradigm.

The process-dissociation procedure (Jacoby, 1991) is based on the earlier mentioned item/associative recognition method. This procedure is based on the idea that only recollection would let the participant remember *where* or *when* they saw the studied item. An example of this is in a study by Yonelinas (1994) in which participants had to learn words from two lists. Later participants had to respond positively to words from the first list (inclusion condition) but not from the second one or new words. In the second condition, participants had to respond negatively to word from the first list (exclusion condition) or new words and positively to words from the second list. The basis behind this model is that by comparing inclusion and exclusion performance it is possible to estimate recollection and familiarity parameters.

The receiver operating characteristic (ROC) procedure in EM examines the effect of varying response criterion on correct and ‘false alarm’ responses to estimate how recollection and familiarity contribute to recognition memory (Yonelinas, 1994). ROCs are usually created by taking participants’ confidence ratings in their yes/no recognition responses and plotting them against false alarms creating a response-confidence function. The main drawback is the number of different theories that try to explain the results and predict the outcome of the test (for a review of related theories see Wixted, 2007; Yonelinas, 2002), showing a lack of clarity on how the ROC procedure relates to the structure of EM. However, a more useful version of ROC is the d’ sensitivity index. The d prime (d’) index is based on the signal detection theory and measures sensitivity or discriminability. It is a standardized score which is computed as the difference between the standard scores for the false-alarm rate and the hit rate (Macmillan & Creelman, 2005). Overall, it shows how well participants can discern between stimuli that they have seen (old) and stimuli that they have not seen before (new) with higher d’ scores indicating better discrimination between different stimuli (Haatveit et al., 2010).

1.4.6.3. The remember/know procedure

Developed by Tulving (1985), the Remember/Know (R/K) procedure is now one of the most widely used methods to investigate recognition memory. As explained by Tulving (2002), the R/K procedure maps accordingly to autonoetic and noetic awareness and thus to episodic and semantic memory. When the R/K procedure is used to investigate recollection and familiarity, remember responses are taken as recollections and know responses as familiarity judgements (Diana et al., 2008; Mayes et al., 2002).

During the R/K procedure participants are asked to study a list of items. These items are then presented again but this time they are interspersed with lures. For each of the item, participants judge if the item was presented in the original list. If a participant responds 'yes' they then need to respond if they either 'Remember' or 'Know' whether that item was in the original list. A 'Remember' response indicates that the participant can consciously recollect studying that item while a 'Know' response indicates that they cannot but they still recognise the item using some other criteria. The probability of a 'Remember' response is then used as an index of recollection. On the other hand, the probability of familiarity is equal to the conditional probability that the item received a 'Know' response given it was not recollected (Yonelinas, 2002). However, this creates a problem as participants are instructed to respond "know" when an item is "familiar but not recollected" and not just "familiar". This leads to an underestimation of the probability that an item is familiar.

While the R/K description might lend support to dual-process theories (Rugg & Yonelinas, 2003; Wixted, 2007; Yonelinas, 2002) there is strong opposition from single-process signal theories (Ratcliff et al., 1995; Slotnick & Dodson, 2005). Single process theories instead maintain that 'remember' and 'know' judgements reflect different degrees of memory strength or confidence and not different memory processes (Donaldson, 1996; Dougal & Rotello, 2007; Dunn, 2004, 2008; Hirshman & Henzler, 1998; Shimamura & Wickens, 2009; Slotnick & Dodson, 2005; Xu & Bellezza, 2001). As such, 'Know' responses, compared to 'Remember' responses, might actually reflect weaker memories with lesser degrees of recollection (Wais et al., 2008).

The main downside of recognition tests is that they do not explore the WWW components of memory, or more specifically the Where and When components of EM.

While asking participants if they can recollect/remember a stimulus or are just familiar with it/know about it, link to the needed autonoetic consciousness, it does not assess mental time travel. As a result, it is possible to argue that it is not a valid measure of EM.

1.4.7. Interim summary

As can be seen from this review, each EM tests has its advantages and problems. A reoccurring problem is the lack of links to the definition and/or the main components of EM, such as the WWW triad or the autonoetic consciousness. Research on the tests themselves shows that not all of them relate to one another (Cheke & Clayton, 2013, 2015). This suggests a contribution from multiple psychological processes and that not all of these tests necessarily measure the same thing. As such it is important to discuss what tests and why were chosen to be used in the empirical chapter of the present thesis and why.

Following the discussed research, EM in this thesis has been described as a collection of factual, spatial and temporal information about an episode. This translates to the What-Where-When information and as such the corresponding memory test was chosen to be included. In addition to this theoretical compatibility, the WWW test is extensively used in EM assessment (Cheke, 2016; Cheke & Clayton, 2015; Easton & Eacott, 2008; Guillery-Girard et al., 2013; Hampton & Schwartz, 2004; Martin-Ordas et al., 2017) thus being both theoretically and practically supported choice.

Following the mentioned theoretical aspect of EM, the free recall test was also chosen to be included in the experiments performed in the present thesis. The reason for this inclusion is identical to the inclusion of the WWW test – theoretical and practical strengths. As Tulving explained, the use of internal cues during the free recall task lead to items being remembered more episodically (Tulving, 1972). What is more, internal queuing is more reflective of an everyday remembering (Morris & Frey, 1997) which is an important aspect of this thesis and will be discussed in more detail later. While the test is not as common in EM research as the WWW test it is still widely used (Bäckman et al., 2001; Herlitz et al., 1997; Howard & Kahana, 2002; Tulving, 1985; Tulving et al., 1995).

Lastly, the experiments in the thesis included two recognition tasks – the Remember/Know judgements and the d’ sensitivity index. While at first glance this test deviates from the previous two, it does provide an important insight into EM. The reasoning behind employing the R/K judgement task is based on all of the discussed literature showing its links to the EM (Suddendorf & Corballis, 2007; Tulving, 1972, 1985, 2001, 2002; Wheeler et al., 1997) and its use in exploring EM (Dewhurst et al., 2009; Dunn, 2008; Rotello & Macmillan, 2006; Wais et al., 2008; Wixted, 2009). The inclusion of d’ index was based on similar reasoning.

It is important to point out that the number of tests and their downsides is not the only problem concerning EM research. A greater problem lies in defining and therefore operationalising EM. The differences across tests stem from the varying views regarding EM.

1.5. Problems with Episodic Memory

There are problems with the concept and definition of EM that need to be addressed. For example, Tulving (Tulving, 2005,) stated that EM requires but also ‘goes beyond’ the semantic memory system (p. 9). However, a number of studies have shown that patients with semantic dementia are able to remember episodically (Graham et al., 1997, Graham et al., 2000, Simons et al., 2001, Graham and Hodges, 1997, Hodges et al., 1992, Snowden et al., 1996, Simons et al., 1999). As a response, Tulving might argue that because subjects with semantic dementia successfully completed the tests, the tests cannot, by definition, have been testing EM. A similar problem can be seen with EM in animals. Tulving has pointed out that “before we can undertake the task of evaluating the presence of EM in nonhuman animals, the concept needs to be sharpened… We need to be clear about the kind of memory that I am denying to our feathered and furry friends” (Tulving, 2005, p.9). The problem comes from his definition of EM as “recently evolved…and probably unique to humans” (p.9). Due to this definition, it is not possible for animals to have episodic cognition.

Conway and Pleydell-Pearce (2000) argued that Tulving’s description of EM is not conceptually clear. For example, if EMs contain semantic information where is the distinction between episodic and (personal) semantic memories? This point was explored in a study by Conway, Gardiner, Perfect, Anderson & Cohen (1997). In that study students' knowledge gained during lectures was tested at two intervals, six months apart. Students had to indicate if they remembered or knew if a given answer was correct using the information learned during the lectures. Over the two testing sessions, participants that were providing the correct answers showed a clear remember-to-know shift. This was used as evidence that EMs with time lose detail and become more abstract which leads to the 'know' judgements (Cermak, 1972; Dewhurst et al., 2009; Herbert & Burt, 2001, 2004). This creates a problem, as to where is the line after which EM is not episodic anymore. To overcome this problem one needs to rely on the rememberer's subjective judgement if they think they remember or just know that the event happened. However, this subjective judgement does not always reflect memory accuracy. For example, memories for highly emotional events can be reported as more vividly remembered than everyday events but the accuracy for details might not differ and both decline over time (Rimmele et al., 2011; Sharot & Yonelinas, 2008; Talarico & Rubin, 2003).

Conway (2001) re-characterised EM as 'a memory system that retains highly detailed sensory-perceptual knowledge over retention intervals measured in minutes and hours' (p. 1375). While the sensory-perceptual knowledge is typically short-lived, in cases of significant events it can become stabilised and retrieved over longer retention periods. Several papers by Conway (1992; 1995; 1996) have shown the hybrid nature of long-term autobiographical memories and that they can contain details at various levels of specificity and semanticity.

As discussed, EMs are argued to not endure unless they become part of the autobiographical memory system (Conway, 2001). According to this view, access to EMs rapidly degrades and most memories are lost within 24h of encoding. Only EMs which become integrated into the AM system remain accessible. Why this happens and why it is important in EM will be discussed just after a brief overview of the neuroanatomy of EM.

1.6. The Neuroanatomy of Episodic Memory

As shown in section 1.2, evidence for autonoetic consciousness has been drawn from clinical observations of amnesia (Cermak & O'Connor, 1983; Hornberger & Piguet, 2012; Tulving, 1985). This shows that to better understand EM and processes associated with it is important to understand the underlying brain structures (Tulving, 1985).

1.6.1. The hippocampus

It is most appropriate to start discussing the brain structures related to EM by discussing the hippocampus (HPC). Working with Tulving's ideas, Moscovitch & Winocur (1992; also see Moscovitch, 1992), proposed that HPC binds together the neural elements in the medial temporal lobe (MTL) and neocortex which give rise to the multimodal representations of our conscious experiences. These experiences include the phenomenological awareness and the network interactions that create the experience itself (Moscovitch, 1995). As a result, the phenomenology of experience, or consciousness, is inseparably linked to EM. EM trace is made up of an ensemble of HPC-neocortical neurons with a sparsely coded HPC component creating a spatial framework acting as an index (Nadel, 2008; O'Keefe & Nadel, 1978; Teyler & Rudy, 2007) to neocortical components, creating the representation of experience with perceptual, emotional and conceptual details, infusing it with the sense of autonoetic consciousness.

The HPC is at the top of largely cortical systems, made up from the ventral and dorsal streams, and at the later stages integrate information from the previous ones. This results in the building of more complex representations and, through back projection, influencing the operation of earlier stages (Nadel & Peterson, 2013).

When the HPC receives input from the entorhinal cortex, which in turn receives its input from the perirhinal (PRC) and parahippocampal cortices (PHC), it integrates the information about object representations from the PRC and scene representations from PHC. This frames the spatial relations amongst the numerous parts of the environment and locates those relations and features within that spatial frame (Bird & Burgess, 2008; Hassabis & Maguire, 2009; Nadel, 2008). It is argued that HPC-mediated memories reflect relational associations due to both the separate elements of an event and their relations preserve their distinctiveness (Eichenbaum, Yonelinas, & Ranganath, 2007; Olsen, Moses, Riggs, & Ryan, 2012).

The HPC receives inputs from both posterior neocortex and anterior structures such as the amygdala, the anterior temporal cortex and the prefrontal cortex with all of the structures playing important roles in EM through the interaction with the HPC. While it is not part of the medial temporal memory system amygdala still has connections to it and many other cortical and subcortical regions. Through these connections, amygdala influences and enhances memory for emotionally arousing information (Kensinger, 2009; LaBar & Cabeza, 2006; Phelps & LeDoux, 2005; Roozendaal et al., 2009). This enhancement reflects the encoding and consolidation of emotional information with the special part being played in the recollection of emotional information (Sharot et al., 2004). This has been observed even when the emotional information being retrieved is 1-year old (Dolcos et al., 2005). At the posterior end, based on the input from the posterior cortex, the HPC representations capture information about local spatiotemporal aspects of an episode. On the other hand, representations at the anterior end, capture global aspects of an episode such as the general context and the emotions attached to it. Thus the two types of representations arise from differences in input-output connections in the HPC (Poppenk et al., 2013; Strange et al., 2014). It can be seen that the mentioned integration maps onto the Tulving's WWW components of EM. The relational associations are the links connecting the item and spatio-temporal information about an event. In addition to supporting Tulving's original definition of EM, these findings also provide more weight for the use of the WWW task is EM testing.

It is important to note that not all components of the HPC or cortical structures that interact with it are activated at the same time or in the same tasks. Instead, depending on a task, subsets of components form process-specific alliances (PSAs) (Cabeza et al., 2018; Moscovitch et al., 2016). As mentioned above, posterior neocortical components with the posterior hippocampus (pHPC) determine the local spatio-temporal aspects of the episodes whereas anterior components of the HPC with anterior temporal lobe, PFC and amygdala represent emotional aspects. Encoding and retrieval of information are also regulated by control structures encompassed by the PSAs. All these components interact with each other earlier or later in the hierarchy which leads to the involvement of the HPC and in turn EM (Moscovitch et al., 2016).

1.6.1.1. Episodic details

Research shows that extensive bilateral damage leads to global anterograde amnesia affecting acquisition, retention and retrieval of EMs including specific details, themes and general structure (Hornberger & Piguet, 2012; Rosenbaum et al., 2009). However, the episodic details are most severely affected. When damage is more focused on smaller parts of the HPC, or damage is unilateral, the acquisition of semantic or gist information is relatively spared compared to EM (Winocur et al., 2010; Winocur & Moscovitch, 2011) with the best example of that being the patient H.M (Squire & Wixted, 2011).

Patients with unilateral temporal lobe epilepsy or lobectomy that included HPC showed that memory for perceptual details was affected the most (St-Laurent et al., 2014). In comparison, memory for more global details such as story elements showed less impairment with external details (not unique to the episode) being preserved (St-Laurent et al., 2014). Similar patterns have also been observed in patients with MTL lesions or impairments within a number of different disorders such as Alzheimer's, later stages of frontotemporal dementia or transient epilepsy (Piolino et al., 2009; Viard et al., 2012; Winocur & Moscovitch, 2011).

Functional magnetic resonance imaging (fMRI) studies have shown that HPC activation is modulated by the number of details or the vividness of the event. This HPC sensitivity to details and vividness have also been observed for more generic and repeated events such as family dinners which is consistent with the findings that HPC is inclined to represent details (Rosenbaum et al., 2009; Rubin & Umanath, 2015; Winocur & Moscovitch, 2011).

Recognition studies have shown that similar regions are activated during recollection (but not familiarity). This has been observed when recollection was tested both subjectively, such as asking participants to rate if items evoked recollection or knowing (familiarity), and objectively, by asking if participants recognised elements of the context in which the item appeared (Skinner & Fernandes, 2007). The regions in the so-called recollection network that are active during recall of vivid memories include HPC, PHC and medial prefrontal cortex (Rugg & Vilberg, 2013; Svoboda et al., 2006). What is important is that the activation of the HPC is associated with the amount of detail that is being retrieved (Rugg & Vilberg, 2013) but not with the memory strength (Migo et al., 2012; Montaldi & Mayes, 2011; Squire & Wixted, 2011).

1.6.1.2. Spatial details

It has been argued that the HPC is necessary for the construction of scenes which work as a scaffolding supporting memory for events (Hassabis & Maguire, 2009). Studies investigating HPC lesions have shown that coherent scene construction is dependent on the HPC (Maguire & Mullally, 2013). What is more is that there is a great overlap of brain regions activated during both spatial and EM tasks even if the EM tasks do not have a clear spatial component (Spreng et al., 2009). Interestingly, Robin et al., (2015) found that even if the episodic narrative lacked any spatial information, participants added it during encoding or recall which might explain some of the overlaps.

Evidence also shows that memory for events is aided by familiar spatial information (Robin & Moscovitch, 2014) and that regions in HPC associated with memory for events interact with regions in HPC associated with spatial memory even at a single cell level (Miller et al., 2013). Chadwick, Hassabis, Weiskopf, and Maguire (2010) asked participants to watch and recall film clips in which two different events happened in two different spatial locations. They then used multivariate pattern analysis to neutrally differentiate the retrieved memories from one another. The authors found that classification accuracy for distinct episodes was better than chance only in the HPC. When classification was in regards to differences in spatial location and not events, only locations were accurately classified and only in the HPC.

1.6.1.3. Temporal details

Evidence has been accumulating for the HPC involvement in the formation, maintenance and retrieval of temporal associations and their relations to events. Researchers consider the management of temporal information to be one of the core function of the HPC (Dalla Barba & La Corte, 2013; Davachi & DuBrow, 2015; Eichenbaum, 2014).

Schacter et al. (2012) present three main aspects of temporal processing: 1 – the temporal tag associated with different moments in the event; 2 – the coding of the temporal order of elements within and across episodes; 3 – the subjective sense of time which allows identifying if the experience occurred in the distant or near past. The first two aspects are accounted for by time cells (Eichenbaum, 2004). It has been proposed that there is a hippocampal mechanism that constructs scale-invariant representations of time which serve as the contextual/neuronal settings in which events are embedded (Howard & Eichenbaum, 2013). This could explain the temporal order effects in memory and the reduction of temporal precision with increased temporal distance. Contiguity has been shown to be a determinant of temporal order and is shown to be better for elements within an event, compared to elements across events, which shows contiguity's importance for the event segmentation (Davachi & DuBrow, 2015). Nevertheless, a question remains why memory for temporal information is poorer when compared to spatial information. It has been shown that cerebellum, PFC, posterior parietal cortex and basal ganglia mediates memory for the duration and temporal order (Danckert et al., 2007; Davachi & DuBrow, 2015; Moscovitch, 1992) however the exact neural mechanisms that underlie our subjective sense of time are not well understood.

1.6.2. The frontal lobe

In addition to the central role of the HPC in EM encoding and retrieval, the role of the frontal cortex in learning and memory is also well recognised (Benjamin, 2007; Mackey & Curtis, 2017; Wheeler et al., 1997) and damage to it has been linked to impaired episodic recognition and recall with recall showing a more pronounced impairment (MacPherson et al., 2008, 2016; Stamenova et al., 2017). The deficit in recognition shows the frontal lobe's contribution to the encoding of information while the more distinct deficit in recall shows its importance to the retrieval of information (MacPherson et al., 2016). Functional imaging reveals that the degree of activity in the frontal regions during incidental encoding predicts EM performance (Wagner et al., 1998).

Frontal activity, notably in the right hemisphere, is associated with retrieval mode – a basic and necessary condition of remembering past experiences. It refers to a neurocognitive state in which one mentally holds, in the background of attention, a fragment of one's own past. In this mode, one treats incoming information as retrieval cues for a specific event and refrains from task-irrelevant processing (Herron & Wilding, 2004; Lepage et al., 2000; Nyberg et al., 1995; Tulving, 1983). This translates to a brain state, established and maintained by instructions given in an episodic retrieval task. Episodic memory retrieval mode manages item-related processes such as the recollection of past events cued by a trigger and thus is a critical condition for remembering past events (Lepage et al., 2000; Simons & Spiers, 2003).

It is important to mention that the episodic retrieval mode is based on the hemispheric encoding retrieval asymmetry (HERA) model (Lepage et al., 2000; Nyberg et al., 1996; Tulving et al., 1994) according to which episodic retrieval is based on the right frontal regions whereas left regions are important for episodic encoding. However, the activation pattern observed in the HERA model is not an absolute feature of cortical activity during memory tasks. It has been shown that the activity is affected and in some cases eliminated depending on the nature of the material that is being memorised such as verbal content, difficulty, level of detail etc. (e.g. Buckner, Kelley, & Petersen, 1999; Ranganath, Johnson, & D'Esposito, 2000; Sandrini, Cappa, Rossi, Rossini, & Miniussi, 2003).

Now that the neuroanatomy of EM has been overviewed, it is possible to move on to discussing how memories get integrated into the knowledge networks. This happens through a process called memory consolidation.

1.7. Episodic Memory consolidation

So far, EM has been discussed in relation to encoding and retrieval. What has been left out is memory storage and the involved changes. While some experiments tend to test retrieval immediately after encoding (e.g. Plancher et al., 2012) some do not (e.g. Holland & Smulders, 2011). The time interval between encoding and retrieval is crucial as the encoded information passes through a process called consolidation and can be stabilised or weakened, if insufficiently activated or useful (Benson & Feinberg, 1977; Payne et al., 2008; Talamini et al., 2008). The reason for discussing this process in more detail is that it has been long suggested that new memories need time to stabilise and that such memories are prone to interference from other incoming information (Dudai, 2012). As such, depending on the length of the retention period, EM can undergo great changes. One example of that is the earlier discussed (see section 1.5) Remember-to-Know shift, which as it will be discussed later on is based on memory consolidation (Cermak, 1972; Dewhurst et al., 2009; Herbert & Burt, 2001, 2004).

The original account of memory consolidation was proposed by Müller and Pilzecker (1900). During their studies, they concluded that the physiological processes that were strengthening the associations between syllables read during their experiments continued to strengthen the associations for a period of time even after the experiments, albeit with a reduced effect (Lechner et al., 1999). At present, consolidation refers to the progressive stabilisation of long-term memory after its acquisition. It also includes the phases during which the stabilisation is presumed to take place (Axmacher et al., 2009; Diekelmann & Born, 2010; Dudai, 2012; Piolino et al., 2009). A large body of literature now suggests that consolidation mechanism plays a role in memory enhancement and reorganisation, with newly formed memories going from weak and labile to strong and enduring over time (for a review see Rasch & Born, 2013).

After a memory is initially acquired, a series of cellular, molecular and systems-level changes take place. At the neuronal level, consolidation occurs within minutes to hours resulting in memory stabilisation. This was found by a number of studies that inhibited particular proteins needed for the memory consolidation process (Born et al., 2006; Dudai, 2004; Dudai et al., 2015; Furman et al., 2012; McClelland et al., 1995) with the most used example being the goldfish study by (Agranoff et al., 1966). The study showed that memory was resistant to a protein synthesis inhibitor after an hour. It is important to note that the majority of the literature investigating the neuronal changes associated with consolidation are based on animal models, however, there is some research done with humans (Kandel, 2001).

Systems-level consolidation builds on synaptic consolidation and refers to the redistribution and reorganisation of memory representations for long-term storage (Dudai, 2004; Rasch & Born, 2013; Stickgold & Walker, 2007). While the exact processes and mechanisms behind the systems-level consolidation (and memory consolidation in general) are debatable (Stickgold, 2005) the two main views will be discussed – the standard model of consolidation and the memory trace theories.

1.7.1. Models of consolidation

One of the main models that try to explain memory consolidation is the standard consolidation theory (SCT) (Squire, 2004; Winocur et al., 2010; Winocur & Moscovitch, 2011). The model proposes that memory consolidation is dependent on two memory stores: a hippocampus-dependent short-term store and a long term store distributed throughout the neocortex (Frankland & Bontempi, 2005). Memories are initially encoded in both hippocampus and neocortical networks however the neocortex is not able to support the memory on its own. This is due to the distributed nature of the memory representations, encompassing various multimodal components of an experience which links back to the WWW EM information. As a result, the hippocampus is critical in the early stages of memory encoding to act as an 'integrator' and 'binder' of the cortical patterns of activation which lead to a coherent memory representation. With time, reactivations of these memories lead to a gradual strengthening of the neocortical connections, meaning that memories are integrated into the pre-existing knowledge networks and become independent of the hippocampus (Moscovitch et al., 2016). These two processes benefit memory by maintaining the hippocampus capacity for future learning and by reducing the risk of interference and memory 'overwriting' (Frankland & Bontempi, 2005; McClelland et al., 1995; Squire, 2004). As such EM generally have a relatively short lifespan as they become semanticised as part of the consolidation process.

In contrast to the SCT, the multiple trace theory (MTT) (Moscovitch & Nadel, 1998; Nadel & Moscovitch, 1997, 1998) and its more recent variant Competitive Trace Theory (CTT) (Yassa & Reagh, 2013) posit that all episodic information is encoded by the hippocampal neurons which act as an index for the neocortical neurons. These links between the hippocampal and the neocortical neurons are constituted as memory traces for the episodes. As reactivations of these traces usually occur in different contexts, it results in the creation of multiple traces that share some or all of the information about the initial episode. Because of the multiple traces related to the same episode, the extraction of semantic information is facilitated, leading to the extraction of semantic representations from the episodes. This information gets integrated into the wider network of sematic knowledge and becomes independent of the initial episode.

Trace theories differ from the SCT in that the hippocampal complex remains involved in the storage and retrieval of episodic representations regardless of their age, in the trace theory models. According to the MTT, the hippocampal traces contain contextual information about the episode whereas cortical traces are thought to be semantic and context-free. This relates to the definition of EM as a composition of the WWW information and the hippocampal traces containing the spatio-temporal Where and When information. Due to the two types of traces, the retrieval of remote semantic memories does not involve the hippocampus, but retrieval of remote episodic (context-rich) memories does. This creates a critical distinction between episodic and semantic memories. While both of them are influenced by the hippocampal complex, only the semantic memory becomes independent of it throughout the consolidation process (Kinsbourne & Wood, 1975; Kisker et al., 2019; Tulving, 1972).

While the MTT provides the best explanation for memory consolidation and is a very valuable theory that helped memory consolidation research move forwards (Sutherland et al., 2019) it does come with flaws. It has been shown that the hippocampal activation for remote memories can be explained by scene construction and not reactivation of hippocampal circuits used in the long-term storage. Additionally, a number of animal studies have failed to support the notion that gist-like semanticised memories get strengthened in the neocortex and become less reliant on the hippocampus (Sparks et al., 2011; Thapa et al., 2014). Instead, it was found that memories remain hippocampus-dependant. As such an alternative to MTT was proposed - Competitive Trace Theory (Reagh & Yassa, 2014; Yassa & Reagh, 2013).

According to CTT, and unlike MTT, memory traces are not stored in parallel but compete for representation in the neocortex. The main tenet of CTT is that memories go through reconstruction and reconsolidation during retrieval. As mentioned earlier, when memories are retrieved they become liable and susceptible to interference. After this retrieval and reconsolidation memories change and are updated. Over time, this process leads initially rich EM to become semanticised and lose contextual details. As in the MTT, retrieval of a memory creates a new trace, however instead of both traces coexisting, CTT posits that the traces compete for representation in the neocortex.

While the present thesis was not aimed at investigating specific consolidation theories, due to the amount of supporting evidence, the memory trace theories were appraised and preferred as the underlying mechanism behind memory consolidation. Regardless of the theory, the evidence for consolidation comes from the delayed recall of memories. While some EM retrieval might occur 24 hours after original encoding (e.g. Takashima et al., 2006), some might be years later (e.g. Barry & Maguire, 2019), nevertheless, there is one underlying and important aspect connecting these studies – sleep. The importance of sleep lies in the fact that most of the memory consolidation relies on it (Inostroza & Born, 2013).

1.7.2. The role of sleep in memory consolidation

Sleep can be defined as a natural and reversible state of reduced inactivity, reduced responsiveness to external stimuli and loss of consciousness (Rasch & Born, 2013). The effects of sleep deprivation, conservation of sleep in mammals and the sleep rebound observed after sleep loss shows that sleep serves a highly important purpose in memory consolidation. Nevertheless, there is no unified theory of sleep function (Fuller et al., 2006; Saper, 2013). In addition to the conservation of energy (Schmidt et al., 2017), brain thermoregulation and detoxification (McCarley, 2007; Saper, 2013), sleep offers optimal conditions for 'offline' memory consolidation (Diekelmann et al., 2009; Dragoi & Tonegawa, 2011; Stickgold, 2005). While it has been shown that both declarative and non-declarative memories benefit from sleep (Diekelmann & Born, 2010; Stickgold, 2005; Stickgold & Walker, 2007), the research regarding it started with an interest in forgetting.

A number of studies by Ebbinghaus (1885) on forgetting of lists of nonsense word pairs showed a forgetting curve. During the first hours, learning was followed by rapid forgetting which levelled out after several days. In addition to that, he noticed that forgetting was reduced if during the retention interval that contained sleep (for an early review see Ormer, 1933). Indeed, research on sleep deprivation showed impaired remembering (Patrick & Gilbert, 1896). Heine (1914) was one of the first to show that learning before a night's sleep resulted in less forgetting than learning after a night's sleep. All these findings created the early groundwork for research in sleep's role in memory.

Forgetting and its cause became the main research interest during the 20th century. Two concepts were proposed to explain the cause of forgetting: decay and interference. Decay account explains forgetting as a decay of memory traces which occur over time and results in time-dependent forgetting. In interference account, new information interferes and overwrites old memory traces which results in forgetting (McGeoch, 1932). These accounts were investigated in a study by Jenkins & Dallenbach (1924). Two participants were examined for their nonsense syllable retention after. The participants were tested after 1, 2, 4 and 8 hours after learning, every day for two months. The time between the testing was filed either by wakefulness or sleep. The results showed that when the retention period was filled with sleep, forgetting was lower. Authors concluded that forgetting is not about the decay of old information but about interference between newly acquired and already existing information.

Following these findings, many studies have confirmed that sleep has a positive effect on memory (Idzikowski, 1984; Koulack, 1997; Newman, 1939). The fundamental idea was that sleep acts as a passive shelter from interference (Ellenbogen et al., 2006). However, the hypothesis that forgetting is simply based on the time elapsed from learning is contradicted by the fact that interference is much stronger just after learning compared to later times (Ebbinghaus and his learning curve). This shows that consolidation is time dependant and memory traces are strengthened with time (for a review see Lechner, Squire, & Byrne, 1999). Indeed, the earlier discussion of memory consolidation shows that consolidation is an active process through which (episodic) memories become semanticised over time and lose detail (Donix et al., 2010; Martinelli et al., 2013; Murphy et al., 2008).

A time-dependent effect of sleep on the formation of memories was shown by studies which compared the effect of sleep just before learning to sleep at a later time (Benson & Feinberg, 1977; Payne et al., 2008; Talamini et al., 2008). In other words, the closer sleep is to learning the better memory retention becomes. For example, in a vocabulary learning experiment participants that went to sleep after 3 hours showed better retention compared to participants that went to sleep after more than 10 hours (Gais et al., 2006). Similar results were found in a study in which, after 24 hours, word pair recall was better if learning was immediately followed by sleep compared to a full day of wakefulness (Payne et al., 2012). These findings cannot be explained by simple interference reduction as the time between learning and retrieval and general time spent asleep was identical to both wake and sleep conditions. Instead, these findings show the importance of a time window of reduced interference just after learning.

While it can be seen that sleep does play a role in memory consolidation, the exact role and processes are still being investigated. As with the earlier discussed general memory consolidation, there are a number of main theories that try to explain how memories are consolidated in sleep.

1.7.3. Models of sleep-dependant memory consolidation

Sleep is composed of 90-minute cycles of non-rapid eye movement (NREM) and rapid eye movement (REM) sleep stages (Iber et al., 2007; Marshall & Born, 2007). The first part of the night is dominated by deeper, slow-wave sleep (SWS) whereas the second part of the night is dominated by REM sleep. NREM is further divided into three stages; N1, N2 and N3 with each stage representing a deeper level of sleep (Fuller et al., 2006). All these stages cycle through the night. This cycling gives rise to the dual-process hypothesis and the sequential hypothesis.

The dual-process hypothesis posits that SWS and REM sleep affects different types of memory (Born & Wilhelm, 2012; Diekelmann et al., 2009; Rauchs et al., 2004). Studies have found that declarative memories such as word-pairs, spatial locations and word recognition showed more benefit from early SWS-rich periods of sleep (early part of the night). On the other hand, REM-rich periods of sleep (late part of the night) enhanced non-declarative (e.g. procedural) and emotional declarative memories (Born et al., 2006; Marshall & Born, 2007; Peigneux et al., 2001; however see Rauchs et al., 2004). However, it is important to note that early sleep is only associated with SWS and late sleep with REM. As mentioned earlier, sleep goes through cycles and as such both types of sleep can occur in both halves of the night. This cycling gives rise to the sequential hypothesis.

The sequential hypothesis (Giuditta, 2014; Giuditta et al., 1995) suggests that the different stages of sleep and their cycling work together in memory consolidation. For example, SWS might help with the consolidation of the temporal information of an episode, while REM sleep helps with the consolidation of emotional and spatial parts of the memories (Rauchs et al., 2004). The hypothesis states that this cycling of sleep stages help to integrate memories to the knowledge networks (Cairney et al., 2015).

However, the sequential hypothesis suffers from a problem, which can also be seen in the dual-process hypothesis, of trying to map specific sleep stages to specific memory types. For example, research has shown that certain tasks that should be associated with consolidation in REM sleep, show consolidation in SWS and vice versa (Backhaus & Junghanns, 2006; Nielsen et al., 2015). This shows that sleep dependant consolidation is more complex and cannot easily be grouped to specific memory types and sleep stages. What is more, these hypotheses are difficult to relate to the general memory consolidation theories discussed previously. As such, another model of sleep dependant memory consolidation is presented - The Active Systems Consolidation Model.

The Active Systems Consolidation Model (ASC) proposes that memories are redistributed through the systems-level consolidation which is driven by slow oscillations, SWR (sharp-wave ripple) and sleep spindles taking place during SWS. (Frankland & Bontempi, 2005; Walker, 2009). This model hypothesises that SWRs help with the communication between the hippocampus and the neocortex. During this, slow-wave oscillations and sleep spindles synchronise to induce long-term plastic changes within cortical networks (Rasch & Born, 2013). Due to this, it is thought that sleep facilitates memories to become less dependent on hippocampus and more dependent on the neocortex through memory reactivations. Even from this brief description, it can be seen that ASC links well with the previously discussed synaptic and system consolidation and more importantly with the trace theories of memory consolidation.

Slow oscillatory activity synchronises activity from the thalamus and hippocampus leading to spindle-ripple events which mediate the hippocampal-neocortical information shift (Born & Wilhelm, 2012). It is hypothesised that sleep spindles that reach the neocortex prepare the needed neural networks for the synaptic adjustments for the long-term storage of information. Consequently, the synchronous connection from the thalamus and hippocampus to the neocortex is critical for the redistribution of declarative information.

The evidence supporting the ASC model comes in many different forms. Notably, positive correlations have been observed between memory performances, time spent in SWS and spindle activity after many different memory tasks (Clemens et al., 2005; Durrant et al., 2011, 2013; Gais et al., 2002). Investigations of local brain regulation during sleep show increased coherence of slow oscillations in brain regions that were active in pre-sleep learning (Huber et al., 2004). Transcranial direct current stimulation (tDCS) studies have also shown increases in slow oscillations, sleep spindles and memory retention after tDCS application. This is due to the tDCS inducing slow-oscillation field potentials (Barham et al., 2016; Marshall et al., 2006, 2011).

As mentioned before, the ASC model suggests that memory consolidation is driven by memory reactivations. This finding gives support to the earlier discussed competitive trace theory which was chosen as the underlying memory consolidation theory for this thesis. The support comes from animal studies and the research on the hippocampal place cells (Pavlides & Winson, 1989). It was found that these cells, which fire in specific spaces during exploration and thus encode place representations, 'replay' those representations during sleep. During that 'replay', the order of cell firing was largely similar to the order observed during the initial exploration task (Deuker et al., 2013; Peigneux et al., 2004; Skaggs & McNaughton, 1996). This replay has been observed during SWS or more particularly during SWR (Diba & Buzsáki, 2007; Ji & Wilson, 2007; Roumis & Frank, 2015) successfully predicting memory performance (O'Neill et al., 2010). The replay has been observed in both the hippocampus and in the neocortex and during both SWS and REM. The difference is that the replay during SWS is 'fast-forwarded' by about 15-20 times faster than in the real world, whereas replay during REM is close to real-time (Ji & Wilson, 2007; Lee & Wilson, 2002; Louie & Wilson, 2001).

However, this replay was also found during wake and not only forwards but also backwards. Foster & Wilson (2006) have observed a reverse replay in rats immediately after a run on a track. This reverse replay declined with familiarity. Similar results were observed by Diba & Buzsáki (2007) with a forward replay at the beginning of the run (as if rats were anticipating) and reverse replay at the end of the run. With this evidence, it has been proposed that the replay that happens in both wake and sleep after the experience is due to consolidation whereas reverse replay, that happens during wake, may subserve episodic binding (Carr et al., 2011). Indeed, human functional brain imaging has shown post-stimulus activity in the hippocampus which predicted later memory performance (Ben-Yakov & Dudai, 2011). This activity might indicate the EM binding and the beginning of the consolidation.

1.7.4. Sleep-dependant consolidation of Episodic Memory

Neuroimaging and lesion studies have shown that contextual information of EM depends on the hippocampus whereas item memory is mainly supported by the extra-hippocampal structures with the main one being the perirhinal cortex (Davachi, 2006; Eichenbaum et al., 2007). Even short periods of sleep such as napping have been shown to enhance contextual but not item information (van der Helm et al., 2011) with the enhancement correlating with the amount spent in SWS. This was shown in a study in which participants had to learn two lists of words (item information) while facing two distinct posters (context information). Similar findings were also found in a study looking at all of the episodic components. Rauchs et al. (2004) asked participants to learn two lists of words (item – what), one after another (temporal – when), at the top or the bottom of a page (spatial – where). Again, forgetting was lower for the temporal information in SWS-rich sleep and spatial information showed enhancement after REM-rich sleep. This type of contextual strengthening has been observed in a number of studies (Drosopoulos et al., 2007; Foster & Wilson, 2006; Griessenberger et al., 2012).

Sleep's effect on EM can also be seen from studies using the Remember/Know paradigm (Rugg & Yonelinas, 2003; Yonelinas, 2001). As it has been explained previously, remembering or recollection relies on hippocampus while knowing or familiarity can be achieved by extra-hippocampal regions alone. A number of studies have shown a post-sleep enhancement of explicit recollection of memories while familiarity based judgements were not affected (Atienza & Cantero, 2008; Daurat et al., 2007; Drosopoulos, 2005; Rauchs et al., 2004). Some studies have also shown a link between this enhancement of recollection and the SWS occurring post-learning (Daurat et al., 2007; Rauchs et al., 2004).

Evidence also shows that sleep preferentially consolidates EM that is emotionally arousing (Rauchs et al., 2011; Wilhelm et al., 2011). In general, emotional events are remembered better (both in accuracy and vividness) than neutral events and this is shown by a number of EM studies using emotional stimuli (Kensinger & Ford, 2020; LaBar & Cabeza, 2006). For example, in a study by Payne et al. (2008) participants were presented emotional (car crash) and neutral (car) scenes. They have found that participants had a superior memory for the emotional objects when learning was followed by sleep compared to wake. Their study has shown that sleeping preserves emotionally salient information. This shows two important things: EMs are not consolidated equally and EM consolidation is affected by sleep.

Research shows that the amygdala plays a major role in in the enhancement of emotional memory through modulation of other brain regions, the main one being the hippocampus (Kensinger & Ford, 2020; LaBar & Cabeza, 2006). If the amygdala is activated during encoding, performance at retrieval is enhanced, therefore amygdala's involvement might persists after encoding and influence consolidation. Indeed studies have shown that amygdala activity is enhanced during sleep, which suggests an interplay between the amygdala and the hippocampus and its influence on emotional EM consolidation. Nevertheless, as discussed earlier, the consolidation of EM is not bound to just one sleep stage. While some studies have shown the importance of REM sleep on emotional memory consolidation (Groch et al., 2013; Hutchison & Rathore, 2015; Nishida et al., 2009; Wiesner et al., 2015) others have shown the role of NREM sleep (Cairney et al., 2014; Hauner et al., 2013; Lehmann et al., 2016) or found no clear stage dependence (Ashton et al., 2018; Cairney et al., 2015; Cellini et al., 2016; Morgenthaler et al., 2014).

The earlier mentioned emotional EM study by Payne et al. (2008) also brings up an important point about EM testing as a whole – ecological validity. Instead of using lists of words as in more traditional paradigms, the study used arguably realistic images of car crashes. This use of realistic and more life-like stimuli in EM research is one of the main interests of the present thesis and as such will be explored in more depth.

1.8. Ecological validity of Episodic Memory testing

In the past couple of decades, a change can be seen in cognitive psychology. A trend is emerging towards a more ecological approach to investigating human behaviour (Gibbs, 1979; Grewe et al., 2014; Neisser, 1985; Reggente et al., 2018). The term 'ecological validity' has become popular among cognitive researchers, undergraduate texts, research methods and dictionaries of cognitive psychology (Ashcraft, 1994; Coolican, 1992; Eysenck, 1990; Eysenck & Keane, 2000). Diverse areas of psychology such as neuropsychology (Gioia & Isquith, 2004; Sbordone, 2008; Sbordone &

Guilmette, 1999; Silver, 2000), child development (Fabes et al., 2000) and cognitive ergonomics (Hoc, 2001) shown the need for higher ecological validity.

For example, while the initial purpose of the neuropsychological assessment was to diagnose a person with a brain injury or a disease and then define the brain-behaviour relationship, today clinical neuropsychology is more interested in making prescriptive statements about person's everyday functioning (Long, 1996). This change in the role of neuropsychologists led to an increased emphasis on the ecological validity of neuropsychological assessments. Neuropsychologists need to demonstrate either (or both) verisimilitude and veridicality to establish ecological validity of a measure (Franzen & Wilhelm, 1996). Verisimilitude means that researchers need to emphasise the need for the data collection methods to be as close to real-life tasks as possible. For a measure to show veridicality, it needs to reflect and predict real-world tasks (Chaytor & Schmitter-Edgecombe, 2003; Ready et al., 2001; Silver, 2000).

Correlations between classical neuropsychological tests, subjective memory complaints and everyday memory functioning have been shown to be inadequate (Chaytor & Schmitter-Edgecombe, 2003; Reid & MacLullich, 2006). The majority of the tests have been developed following a "construct-driven approach" (Burgess et al., 2006; Parsons et al., 2017). The tests are built starting with a solid theoretical base and evaluate abstract constructs without referencing real-life performance or behaviour (Parsons, 2015; Parsons et al., 2017). Due to the movement towards more ecologically valid assessment, a "function-led approach" became more popular. The "function-led approach" focuses on the direct observation of behaviour which should lead to a more valid measure (Parsons, 2015; Parsons et al., 2017). An example of this approach is the Rivermead Behavioural Memory Test (RBMT; (Wilson et al., 1985, 2013). To evaluate the memory abilities of people with brain injuries the test includes a series of daily-life tasks such as remembering an appointment, recognising a picture or encoding and storing a route.

Another aspect of ecological validity is that conventional memory tasks are unreliable in capturing is the complexity of memory functioning and the components that make it up. As discussed before, EM is made up of a number of components which are merged through a process called binding (Kessels et al., 2007). In a clinical setting, EM is usually assessed by asking patients to remember a verbally presented story

(Spinnler & Tognoni, 1987), a list of words (Delis et al., 2000) or a picture (Shin et al., 2006). As a consequence, these tests evaluate memory components in isolation which results in the lack of episodic retrieval which is essential if one wants to follow Tulving's description of EM (Tulving, 2002).

The same problem can be said about many other EM tests that have been discussed previously. For example, in many of the discussed experiments that used the WWW, Source or Recognition tasks, the stimuli were lists of words or pictures. In the study by Pause et al., (2010) participants had to choose from quadrants on a computer screen. In a study by (Davachi et al., 2003) participants had to complete both recognition and source tests with the stimuli being visually presented adjectives. A study by (McElree et al., 1999) investigated recognition memory by using the response speed method and used visually and verbally presented words. None of these studies investigated EM in an ecologically valid and function-led fashion. When discussing the autobiographical interview, while it looks at memory for real-life events, there is the problem of semantisation. The WWW study (discussed in section 1.4.1) by Holland & Smulders (2011) can be said to explore EM, both as components and as a combined WWW, in a (relatively) ecological fashion. However, these kinds of experiments involve a lot of planning, use space that might not be available to some and finally cannot be completely replicated. The last point is due to the fact that while the procedure and objects could be recreated by other researchers, the spatial layout might not. This creates another problem of reproducibility. What is needed is a way of testing EM in an ecologically valid fashion that could be fully replicated by anyone. One way of achieving this is by using virtual reality.

1.9. Virtual Reality as a solution to the problem of ecological validity

1.9.1. What is virtual reality?

The term "virtual reality" (VR) is frequently and interchangeably used to refer to many different experimental apparatuses (Wilson & Soranzo, 2015). A good definition of VR is given by Fuchs et al. (2011, p. 8): "Virtual Reality is a scientific and technical domain that uses computer science and behavioural interfaces to simulate in a virtual world the behaviour of 3D entities, which interact in real-time with each other and with one or more users in pseudo-natural immersion via sensorimotor channels.” The real-time interaction means that a user should be able to directly interact with the system (e.g. navigation) without or with a minimal delay between the user’s input and the associated response from the virtual environment. Using this criterion, VEs observed by a subject without any interaction does not fall under the VR description. The concept of immersion is more complicated. In VR research, immersion can be described as an extent to which the VR system creates a naturalistic representation of the sensory and interactive elements of a specific VE. It is a degree of how well the VR system recreates the sensorimotoric richness of the real world (Fuchs et al., 2011).

However, this definition of VR preludes the kind of apparatuses can satisfy the mentioned conditions. A number of existing systems, to a varying degree, can be categorised as VR. What is important is to understand that, depending on the used apparatus, an experimental task could change considerably. For example, an act such as walking in a VE can range from a sophisticated treadmill system to a simple press of a button. This difference might lead to inappropriate or impossible comparisons between experiments (e.g. Ruddle et al., 1999; also see Smith, 2019). Due to this, it is important to understand the different types of VR systems.

1.9.2. Types of virtual reality

1.9.2.1. Desktop-VR

Desktop-VR uses a standard computer screen to display VEs (Furht, 2008, p. 963). Interaction with VEs is usually performed using a mouse and a keyboard. Due to this, Desktop-VR is widely available and cost-effective. The hardware needed to run VEs and software to create them is easily accessible. Additionally, the fact that desktop computers are becoming an integral part of our everyday lives, subjects are much more familiar with their input devices which leads to quicker training phases. Desktop-VR has been widely used in psychological research for decades, however, the exact name can differ (e.g. screen-based VR).

The main drawback of Desktop-VR is the two-dimensional presentation of VEs. Due to no stereoscopy, only monocular depth cues are available to indicate the distance of objects presented in the environments. Additionally, the interactions with the VEs are not analogous to the ones that it tries to simulate. For example, looking around is done through a movement of a mouse and walking through a push of a button on a keyboard. The lack of motoric component reduces the levels of immersion and limits the usefulness of Desktop-VR in the exploration of ecological memory.

1.9.2.2. Simulator-VR

The main difference between the Simulator-VR and other types of VR is its use of external visual displays and specialised input devices. A usual Simulator-VR system consists of multiple projector screens or display panels which lead to the user feeling surrounded by the imagery. The most sophisticated and well-known systems used in research are the Computer-Aided Virtual Environments or CAVEs (Furht, 2008). These systems comprise of whole rooms dedicated for the display of VEs with features such as head tracking, special glasses for stereoscopic vision and floor-to-ceiling displays.

The main problem with the Simulator-VR systems is the cost associated with running them. The setups require separate customised rooms filled with projectors or screens, custom headsets and custom input devices. However, it is possible to create more affordable setups at a price of reduced immersiveness. For example, a study by Maillot et al. (2017) used a number of screens to create a U-shaped configuration around a participant. However, it is possible to argue that such kind of set-up may fall under the Desktop-VR systems.

1.9.2.3. HMD-VR

HMD-VR is characterised by the use of special viewing equipment. VEs are presented directly in-front of users eye through a head-mounted display (HMD) placed on the users head. HMDs are able to detect user’s head motion, such as the angle and velocity, and use that information to update the visual information that is being presented resulting in an ability to naturally look around in the VE (Furht, 2008). Additionally, the HMDs present images to each eye with a slightly shifted perspective. This allows viewing VEs stereoscopically due to the availability of binocular depth cues. Latest HMD-VR systems also come with hand-held controllers as input devices.

By tracking the spatial locations of these controllers, HMD-VR systems are able to map them in the 3D space, allowing users to 'see' where the controllers, are in the VEs.

There are a number of ways how movement can take form through HMD-VR. Users can be stationary (sitting or standing) and use a keyboard or a handheld joystick for movement and interaction. Using the VR controllers it is possible to walk around the VEs or teleport around them while staying stationary. Newer HMD-VR systems now allow users to physically walk around with their location being updated in the VE. This is done by sensors which track the location and rotation of the headset.

The main limitation of HMD-VR is the cost. While the hardware cost has been declining it is still relatively expensive. In addition to the cost of the headset, users need to have a powerful enough desktop PC (or a laptop) to run the software which adds to the overall cost. However, unlike with the Simulator-VR, it is not necessary to dedicate special rooms for it and also as the whole software runs on a desktop computer it is possible to run both Desktop-VR and HMD-VR setups.

1.9.3. Virtual Reality in memory research

The advantage of the large scale and realistic environments that can be provided by virtual reality (VR) was first utilised in spatial learning research. The same kind of problem of the, just discussed, ecological validity, was also pointed out in this field. It was acknowledged that navigation is not the same as table-top tests of spatial memory and that direct inferences should not be made between them. In a table-top test, all information is within one's field of view, which is not true in a complex real-life environment where most of the information cannot be seen. An example of what can be seen in studies on patients with topographical memory deficits. In these studies, patients have shown difficulties navigating in their (real-life) environments but displayed no impairments in the table-top spatial knowledge tests (Habib & Sirigu, 1987; McCarthy et al., 1996).

Some examples of the utilisation of VR in spatial memory research was the recreation of the Morris water maze (Hamilton & Sutherland, 1999) and radial arm maze (Leplow et al., 1998) - behavioural procedures mostly used with rodents. Another, a more sophisticated example, is a study by Maguire et al., (1998) which looked at brain activations and spatial navigation using a VR town that was presented on a computer screen. Participants had to navigate a complex virtual town to reach certain locations. In one condition, it was possible to reach those locations directly, while in the second condition, direct routes were not available and participants had to take detours. Even early research showed that cognitive maps of the environments created in this type of virtual exploration are comparable to those acquired in the real environments (Ruddle et al., 1997). Furthermore, the general representations of the environments are transferred when subjects subsequently navigate in the real environments (Arthur et al., 1997; Waller et al., 1998; Wesley Regian & Yadrick, 1994; Witmer et al., 1996). What is more is that VR environments elicit a stronger sense of 'presence' compared to table-top experiments (Held, 1992). The definition of 'presence' being the subjective experience of being in one place when one is physically in another. This was found, both in the early and more recent days of VR use. Research by Witmer and Singer (1998) observed a significant correlation between experienced 'presence' and performance in VR whereas a study by Schomaker, Roos and Meeter (2014) showed a correlation between ratings of presence and memory performance. Due to all of these findings and factors, the use of VR in spatial learning research is prevalent even today (Guderian et al., 2015; Horner et al., 2016; Konishi et al., 2017).

1.9.4. Virtual Reality in episodic memory research

VR would be a useful tool in the EM research especially with the spatial relationship component which links back to the spatial learning and cognitive map testing studies mentioned earlier. Indeed, one of the first EM studies, using VR, used a virtual town very similar to the one used in the earlier mentioned study by Maguire et al. (1998). The same group of researchers, that had a deep interest in spatial memory, tested the EM by asking participants to explore a virtual town in which they received a set of objects from two different virtual characters in two different places (Burgess, Maguire, Spiers & O'Keefe (2001). Participants then were placed back with each of the virtual characters and were asked which object was given to them, where it was given and which person gave that object to them. While their study was focused on brain region activations during the memory retrieval they did note that the lifelike events experienced in the VR led to brain activation not observed in non-VR laboratory studies. This links back to the problem of artificiality and lack of ecological validity in memory research.

As mentioned earlier while the non-VR experiments allow for great experimental control, the experiences in the real world, and therefore the EM for those experiences, consist of richer and more complex interactions and events (Burgess et al., 2001). Indeed, early EM research used verbal material. In these verbal paradigms, participants typically study a list of words and then are tested on that list. And again, even in the early days, it was pointed out that there was a discontinuity between spatial, temporal and intermodal continuities of real objects and events presented in laboratory-based research (Trevarthen, 1977). This can be seen by looking at the classical neuropsychological tools used to assess EM. For example, verbal material, such as words and sentences in the Logical Memory Test (Wechsler, 2008) or abstract figures in the Rey-Osterrieth Complex Figure (Meyers & Meyers, 1995) which none of them bear a close resemblance to the everyday memories. Due to this ‘closeness’ to the real-world experiences and the availability of VR software, more studies started coming out investigating EM using VR (King et al., 2002; Spiers, Burgess, Hartley, et al., 2001; Spiers, Burgess, Maguire, et al., 2001).

VR based EM tests represent a good compromise between the needed experimental control and an everyday-memory-like assessment of EM. Much of what people remember in the everyday life refers to complex events composed out of elements such as what happened, where it happened when it happened and a number of other multimodal details related to the event (Tulving, 2002). This, EM defining, What-Where-When approach benefits greatly from the advantages of VR. However, it is important to point out that all of the earlier mentioned VR studies presented their virtual environments (VEs) on a computer screen using Desktop-VR. While this type of memory testing is a better reflection of the real-world experiences there is a major issue with it. Desktop-VR still does not fully reflect the real-world as participants are focusing at the screen in front of them which reduces the immersion aspect of the screen-based exploration compared to the real-world experiences (Kinugawa et al., 2013; Zlomuzica et al., 2016).

1.9.5. Interim summary

From the earlier discussion of VR use in research and the descriptions of the different VR systems, it can be seen that VR can be a great tool in EM research. VR, in general, has been successfully used in EM research regardless of its type (refs). However, when considering the closeness to the real-life experiences and the cost-effectiveness, HMD-VR seems to be the best system to use over the Desktop-VR or Simulator-VR.

Indeed, a move can be seen towards HMD-VR use in psychological and more importantly EM research with more and more studies emerging utilising HMD-VR (Corriveau-Lecavalier et al., 2018; Davison et al., 2018; Ouellet et al., 2018; Parsons & McMahan, 2017). The studies argue that HMD-VR measures are better predictors of cognitive decline and also positively correlate than the more traditional clinical memory tests. As a result, the present thesis chose to use HMD-VR as the main method of experiencing episodes. The specific studies and their relevant findings such as the examples of HMD-VR use in EM research and comparisons of HMD-VR, Desktop-VR and the real-life, will be discussed in the further empirical chapters.

1.10. Thesis Aims

The overarching aim of this thesis was to explore how HMD-VR can be used to increase the ecological validity of EM testing.

While the definition of EM went through a number of revisions (Tulving & Markowitsch, 1998) the main concept has stayed the same – EM holds information about events (The What) and their spatio-temporal relationships (The Where and The When). The beginning of the present chapter was dedicated to the discussion and operationalisation of EM. This was followed by a discussion of memory consolidation which is an important topic in long-term memory research. Memory needs to go through a series of processes which stabilise it and make it resistant to change and forgetting. As our memories about events are typically from times other than the present it is important to test EM not just immediately after the event. Due to this, the next part of Chapter 1 was dedicated to the discussion of memory consolidation and its effects on

EM, that is, EM over time. The next part of Chapter 1 covered the different ways of testing EM and the common problem of a low ecological validity of those tests.

Virtual reality is increasingly popular as a tool in the cognitive sciences as it allows researchers to create virtual environments and situations that are very close to daily life while still having high experimental control (Lloyd et al., 2009; Plancher et al., 2010, 2012). Virtual reality use in EM research has become more prevalent with virtual environments becoming more sophisticated and more life-like. However, the majority of the virtual environments are still presented on a computer screen which still lacks the real-world immersion (Kinugawa et al., 2013; Zlomuzica et al., 2016). With the increase in popularity and availability of virtual reality, studies started to emerge, utilising the head-mounted display based virtual reality. By letting participants interact with rich multimodal environments and carry out sensorimotor activities, assessment of memory using this type of virtual reality allows obtaining data that is closer to the real-life compared to pen-and-paper tests, standard computer interfaces or virtual environments presented on computer screens (Mestre & Vercher, 2011). The general introduction to virtual reality and its uses in various types of memory research was discussed in the last part of Chapter 1 before focusing on its uses in EM research.

The following chapters explored how HMD-VR could be used in testing EM in a more ecologically valid fashion while also investigating the effect of sleep and time-based consolidation. Chapter 2 described the general methods used to investigate EM throughout the rest of the thesis. Following this, four chapters each present an experimental investigation. Chapter 3 and 4 examined how virtual environments and events within them can be presented through the HMD-VR setup and how EM changes over 24 hours (Chapter 3) and 30 day (Chapter 4) periods. The main aim of Chapter 3 was to explore how EM for life-like events differs to memory for static objects which are typically used in research. The secondary aim was to explore how these differences are affected by sleep-dependant memory consolidation. Chapter 4 continued with the previous two aims but extended the exploration of EM consolidation by testing EM over a 30 day period. Additionally, Chapter 4 explored how the different EM measures relate to each other. Both of these experiments were enclosed in a general interest on how custom-built VEs, presented through HMD-VR, can be utilised in function-led and more life-like EM testing.

Chapter 5 and 6 moved towards exploring the differences in EM performance between HMD-VR and Desktop-VR which, as discussed in Chapter 1, still dominates the literature. Chapter 5 compared how EM for the life-like events introduced in Chapters 3 and 4 differs when the same virtual environments are presented on a computer screen and a head-mounted display. Chapter 6 continued with the comparison of EM in the two VR systems with an addition of real-life condition. The thesis was finished with Chapter 7. The first part of Chapter 7 was a brief overview of all the experiments and their findings. Further sections discussed what the findings added to the literature on EM, consolidation and the use of virtual reality. The last part of the chapter and the thesis itself was finished with a general conclusion and the directions for future research.

Chapter 2 - General Methods

The current chapter presents details of methods that apply to the empirical studies in the thesis. Due to the first three studies using the same virtual environments and almost identical testing procedures, this chapter will mostly focus on them with only minor information regarding the last, fourth study. Full information regarding the last study will be presented in its own chapter (Chapter 6). The current chapter provides a general overview of participants, virtual environments, study materials, procedures and data analyses. Details of experimental methods that deviate from the ones described here will be provided in the relevant experimental chapters.

All four studies in the thesis involved using HMD-VR system to present some or all of the experimental virtual environments. All of the studies in the thesis followed a similar pattern of exploring or performing tasks in three different environments which were followed by memory tests. The environments in Experiments 1 and 2 were presented just through HMD-VR while in Experiment 3 half of the participants experienced the three environments through the HMD-VR and the other half through Desktop-VR. Experiment 4 differed significantly as one of the environments was presented through HMD-VR, one through Screen VR, and one was created in real-life.

2.1. Participants

The Research Ethics Committee of Bishop Grosseteste University approved all experiments individually in this thesis. Participants were a mix of students from the university and the general public. All participants had a normal or corrected-to-normal vision and aged 18-51. All participants were screened using The Motion Sickness Susceptibility Questionnaire (Golding, 1998), the Epilepsy Screening Questionnaire (Placencia et al., 1992). The data were used for screening purposes and did not form a part of the main analyses. Participants who scored more than 20 points on the Motion Sickness Susceptibility Questionnaire or answered positively to the majority of questions on the Epilepsy Screening Questionnaire were not allowed to continue with the studies. This was not encountered in any of the four experiments. All participants in Experiments 2 and 4 were reimbursed using Amazon vouchers (£20 and £10 respectively). Undergraduate psychology students also received course credits.

Participants who scored more than 20 points on the Motion Sickness Susceptibility Questionnaire, more than 7 points on the Insomnia Severity Index or more than 5 points on the Pittsburgh Sleep Quality Index global score were not allowed to continue with the experiment.

2.2. Equipment

HMD-VR VEs (used in Experiments 1, 2 and 4) were presented using the HTC Vive virtual reality system (HTC Corporation, 2015) connected a desktop computer. The HMD contains two 1080×1200 px resolution OLED screens with a 90 Hz refresh rate and a 100° (horizontally) × 110° (vertically) field of view. Additionally, it has multiple infrared sensors, an accelerometer, and a gyroscope. The HMD-VR system comes with two wireless controllers to represent hands in the VEs. Only Experiment 4 utilised one of the controllers as none of the other experiments needed that functionality. The location and rotation of the HMD and controllers were tracked by the Lighthouse system containing two infrared sensors positioned in two opposite corners of the room. The VR sensor positioning allowed participants to physically explore a 300cm x 250cm area in the middle of the laboratory. Everything that participants saw through the HMD was also visible to the experimenter on the computer screen. The Desktop-VR VEs (Experiments 3 and 4) were shown on the same desktop computer.

Participants in Experiment 1, 2 and 3 had to wear stereo headphones (Sony MDR-XB950BT). In Experiments 1 and 2, the headphones were connected to the headset, whereas in Experiment 3, they were connected to the desktop computer. Headphones were not used in Experiment 4, as there were no sounds.

2.3. The Environments

The Virtual Environments (VEs) used in the thesis were created using either Unity3d (Experiments 1, 2 and 3) or Unreal (Experiment 4) game engines. The 3d models used in the environments were obtained from internet websites such as Turbosquid and CGTraders, the game engine marketplaces or video games such as Counter-Strike: Global Offensive. The sounds used in Experiments 1, 2 and 3 were obtained from various internet websites and were all free to use.

All VR environments were designed around the 300cm x 250cm walkable area, trying to keep it clear of any virtual objects. This was done to provide the participants with as much physical walkable space as possible. Almost all of the objects that participants saw in the VEs were outside of the walkable area but still in reach for the participants. Experiments 1, 2 and 3 used identical VEs consisting of one practice and three experimental rooms. Experiment 4 differed significantly by having one real-life and two virtual experimental rooms. The real-life environment was created in one of the Bishop Grosseteste University's rooms and used solely for that experiment. The practice room in Experiment 4 was for the two virtual experimental rooms.

2.3.1. The virtual practice environments

The virtual practice rooms were created to make participants comfortable with the VR headset and general VE exploration. Participants were taught how to stay inside the walkable area and not to walk into any real objects in the laboratory. In the cases where additional controls were needed for the experimental rooms (Desktop-VR group in Experiment 3 and both VR conditions in Experiment 4) participants were also taught the keyboard and mouse controls and how to use the HMD-VR controllers.

2.3.1.1. Experiments 1, 2 and 3

The practice room used in Experiments 1, 2 and 3 involved participants exploring a simple room made out of colourful shapes (Figure 2.1). Two of those shapes, one on each side of the room, played animations when participants got close to them: a square fell down the wall, and a cube spun around its axis. Both animations contained sounds: square hiding a surface and the cube creaking as it spun. When both animations were triggered and the participant indicated that they feel comfortable in the VE, the headset was removed and the experimental phase commenced.

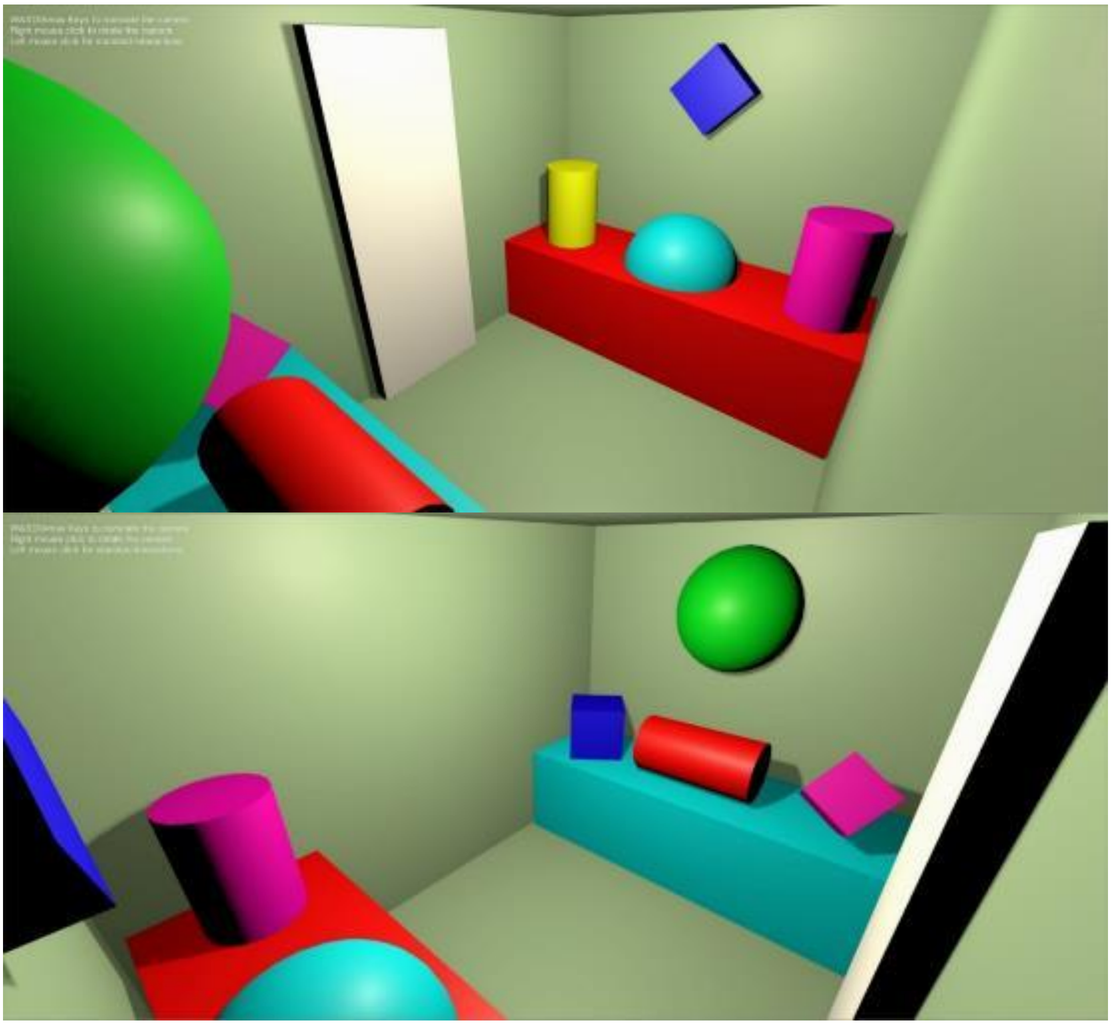
2.3.1.2. Experiment 4

The practice room used in Experiment 4, involved participants hiding four objects in four corners the room (Figure 2.2). A platform was located in the middle of the room, on which participants were presented with three objects. Participants were asked to hide two of those objects in two corners of the room. After hiding the second object, a new group of objects replaced was placed on the platform. Participants again had to hide two of the three objects in the two remaining corners of the room. The objects and the order in which they had to be hidden were provided differently depending on the VR condition. In the HMD-VR condition, a button press on the VR controller brought up a virtual plane, attached to the controller. On that plane, participants were shown the two objects they had to hide and numbers indicating in which order. In the Desktop-VR condition, an experimenter sitting next to the participant held a piece of paper with the same objects and numbers as in the HMD-VR condition. When hiding the second group of objects, the HMD-VR plane was instantly replaced with one containing the needed objects. Similarly, experimenter switched to a different piece of paper in the Desktop-VR condition.

The same room and the same task was used for both HMD-VR and Desktop-VR conditions. The task ensured that the participants were comfortable with the VR headset and that they knew all the needed controls. In the HMD-VR condition, the needed controls included physical movement in the VE, picking up and dropping an object and turning on and off the object order plane. In the Desktop-VR conditions, the controls included walking around the VE using the keyboard, using the mouse to look around, using the mouse to pick up and drop objects, using the mouse to make the object that is being held closer or further away from the screen and how to ‘crouch’ to reach locations closer to the ground. In this experiment participants also engaged with these tasks in a comparable real room, as described later.

Figure 2.1.

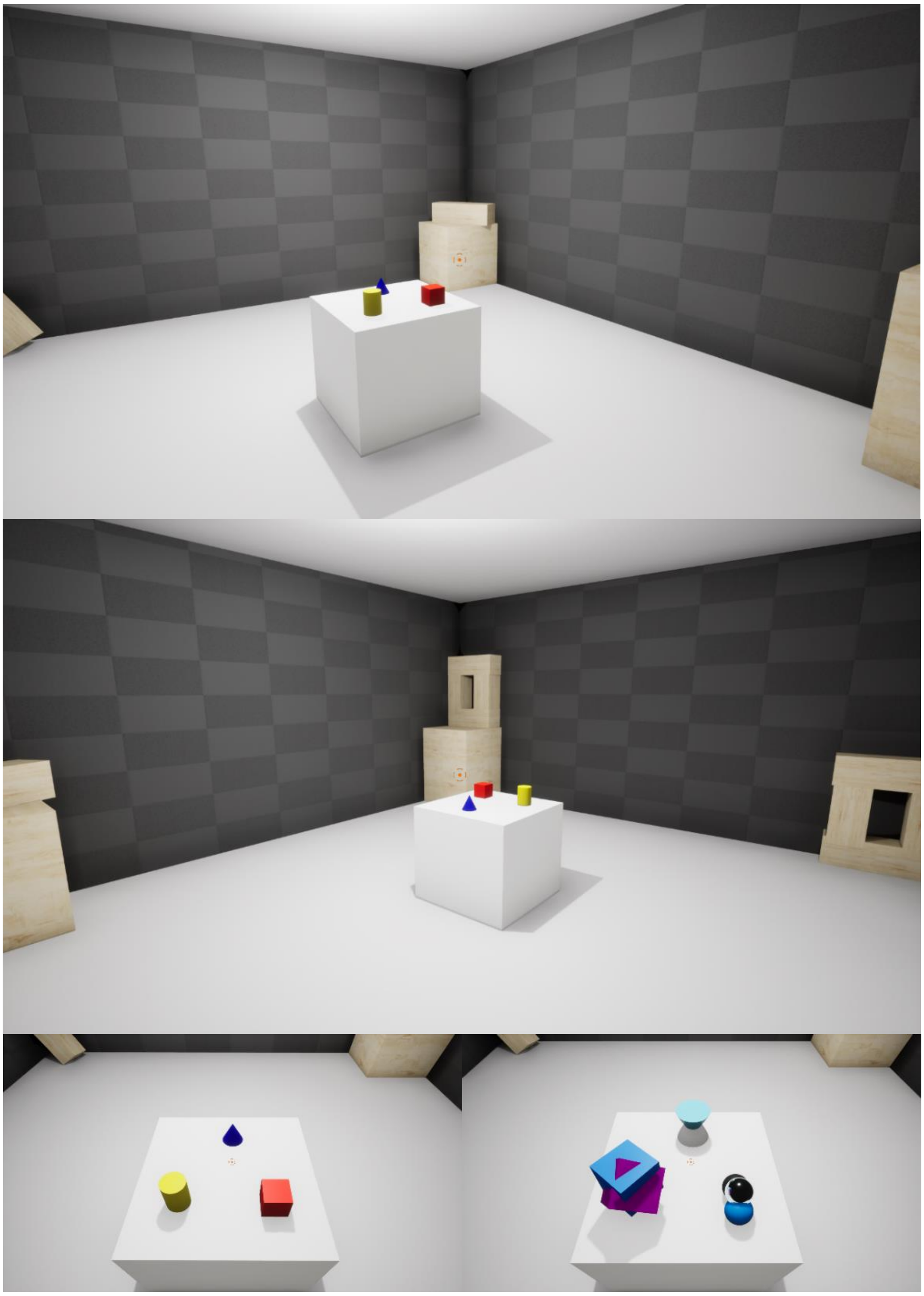
The virtual practice room used in Experiments 1, 2 and 3.



Note. To prepare the participants for the experimental rooms, two objects in the room played animations when participants got closer to them: the blue square fell off the wall (top), and the pink cube spun around its axis (bottom)

Figure 2.2.

The virtual practice room used in Experiment 4.



Note. Participants were presented with a group of three objects in the middle of the room (bottom left) and had to hide two of them in two corners of the room. After doing so, another group of objects was presented in the same place as the last one (bottom right). Participants again had to hide two objects in the remaining two corners of the room. The same room and the same task was used for both HMD-VR and Desktop-VR conditions.

2.3.2. Experimental environments

2.3.2.1. Experiments 1, 2 and 3

As mentioned before, Experiments 1, 2 and 3 used identical VEs. The VE's were comprised of three virtual rooms decorated to look like a bedroom, a kitchen and a study (Figure 2.3). The availability of 3d models chose the way the rooms were decorated. As mentioned before, all of the rooms were modelled in a way to leave the middle of the rooms empty so participants could walk around freely.

In addition to static furniture objects such as desks, tables and wall units, each room featured six event and six non-event objects. The reasoning behind having event and non-event objects is explained in Chapter 3 but a brief explanation would be to explore the differences between closer to real-life experiences (event objects) and static objects commonly used in memory research (non-event objects). Event objects constituted as objects which played certain animations (see Figure 2.4). This ranged from the clock falling off the wall and the books tumbling on their sides to the phone ringing and the microwave turning itself on and off (for a list of all the event objects and their events see Appendix B). These events were triggered by walking over invisible triggers placed close to the event objects and looking at invisible triggers surrounding those objects (see Figure 2.5). This was done to stop the animations playing by accident when participants walked past the objects. All events lasted no more than two seconds and all of them had accompanying sounds (e.g. the clock hitting a surface, sound of the microwave working etc.). Non-event objects were static objects that did not play any animations. Both event and non-event objects were chosen to be related to the room they were in (e.g. chopping board in the kitchen, PC in the study, TV in the bedroom etc.). In total there were thirty-six testable objects present in all three rooms (6 x 3 event and 6 x 3 non-event objects).

Figure 2.3.

The three virtual environments used in Experiments 1, 2 and 3.

Note. Top - Bedroom, middle - Study, bottom - Kitchen. Note the empty space in the middle of each room. This space represents the physical walkable space created with the HTC Vive VR system.

Figure 2.4.

An example event experienced by the participants in Experiments 1, 2 and 3.

Note. After looking at it, the painting moves to a side as if one of the nails holding it came off. This was accompanied by a sound.

Figure 2.5.

An example of the triggers used in Experiments 1, 2 and 3.

Note. The triggers are made visible for this example. The event (popping of the lightbulb) plays when a participant walks on the trigger on the floor and looks at the trigger surrounding the object (in this example – the lamp).

2.3.2.2. Experiment 4

Experiment 4 involved participants performing tasks in one real and two virtual environments. For the real-life environment, a 250cm by 300cm room was used in the Bishop Grosseteste University (Figure 2.6). The two virtual environments were modelled to look like two distinct realistically furnished rooms. Unlike in the previous experiments, the walkable area in the environments was not completely empty. In every environment (as in the corresponding virtual practice environment) there was a small table with a number of objects placed on it. The table was small enough not to inconvenience the participants and positioned in a way to provide enough space to walk around. The table was used to provide participants with two groups of six objects at a time. The task used in this experiment involved participants hiding four objects from each of the two groups in each of the environment. This resulted in every participant hiding twenty-four objects in three environments. The reasoning behind the task will be provided in Chapter 6.

Real-life (Figure 2.6). A box in the middle of the room was used as a table on which objects were placed, six at a time. The objects the participants had to hide and the order in which they had to be hidden was provided on two pieces of paper (one for each group) given one at a time. The experimenter showed the locations in which the objects had to be hidden by silently pointing towards them.

Figure 2.6.

The Real-world environment used in Experiment 4.

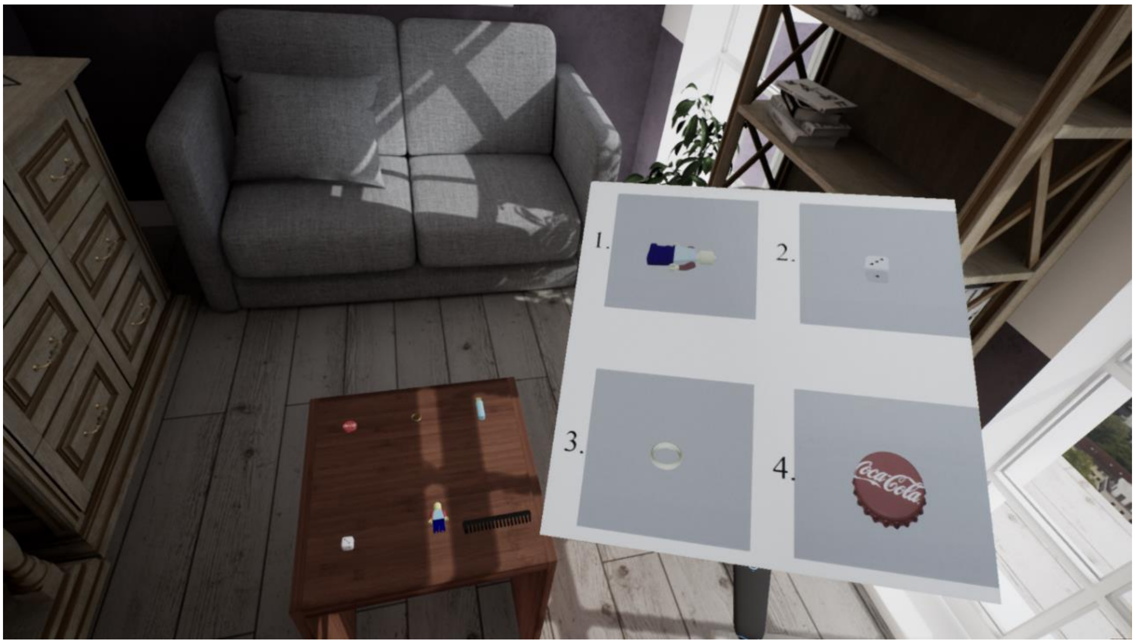
HMD-VR (Figure 2.7). To pick up and hide the objects, participants had to use the VR hand controllers. The order in which the objects had to be hidden was presented on a virtual plane that appeared over the controller when a participant pressed a specific button (Figure 2.8). The hiding worked by picking up one of the objects with the controller and placing it near one of the hiding spots shown by an arrow (Figure 2.9). The second group of objects was instantly placed on the object table as soon as the last object of the first group was hidden while also clearing the table of the two distractor objects. This also updated the virtual plane with a new set of objects and their order of hiding.

Figure 2.7.

The HMD-VR environment used in Experiment 4.

Figure 2.8.

The virtual paper used in Experiment 4.



Note. The virtual paper contained the order in which participants had to hide the objects. The virtual paper could be turned on and off with a press of a button on the hand controller.

Figure 2.9.

Hiding of an object in the HMD-VR environment in Experiment 4.

Note. As soon as an object is picked up from the table, an arrow appeared near the spot where the object needed to be placed. When the object reached the hiding spot, it disappeared along with the arrow.

Desktop-VR (Figure 2.10). The navigation and interaction with the objects were done using a mouse and a keyboard. The left mouse button was used to pick up and drop objects while a mouse scroll was used to bring the object closer or further away from the participant. The right mouse button would allow participants to 'crouch down' to better reach some places. The object hiding was identical to the HMD-VR with the only difference being the object order was shown on a piece of an actual paper as in the Real-world setting. The piece of paper was held by the experimenter so that the participant could easily see it while still be able to focus on the task.

Figure 2.10.

The Desktop-VR environment used in Experiment 4.

2.4. General procedure

The general procedure in all four experiments was to explore three rooms and perform memory tasks concerning some of the items found in those rooms. This section will provide a general overview of the procedures used in the thesis. The more specific information on the procedures is included in each of the corresponding experimental chapters.

2.4.1.1. Experiments 1, 2 and 3

Before entering the main experimental VEs, all participants had to explore a practice VE. This was done to check if the VR headset was not inducing nausea and if the participants were feeling comfortable wearing it. Participants spent 3 minutes in the practice environment. Participants were then assigned a sequence of experimental rooms to explore. This was done for counterbalancing purposes. There were a total of six different sequences (e.g. ABC, ACB, BAC etc.).

Before starting each experimental VE, participants were asked to stand on a predesigned location in the laboratory and face the same wall. This led to participants ‘entering’ each VE facing a virtual door and not the whole room. This was done to stop participants from being instantly ‘overwhelmed’ by the whole VE.

Participants were tested individually. The instruction was to explore the rooms while trying to inspect all of the objects. Unknowingly to them, participants were given three minutes to explore each room. After the three minutes, participants were asked to walk back to the virtual door which triggered a fade-out sequence and the removal of the VR mask and headphones by the experimenter. The point of this exploration was for the participants to trigger and observe all of the events in each of the room (six per room). If after three minutes there were still some events left un-triggered, participants were given an additional minute and again encouraged to explore the room and to inspect all of the objects. If after four minutes in total there were still events left un-triggered, the exploration was stopped and the normal finishing procedure ensued. This was repeated for all three rooms with two-minute breaks after each one, during which participants had to fill in the Simulation Sickness Questionnaire (Kennedy et al., 1993).

After exploring the last room and filling the last Simulation Sickness Questionnaire, participants sat in front of a computer and underwent the memory tests – the free recall test, the What-Where-When episodic memory (EM) test and the object recognition test. All three experiments also had additional testing sessions at various intervals during which participants only performed the memory tests. Experiments 1 and 3 in total had two testing sessions while Experiment 2 had four. The timing of these sessions and the rationale behind them will be discussed in each of the corresponding experimental chapters.

2.4.1.2. Experiment 4

All participants were met in a room separate to the ones used for the Real-world setting or the one with the VR/Screen equipment. Participants were told that they will be hiding objects in three different rooms and then filling a number of questionnaires afterwards. The order of the rooms was decided before the participants came in, in an identical fashion to the previous experiments (e.g. ABC, ACB, BAC etc.) The order of objects and the hiding locations were the same for all of the participants.

The objects that had to be hidden were presented in the middle of each room, four at a time with two additional distractor objects mixed in (for the list of all the objects see Appendix G. After hiding the four objects, another group of objects were presented resulting in eight objects from two groups hidden in each of the three rooms. This created the needed What-Where-When information: What – The object, Where – The hiding location, When – Was the object from the first of the second group of objects. In the VR settings, the change of objects was instant. In the Real-life setting, participants were asked to leave the room and wait outside. During this time experimenter removed the four hidden objects and the two distractor objects. The second group of objects was then placed on the box and the participant was invited back in and given a new piece of paper with the objects needed to be hidden. This took less than a minute.

After finishing each environment participants had a two-minute break. After the HMD-VR and Desktop-VR environments, participants sat down in the same laboratory while after the Real-world environment participants had to walk to the laboratory located in a different building.

After finishing all three environments, participants were taken to the room they were initially greeted in and were asked to fill in the IPQ (Igroup Project Consortium, 2015) and PQ (Witmer & Singer, 1998) questionnaires. Following this, participants were told that they can leave the laboratory and to come back to the same room after one hour. After participants came back they were asked to do the WWW, object recognition and detail tests. All of the tests were done on a computer in the room where the participants were initially met.

2.5. The Memory tasks

Experiments 1, 2 and 3 contained almost identical memory tests with only minor modifications. The order in which the tests were completed was: free recall, WWW and object recognition. As some memory tests in Experiment 4 differed significantly only the common features of the tests used in Experiments 1, 2 and 3 will be provided here. Information regarding the memory tests used in Experiment 4 will be provided in its own chapter (Chapter 6).

2.5.1. The free recall task

This task was only used in Experiment 1, 2 and 3. In the free recall task, participants were given space to write freely about what they had experienced in the main VEs. The instructions were the same in all three of the experiments:

"Imagine that you are telling a friend about the three rooms you have just explored. Try to write down everything you have seen and experienced in those rooms. Try to give as many details as possible."

Participants then had unlimited time to write down their experiences.

2.5.2. The What-Where-When task

The task was used in all four of the experiments. However, while the task was almost identical in Experiments 1, 2 and 3, it was greatly modified in Experiment 4.

Due to this, the present section will only focus on the task used in the first three experiments.

The WWW task was split into five parts: What, When, Where, Event and Detail and presented in that particular order. Instructions for each part were provided on screen. All parts of the task followed a similar pattern – participants were given a name of an object and had to answer the related question (screens of each part of the task can be seen in Appendix D and E).

2.5.2.1. What

Participants were presented with the name of an object and asked to indicate if they recalled seeing that object in the VEs. The pool of objects used in this task included eighteen event (six per room), eighteen non-event (six per room) objects and thirty-six distractor objects (objects that were not present in any of the rooms). This resulted in a pool of 72 objects. The main sampling feature of the distractor objects was that it would be possible to find them in the three different environments (bedroom, study and kitchen) used in the experiments. For example, an iron, a clipboard and a fork (some of the distractor objects) could all be associated with the bedroom, study and kitchen rooms accordingly but were not present in any of the rooms.

In experiments that only had two testing sessions (Experiments 1 and 3), the object pool was equally divided across the sessions resulting in 9 event, 9 non-event and 18 distractor objects. In Experiment 2, the first two testing sessions followed identical object pool division as in Experiments 1 and 3, however during the remaining two testing sessions participants were presented with all 72 objects. The objects that were not recalled during this task were not used in the remaining parts of the task.

2.5.2.2. When

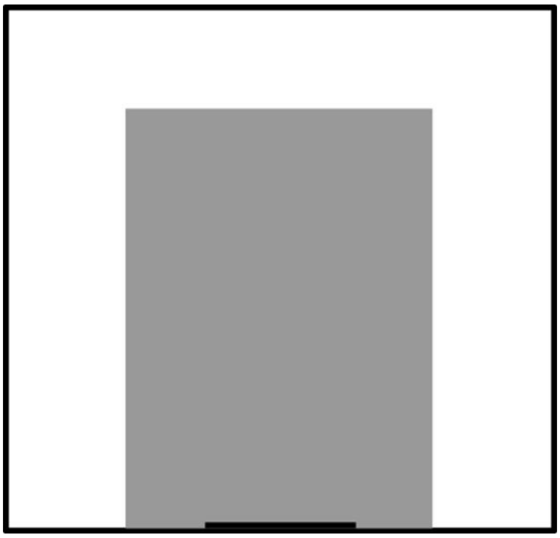
Participants were asked to indicate in which of the three virtual rooms they recalled seeing that specific object: first room, second room or a third room.

2.5.2.3. Where

Participants were given a simple map representing all three rooms (Figure 2.11) and were asked to use the mouse to press where they thought that the object was located in the room.

Figure 2.11.

A map of the rooms used in the Where part of the WWW task in Experiments 1, 2 and 3.



Note. The grey area is the area in which participants were able to walk around. The white area is the area in which all of the room objects were located. The thick black line on the bottom is the location of a virtual door which all of the participants faced before starting the exploration.

2.5.2.4. Event Recall

Participants had to indicate if anything happened to that object. If participants indicated that something happened, they had to write what they thought happened to that object (e.g. “the clock fell off a wall”). A list of all the event objects and their events can be seen in Appendix B.

2.5.2.5. Detail Recall

Participants were asked if they remembered any physical or perceptual details about the objects. If they did, similarly to the Event part, participants had to write that detail down (e.g. “The clock showed ten to one”). Participants were given a chance to provide five details for each object.

2.5.3. The object recognition task

In the object recognition task, participants were presented with images of objects and had to indicate if they recognised them, and how confident they were with their decision. The pool of objects used in this task was the same as in the What part of the WWW task. In the images, objects were placed against a plain grey background. The confidence was assessed by asking participants to use a continuous scale with the leftmost point being “I am not confident” and the rightmost point being - “I am very confident”. Screens of each part of the task can be seen in Appendix D and E.

2.6. Memory retention period

All four of the experiments explored memory retention and consolidation. This ranged from 1h (Experiment 4), 24h (Experiment 1 and 3) or 30d (Experiment 2). In each case, participants were asked to leave the laboratory and continue with their daily activities. In Experiments 1, 2 and 3 memory testing occurred directly after the VE exploration with additional testing sessions: after 24h in Experiments 1 and 3 or after 24h, 7d and 30d in Experiment 2. In Experiment 4, the one and the only testing session was after the 1h retention period.

2.7. Data processing and analysis

Data were analysed using JASP (JASP Team, 2019) and R (R Core Team, 2019) statistics software packages. An alpha level of .05 was used for all statistical tests. For post hoc analyses, p values were corrected using the Bonferroni method.

The free recall task

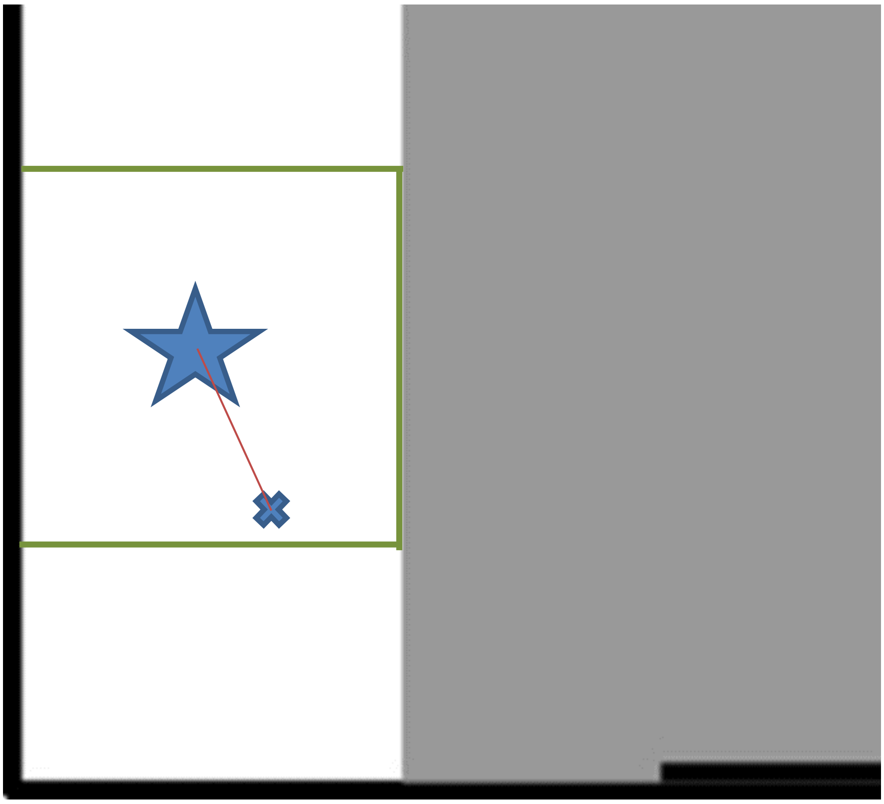
Free recall responses were scored in terms of mentions of objects and object details. They were divided into event, non-event objects and their details. This provided a dataset showing how many objects and how many object details each participant recalled.

The WWW task

The WWW data were converted to a proportion of correctly recalled information on the specific testing session. Objects that were not present in the VEs (distractors) were not included in the analyses. The Where and When proportions were calculated from the number of recalled What objects. For example, if a participant recalled 6 event objects out of 9 possible in the session, the proportions were calculated using the number of recalled objects and not the number of all possible objects. The Where proportion was created by checking if object's location guess was no further away from the real object location than the distance from the virtual wall to the physical walkable area (for a visual explanation see Figure 2.12). In addition to the separate WWW component proportions, a unified WWW proportion was also calculated. This was a proportion of recalled What objects that also had correctly recalled Where and Where components.

Figure 2.12.

An example of how Where binary score was calculated in Experiments 1, 2 and 3.



Note. In this example, the star is the object that the participant is pointing to. The X is the point where the participant thinks the object was in the room. The red line between the real object and the guessed location of that object is the Where pointing error in screen pixels. The green lines indicate the square area around the object. If a participant's guess falls within the area – it is counted as if the participant correctly guessed the object's location.

Object recognition task

Recognition data was converted to d' scores. The d' score relates the correct positive judgments or hits, to the false-positive judgments - d' = z(hits) – z (false positive) (Haatveit et al., 2010; Swets et al., 1961).

Chapter 3 – Experiment 1: The effect of sleep on event and non-event based episodic memory

3.1. Introduction

One of the weaknesses of episodic memory (EM) research is that the stimuli used in studies often lack ecological validity (Burgess et al., 2006; Parsons et al., 2017; Parsons & Rizzo, 2008b). In addition, studies exploring EM tend to test participants' memory straight after learning which does not represent real-life behaviour as the recall of information might not be necessary until the next day, or later. As it has been shown, sleep promotes memory consolidation (Eichenbaum & Cohen, 2004; Rauchs et al., 2011; Walker, 2009), this leads to a question of how is EM is affected by a period of sleep. The present chapter has explored how EM for life-like events is affected by sleep-dependant memory consolidation.

The What-Where-When (WWW) information is considered a cornerstone of EM and is frequently used in EM research (Martin-Ordas et al., 2017; Pause et al., 2013; Plancher et al., 2008). However, while the WWW test is usually the main approach to explore EM, there are a number of other approaches such as free recall or object recognition tasks each having certain benefits that the other approaches do not (Cheke & Clayton, 2013, 2015). As discussed in Chapter 1 (section 1.4.7), all three approaches make useful contributions to better understand EM. However, as discussed in Chapter 1 (section 1.5), there is a general problem of low ecological validity of laboratory-based EM experiments (Burgess et al., 2006). To overcome this problem, and to explore EM in more naturalistic settings, researchers started utilising virtual reality (VR) to create realistic virtual environments (VEs). It is argued in the literature (Bréchet et al., 2019; Dehn et al., 2018; Li et al., 2016; Picard et al., 2017; Reggente et al., 2018) and as the main point of this thesis that VR use should lead to more life-like memory representations and thus more ecologically valid data.

VR has been used to assess object memory (Parsons & Rizzo, 2008b; Sauzéon et al., 2012; Widmann et al., 2012) and object memory in association with contextual information such as the character, the location and the moment associated with each object (i.e., WWW information) (Plancher et al., 2010, 2012; Rauchs et al., 2008). Some relevant examples of VR use in EM research are the studies by Plancher and colleagues (Plancher et al., 2010, 2013, 2008). In their studies, participants had to drive through a virtual city containing a number of scenes composed out of one main (e.g. a newsstand) and some secondary (e.g. a man or a bench) elements. Afterwards, participants were given a free recall test during which they were asked to write down all of the elements (what) they saw in the VE with all of the corresponding perceptual details (detail, where, when). For example – “at the beginning of the route I saw a newsstand with a man sitting next to it. The man was wearing a red shirt”. A similar protocol was also used in a recent study by Picard, Abram, Orriols and Piolino (2017). In this study, participants had to walk through a virtual town to visit a friend and memorise as many elements as possible. During the testing phase, participants were asked to freely recall as many factual items and the associated spatial, temporal and perceptual details as possible. In both these cases, results from the VR were compared to more standard clinical tests of EM such as the Cognitive Difficulties Scale (McNair et al., 1983) or Family Pictures test (Horton et al., 2001). The comparison was done as part of the proposal to use VR as a new ecological tool to assess EM. It was pointed out that VR was a useful and appropriate tool to test EM as it was shown to be more sensitive to memory complaints of daily life compared to the more traditional pen-and-paper tests.

However, in all the mentioned cases, VEs were viewed and explored through Desktop-VR which still lacks the real-world immersion. Participants were focusing at the screen in front of them which has been shown to reduce the immersion aspect of the screen-based exploration compared to the real-world experiences (Kinugawa et al., 2013; Zlomuzica et al., 2016). Recently, studies started to emerge, utilising the fully immersive head-mounted display based VR (HMD-VR). A study by Davison, Deeprose and Terbeck (2018) investigated the use of HMD-VR in the assessment of age-related cognitive functions. Participants completed Stroop colour-word and trail-making tests as traditional assessments of executive function and three tasks in HMD-VR: a seating arrangement task, item location task and virtual parking simulator. In the parking task, participants had to navigate themselves into parking spaces. The arrangement and item location tasks were set in a virtual chemistry lab. During the arrangement task, participants had to create a seating plan by rearranging chairs while in the object location task participants had to find various items located in the lab. The study showed that HMD-VR measures were better contributors in predicting age-related cognitive decline than traditional neuropsychological tasks. The explanation for this finding was that the HMD-VR tasks were more sensitive and ecologically valid assessments of everyday cognitive functions and normal ageing. A more EM focused study was conducted by Ouellet et al. (2018), in which EM was tested using an HMD-VR based Virtual Shop task. The task consisted of remembering and retrieving twelve objects in an environment representing a grocery shop. As in the previously mentioned study, participants have also completed a traditional memory task. In this case, it was a free recall word list test. The results showed that both construct and ecological validity was supported by the data. The HMD-VR task was sensitive to ageing and was related to an everyday measure of shopping abilities. It was found that the task was better correlated to the participants' memory ability than the traditional memory task (Weschler Memory Scale).

As it can be seen, HMD-VR has a positive contribution to ecological validity in memory testing. Due to this, the present experiment used HMD-VR to continue with the earlier discussed (Chapter 1, section 1.8) "function-led" EM testing (Parsons, 2015; Parsons et al., 2017). As EM tests mostly rely on static stimuli such as words or pictures of objects, which is not how we experience the world, it can be argued that more valid measures are required. Even in the earlier mentioned study by Plancher et al. (2008) participants observed static scenes such as a train station with a girl in front of it. The study by Ouellet et al. (2018) overcomes this problem by first of all utilising HMD-VR and then using a task that puts participants in a realistic environment with a task that reflects a real-life behaviour. The present experiment tried to follow this direction and move towards the use of more real-like events and experiences in EM testing. To further explore the use of HMD-VR use in EM research and its capabilities to produce life-like experiences, the present experiment introduced event objects.

In addition to a set of static objects such as the man in a red shirt from the Plancher et al. (2008) study, objects were added that 'performed' an event (for example books falling on their side or a TV turning itself on and off). The reasoning behind this was the research showing that more life-like events and experiences are more likely to be remembered compared to laboratory-based stimuli such as lists of words or pictures (Chen et al., 2017; Roediger & McDermott, 2013; Schöne et al., 2019). However, this leads to the question of whether remembering a word seen on a computer screen and remembering that some books fell off a shelf are equal EMs? In both instances, it is possible to recall the needed WWW and any other perceptual information, and in both instances, their recall can potentially be identified as remembering (associated with EM and not 'knowing' associated with semantic memory). The difference between these two approaches is what has been mentioned earlier and discussed in Chapter 1 (section 1.3) – the representation of real-life behaviour and events. The data collection methods need to be as close to real-life tasks as possible and the measures need to reflect and predict real-world tasks (Chaytor & Schmitter-Edgecombe, 2003; Ready et al., 2001; Silver, 2000). Following this logic, the event objects introduced in this experiment, better represent the What happened part of the WWW triad. However, it is not to say that objects with no events attached to them (here called non-event objects, in comparison), as in most of the aforementioned studies, cannot be recalled using EM. The problem is how these objects are experienced and how these experiences reflect the real-life. Real-life experiences are rarely about observing static objects and more about experiencing events. As such, the present experiment aimed to explore this difference between the two types of stimuli.

Events are more episodic and should lead to more holistic activation of the WWW EM components. This activation should lead to a richer memory trace being encoded. The support for this comes from memory literature showing that stimuli with higher saliency and novelty are better recalled (Fernández & Morris, 2018; Hunt & Mcdaniel, 1993; Neath, 1993a, 1993b; Reggev et al., 2018; Schmidt, 1991; Van Kesteren et al., 2012). It is argued that an event or an object that shows lack of typicality among other events or objects (in the present case event object among non-event objects) will be 'tagged' for preferential consolidation (Fernández & Morris, 2018). As a result of this research, the present experiment used two types of objects – event and non-event. The inclusion of the event objects served two purposes. First of all, to explore the differences in EM recall between event and non-event objects. The prediction was that event objects would be better remembered than the non-event objects. This should be visible in free recall, WWW and recognition tasks. In free recall and object recognition tests, this difference should be evident by a higher number of recalled event objects and higher recognition scores for the event objects. The recognition scores included d' sensitivity index and confidence ratings with the d' score providing a measure of memory sensitivity and the confidence ratings showing the 'ease' of the recognition response. Both measures have been widely used in EM research (Dewhurst et al., 2009; Mickes et al., 2013; Weidemann & Kahana, 2016; Wichert et al., 2013).

A similar difference was also predicted in the combined WWW scores – correctly recalling What, When and Where information about an object. What was difficult to predict was the difference between the two object types when looking at the separate WWW components. While these differences were not the main interest of the experiment, the majority of EM studies employing the WWW task also provide the results for the separate components (Cheke, 2016; Holland & Smulders, 2011; Plancher et al., 2012, 2013; Saive et al., 2015). Due to lack of research using similar methods to the one used here, the prediction was twofold. First of all, the recall of separate WWW components should follow the combined WWW prediction and be higher for the event objects. This prediction was based mainly on the earlier discussed literature regarding object saliency. Due to this effect, it was predicted that event the separate WWW components should be better recalled for the event objects than the non-event objects. Secondly, research shows that item memory (what) should be better remembered over temporal (when) or spatial (where) information (Dobbins et al., 2002; Fujii et al., 2004; Hayes et al., 2004). The difference in recall between spatial and temporal information is less clear with some studies showing higher spatial recall (Hayes et al., 2004; Postma et al., 2006) and some showing similar levels of recall between the two types of memory (Fujii et al., 2004; Pitel et al., 2007). Due to this, we predicted that item information (What recalls in our case) would be better recalled than temporal (When) and spatial Where) information.

An additional avenue of exploration added in the present experiment was the memory for perceptual details. In studies by Plancher and colleagues (Plancher et al., 2010, 2013, 2008), in addition to the main WWW components participants were also asked to recall any perceptual details about the elements that were part of the scenes (e.g. the girl in front of the train station wore a red shirt). While not being the main point of this experiment, it was worthy to explore how memories for perceptual details would differ between the event and non-event objects. As discussed by Conway (2001), episodic information is conceived as being largely sensory-perceptual in nature. In a method akin to Plancher et al. (2008), in the free recall and WWW tasks participants were also asked to recall as many perceptual details about the objects. As with the separate WWW information, the prediction was that there would be more perceptual details recalled for the event objects as compared to the non-event objects. This prediction was based on the previous prediction that event objects will be better recalled and also on the earlier discussed notion that event objects should have richer memory traces.

As discussed in the literature review, EM is often explored while only looking at encoding and retrieval of information. The consolidation phase is often overlooked, and memory is tested immediately after encoding (e.g. earlier mentioned studies Ouellet et al., 2018; Picard et al., 2017; Plancher et al., 2008). While this is not a problem in itself, we argue that the move to increase the ecological validity of EM testing should have a consolidation period as it is rare to recall experiences immediately after they happen. Consolidation and forgetting literature shows that memories do change over time (Diekelmann & Born, 2010; Rasch & Born, 2007) and while it is clear that, for example, a retention period of a day will have a great effect on retained information even a shorter period will have an effect on memory retention (Martini et al., 2019). What is lacking in the EM literature is research using ecologically valid measures HMD-VR how EM is retained over time while still having immediate memory for comparison. Here for this purpose, HMD-VR was employed.

Sleep has been shown to be pivotal in memory consolidation processes (Stickgold & Walker, 2005). Temporal order in EMs (Griessenberger et al., 2012), prospective (Grundgeiger et al., 2014), implicit (Casey et al., 2016), emotional (Nishida et al., 2009) and spatial memory (Guan et al., 2004) all benefit from a period of sleep after encoding (for reviews see Diekelmann & Born, 2010; Rasch & Born, 2013; Stickgold, 2005). Evidence shows a more active role of sleep in memory consolidation than just passive protection from interference (Diekelmann & Born, 2010; Gais et al., 2006; Lewis & Durrant, 2011). During sleep, memories are reactivated and the synapses associated with the memory traces are up or down-scaled, which facilitates their consolidation. During reactivation, the synaptic scaling potentiates important and weakens irrelevant memory traces thus extracting their salient features (Diekelmann & Born, 2010; Genzel, Kroes, Dresler, & Battaglia, 2014). With consolidation, EMs are redistributed to the knowledge networks leading to the loss of episodic detail and forgetting. This is related to the hippocampal-neocortex memory redistribution as discussed in Chapter 1 (section 1.7.1). While all of the discussed theories that try to explain memory redistribution state that hippocampus is important for the initial stages of memory consolidation, there is a debate on its further involvement. The standard consolidation theory suggests that memories become hippocampus independent. However, the memory trace theories, which the present thesis argues underpins memory consolidation, suggest that hippocampus always remains involved in EM retrieval.

Sleep dependant EM consolidation studies have shown that sleep helps to consolidate EMs in particular. Studies comparing EM performance after a period of time filled with sleep or wake have shown that sleep actively helps to consolidate episodic information (Aly & Moscovitch, 2010; Oyanedel et al., 2019; Rauchs et al., 2004; van der Helm et al., 2011). For example, a study by Aly & Moscovitch (2010) explored EM performance for stories and personal events after a retention interval that included sleep and after an equal duration of wakefulness. Participants were tested three times with testing sessions separated by 12-hour intervals and the first testing session being in the early morning or late evening. The EM tests were the two Wechsler Memory Scale III stories and personal EMs for conversations the participants' had 12h ago. The results showed that participants recalled more episodic information (story units recalled) following sleep than wake (e.g. the 12h period being filled with sleep or wake).

The time the sleep takes place after learning also has an effect on memory formation as shown by studies which compared the effect of sleep just after learning to sleep at a later time (Benson & Feinberg, 1977; Payne et al., 2008; Talamini et al., 2008). It was found that the closer sleep is to learning the better memory retention becomes (Benson & Feinberg, 1977; Payne et al., 2008; Talamini et al., 2008). For example, Payne et al. (2012) explored the effect of time of sleep on the relation of word pairs. After 24h with all subjects receiving both a full night's sleep and a full day of wakefulness, it was found that memory performance was a lot better when learning was followed by sleep rather than by wakefulness. A study by Scullin (2014) used a similar design to also see how a period filled with wake or sleep would affect the recall of word pairs. There were three groups – 12h sleep (testing in the evening), 12h wake (testing in the morning) and 24h (mixed). The data showed that 12h sleep group performed better than the 12h wake or the 24h group but the 24h group performed better than the 12h wake group. This again shows the importance of time of sleep after learning.

Taking the discussed effect of sleep on EM and the general thesis aim of increasing ecological validity into consideration, the present experiment explored the effect of time of sleep and the general effect of sleep-dependant consolidation on EM for the event and non-event objects. In a similar manner to Aly & Moscovitch (2010) participants were tested at either morning (9AM) or evening (9PM) and then retested after 24h. This led to two groups of participants – those whose initial encoding was followed by a full day of wake and then sleep (AM group), and those whose encoding was followed by full night's sleep and then a full day of wake (PM group). Based on the discussed literature, the main prediction was that after the 24h participants in the PM (compared to the AM) group should have overall better performance in the EM tests. This should be reflected in a higher number of recalled objects in the free recall task, a higher number of combined WWW recalls and a higher d' sensitivity score.

Continuing with the effects of sleep, a study by Scullin (2014) has shown that time spent in slow-wave sleep (SWS) was positively correlated with episodic recall. Similar results were also shown in a study by Daurat, Terrier, Foret, & Tiberge (2007) in which participants in the SWS rich sleep group performed better at recognition than participants in the rapid-eye-movement (REM) rich sleep. As an additional explorative measure, participants in this experiment wore sleep tracking bracelets throughout the 24 hours which provided sleep data such as time spent in SWS and REM (for a review and usefulness of these bracelets see De Zambotti, Claudatos, Inkelis, Colrain, & Baker, 2015; Saito & Sadoshima, 2016). Based on the mentioned literature, it was predicted that time spent in SWS would positively correlate with the number of objects recalled in the free recall task, combined WWW recalls (Scullin, 2014) and d' recognition scores (Drosopoulos, 2005).

The present experiment was conducted similarly to the earlier mentioned Desktop-VR experiments by Plancher and colleagues (Plancher et al., 2010, 2012, 2008). The present experiment similar design, exploring VE and providing the WWW information, with the main enhancement being the inclusion of events objects (for example a clock falling off the wall). At the time of writing, there were no HMD-VR studies that had controlled episodic events as part of the stimuli. As explained before, the reason for including these events was to see how memory for event objects differs from memory for static non-event objects. This, with the addition of HMD-VR technology, will let us explore EM in closer to the real-life setting while still giving control on what is being observed.

In general, the present study had two main objectives. First, to investigate how EM for event objects might differ to EM for non-event objects. Secondly, to explore the effect of sleep dependant consolidation and the effect of AM/PM testing on the EM. Both of these objectives were enclosed in a general interest on how custom-built VEs, presented through HMD-VR, can be utilised in function-led and more life-like EM testing.

3.2. Method

3.2.1. Participants

Participants in this experiment were 20 students from Bishop Grosseteste University and members of general public (mean age = 23.65; range = 18 - 52; female = 13). Undergraduate participants took part to obtain course credit; everyone else contributed freely. Participants were randomly assigned to the two groups (n = 10): AM (mean age = 25.10; range = 19 - 52; female = 5) or PM (mean age = 22.20; range = 18 - 38; female = 8). There were 15 students and 5 non-students. All participants had normal or corrected to normal vision. The screening procedure is described in Chapter 2 section 1.1.

3.2.2. Materials

The virtual environments were created using the Unity3D game engine and presented using HTC Vive HMD-VR system. A more in-depth description of the equipment and VEs is provided in Chapter 2 section 1.2 and 1.3.

3.2.3. Design

The experiment contained three tests (free recall, WWW and object recognition). Every test was performed at two time points: immediately after VE exploration (Baseline session) and after 24h (24h session). The exact time of testing depended on the participants' group: AM – tests performed at 9AM each day, PM - tests performed at 9PM each day. The Session (Baseline/After 24h) was the within-subject while Group (AM/PM) was the between-subjects independent variables.

There were five dependent variables in the free recall test: the number of recalled event objects, the number of recalled non-event objects, the number of recalled event object details, the number of recalled non-event object details and the total word count of the provided text.

The WWW task resulted in six dependant variables: the What, the When, the Where, the Event and combined WWW proportions and the average number of perceptual details recalled per one recalled (What) object. The Event component was a proportion of correctly recalled event associated with a particular event object.

After each Event and Detail recall participants had to provide a Remember/Know/Guess judgements regarding that information. This resulted in six dependant variables: proportions of Remember, Know and Guess judgements for the Event component and proportions of Remember, Know and Guess judgements for the Detail component.

The object recognition task resulted in two dependant variables: the d' sensitivity index and the confidence rating.

The actigraphy bracelet provided three dependant variables: total time spent asleep, time spent in REM and time spent in SWS.

3.2.4. Procedure

Participants were asked to come to the laboratory at either 9AM or 9PM depending on the group they were asked to be assigned to (AM or PM respectively). The self-assignment was due to a low number of available participants. Participants were naïve to the study aims and were told that the study was about exploring VEs. The general procedure that followed is described in Chapter 2 section 1.4. After the free recall, WWW and object recognition tasks, participants were given Jawbone UP3 trackers which they were asked to wear for the next 24 hours. Participants were then able to leave the laboratory and carry out their normal daily activities. Depending on the group participants were assigned to (Morning or Evening), participants were asked to come back to the lab at either 9AM or 9PM the next day. In the second session of testing (24h condition), participants completed the free recall, WWW and object recognition tasks.

3.3. Results

3.3.1. Free recall

In the free recall task, participants were given space to write freely about what they had experienced in the main VEs. This was done in a form of telling a story to a friend about what the participant experienced in the VEs. Free recall tests were scored in terms of the number of mentions of objects and object details. For example, “I remember seeing a radio next to a red mug and also a grey phone” would be marked as two non-event objects (radio and mug), one event object (phone), one non-event object detail (red mug) and one event object detail (grey phone).

3.3.1.1. The number of recalled objects

The number of recalled event and non-event objects were compared at Baseline and after 24h. An RM ANOVA with Session (Baseline, After 24h) and object type (Event, Non-event) as within-subject variables and Group (AM, PM) as a between-subject variable was performed. The means and standard deviations are presented in Table 3.1.

Table 3.1.

Experiment 1: Mean number of recalled objects in the free recall test

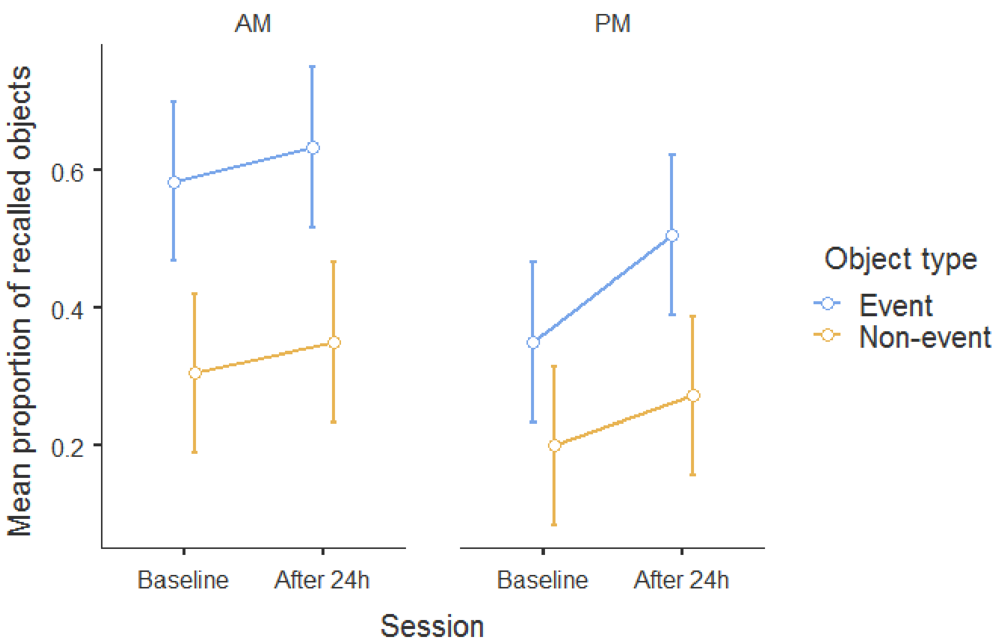
		Object type	
Session	**Group**	Event	Non-event
Baseline	AM	10.50 (3.24)	5.50 (3.47)
	PM	6.30 (3.71)	3.60 (2.67)
After 24h	AM	11.40 (3.03)	6.30 (2.71)
	PM	9.10 (4.15)	4.90 (2.60)

Note. Numbers in parentheses represent standard deviations

An effect of Object type was found, $F(1,18)=54.44$, $p<.001$, with more Event objects being recalled than Non-event objects, $t(18)=7.38$, $p<.001$ (see Figure 3.1). A significant effect of Session was found, $F(1,18)=9.67$, $p=.006$, with object recall at Baseline being lower than After 24h, $t(18)=-3.11$, $p=.006$. The effect of Group was not significant, $F(1,18)=4.17$, $p=.056$. None of the interactions were significant, $Fs<2.48$, $ps>.133$.

Figure 3.1.

Experiment 1: Mean proportions of recalled objects in the free recall test.



Note. Error bars represent 95% confidence intervals.

3.3.1.2. The number of recalled object details

The same analyses were performed for the number of recalled perceptual object details (Event object details, Non-event object details). The means and standard deviations are presented in Table 3.2.

Table 3.2.
Experiment 1: Mean number of details recalled in the free recall test.

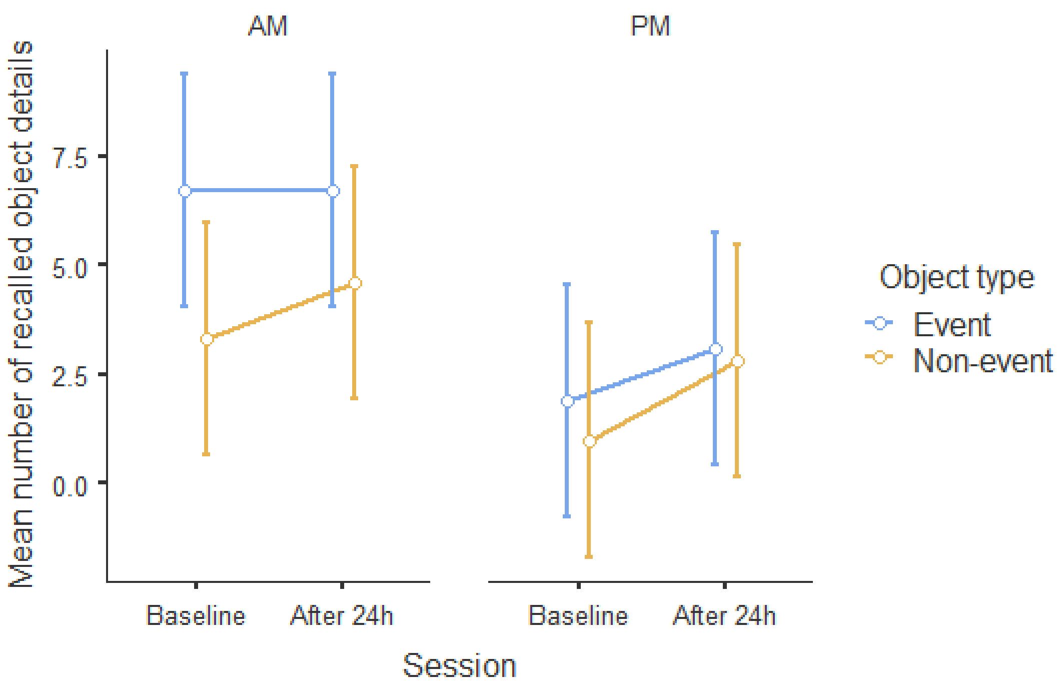
Session	**Group**	**Object type** Event	Non-event
Baseline	AM	6.70 (5.08)	3.30 (3.33)
	PM	1.90 (2.88)	1 (1.56)
After 24h	AM	6.70 (3.98)	4.60 (3.10)
	PM	3.10 (6.23)	2.8 (4.92)

Note. Numbers in parentheses represent standard deviations

An effect of Object type was found, $F(1,18)=16.28$, $p<.001$, with more Event object details being recalled than non-event, $t(18)=4.04$, p<.001 (see Figure 3.2). The effect of Session was not significant, $F(1,18)=2.10$, $p=.165$. The effect of Group was not significant, $F(1,18)=3.75$, $p=.069$. None of the other effects or interactions were significant, *Fs*<2.41, *ps*>.104.

Figure 3.2.

Experiment 1: Mean number of recalled object details in the free recall test.



Notes. Error bars represent 95% confidence intervals.

3.3.1.3. Overall word count

An overall word count produced during the task was analysed using RM ANOVA with Session (Baseline and 24h) being the within-participant and Group (AM/PM) between-participant variables. The means and standard deviations are presented in Table 3.3.

Table 3.3.

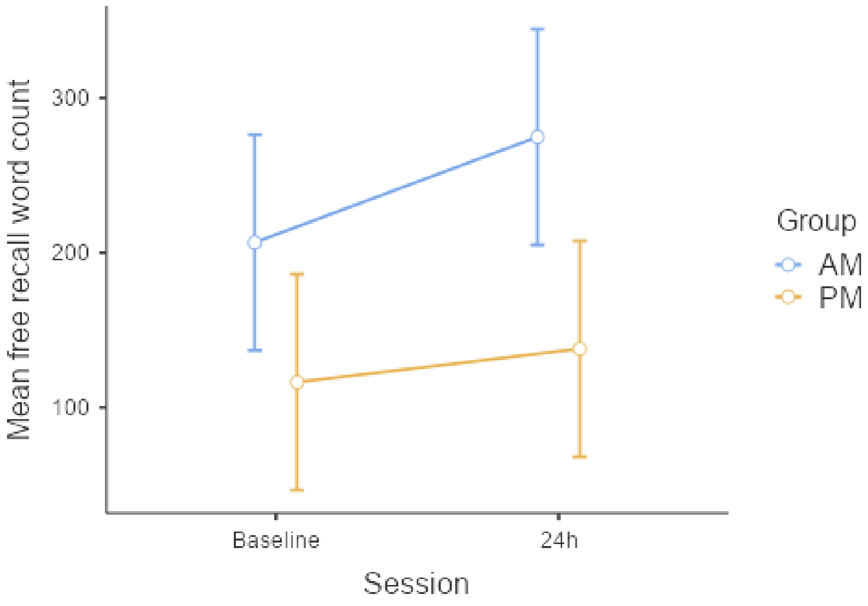
Experiment 1: Means (and SDs) of the number of words written in the free recall test.

Session	Group	Mean	SD
Baseline	AM	207	87.1
	PM	116	107
24h	AM	275	81
	PM	138	147

An effect of Session was not significant, $F(1,36)=1.70$, $p=.201$. There was a significant effect of Group, $F(1,36)=10.91$, $p=.002$, with more words written by participants in the AM group over the PM group. The Session x Group interaction was not significant, $F(1,36)=.46$, $p=.502$ (see Figure 3.3).

Figure 3.3.

Experiment 1: Mean number of words written in the free recall test.



Note. Error bars represent 95% confidence intervals.

3.3.2. Combined What-When-Where components

When a participant correctly recalled the object (What), the room number the object was in (When) and where in the room the object was (Where) it was said that the participant recalled the full WWW information regarding that object. The combined WWW score is the proportion of correctly recalled WWW information out of all given objects (ranges from 0 to 1). For example, a score of 0.5 would mean that a participant recalled combined WWW information for 18 objects out of 36 possible. The means and standard deviations are presented in Table 3.4.

Table 3.4.

Experiment 1: Mean proportions of combined What-When-Where recalls.

		Object type	
Session	**Group**	Event	Non-event
Baseline	AM	.63 (.22)	.54 (.23)
	PM	.39 (.24)	.36 (.27)
After 24h	AM	.44 (.15)	.28 (.23)
	PM	.30 (.22)	.13 (.17)

Note. Numbers in parentheses represent standard deviations

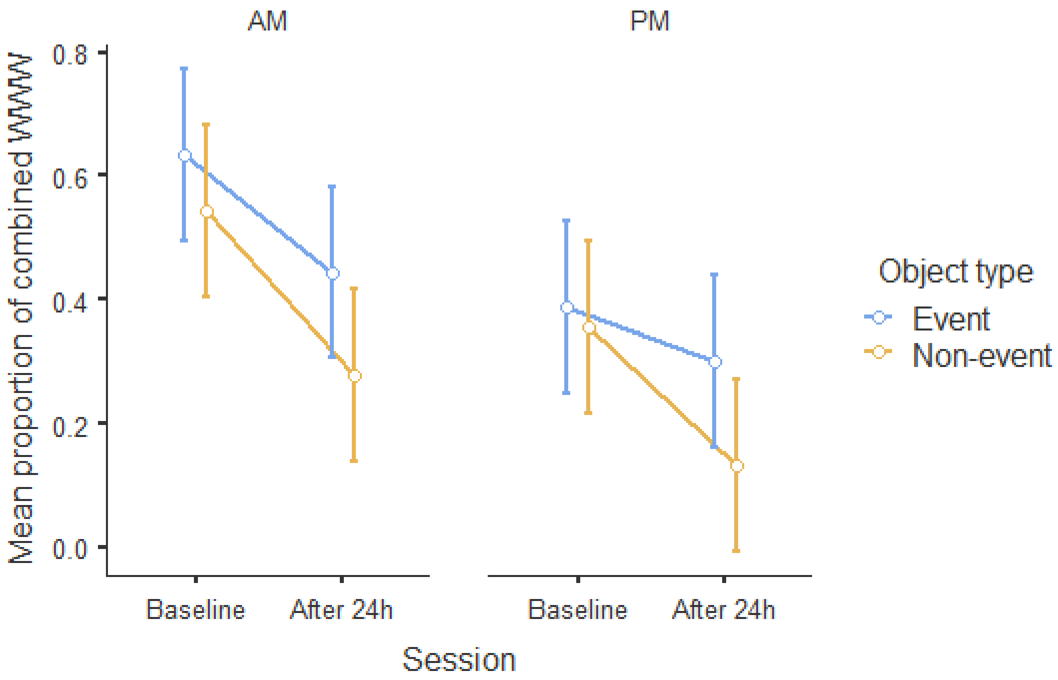
A significant effect of the object type was observed, $F(1,18)=10.99$, $p=.004$, with Event objects having higher correct WWW proportions compared to the non-Event objects (see Figure 3.5). A significant effect of Session was observed, $F(1,18)=15.06$, $p=.001$, with higher correct proportions at Baseline than after 24h.

A significant Session x Object type interaction was detected, $F(1,18)=4.53$, $p=.047$. Multiple comparisons revealed Event objects having higher correct WWW proportions in the 24h Session compared to the Non-event objects, $t(32.7)=3.93$, $p=.002$, higher Non-event proportions at Baseline compared to 24h Session, $t(26.5)=4.42$, $p<.001$, and higher Event object proportions at Baseline compared to Non-event object proportions after 24h, $t(32.1)=5.08$, $p<.001$. Event and non-event proportion did not differ at the Baseline, $t(32.7)=1.44$, $p=.952$.

A significant effect of Group was found, *F*(1,18)=6.18, *p*=.023, with higher proportions in the AM group compared to the PM group.

Figure 3.4.

Experiment 1: Mean proportions of combined What-When-Where recalls.



Notes. The combined WWW score is the proportion of correctly recalled What-When-Where information for all of the possible objects (ranges from 0 to 1). Error bars represent 95% confidence intervals.

3.3.3. Separate components

3.3.3.1. What

The What component represents a recall of an object. On the task screen, it is worded as “Do you recall X?” where X is a name of an object. Similarly to the combined WWW, the What component was measured as a proportion of correctly recalled objects. The means and standard deviations are presented in Table 3.5.

Table 3.5.

Experiment 1: Mean proportions of What (object) recalls

Session	Group	Object type: Event	Non-event
Baseline	AM	.90 (.17)	.87 (.08)
	PM	.83 (.09)	.83 (.15)
After 24h	AM	.71 (.20)	.58 (.13)
	PM	.59 (.22)	.51 (.18)

Note. Numbers in parentheses represent standard deviations

A significant effect of the object type was observed, $F(1,18)=6.00$, $p=.025$, with Event objects having higher proportions compared to the Non-event objects. A significant effect of Session was observed, $F(1,18)=83.71$, $p<.001$, with higher proportions at Baseline than after 24h (see Figure 3.5). An effect of Group was not found, $F(1,18)=2.29$, $p=.147$. None of the interactions were significant, $Fs<1.45$, $ps>.244$.

3.3.3.2. When

The When component represents a correct recall of a room number in which the object from the What task was seen. On the task screen, it is worded as "In which room you have seen X?" where X is the name of an object from the previous task. Similarly to the combined WWW, the When component was measured as a proportion of correct room recalls out of all possible. The means and standard deviations are presented in Table 3.6.

Table 3.6.

Experiment 1: Mean proportions of When (temporal) recalls

Session	Group	Object type: Event	Non-Event
Baseline	AM	.74 (.09)	.83 (.17)
	PM	.60 (.25)	.63 (.28)
After 24h	AM	.56 (.18)	.41 (.24)
	PM	.44 (.31)	.31 (.20)

Note. Numbers in parentheses represent standard deviations

A significant effect of Session was observed, $F(1,18)=48.95$, $p<.001$, with higher proportions at Baseline than after 24h (see Figure 3.5). A significant effect of the object type was observed, $F(1,18)=12.49$, $p=.002$, with Event objects having higher proportions compared to the Non-event objects. Effect of Group was not found, $F(1,18)=3.12$, $p=.095$. None of the interactions were significant, *Fs*<.99, *ps*>.779.

3.3.3.3. Where

The Where component represents correctly recalling the object's location in the virtual room. On the task screen, participants had to use a mouse and point on a top-down map of the room where they thought the object was located. This provided a distance – how far away the participant's guess was from the object's real location. Using a method explained in Chapter 2 (section 2.7) the distance was converted to a binary correct/incorrect outcome. Similarly to the combined WWW, the Where component was measured as a proportion of correct location recalls out of all possible. A high correlation was observed (as a measure of validity) between the Where pointing errors and the Where proportions ($r = -.911$, $n = 80$, $p<.001$). The means and standard deviations are presented in Table 3.7.

Table 3.7. Experiment 1: Mean proportions of Where (spatial) recalls.

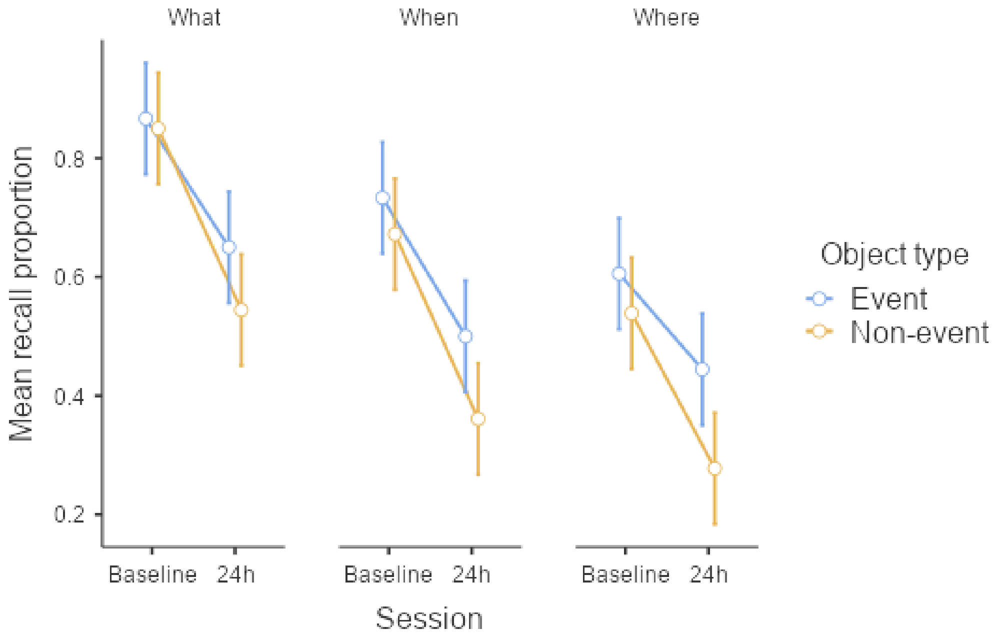
		Object type	
Session	**Group**	Event	Non-event
Baseline	AM	.66 (.23)	.62 (.25)
	PM	.56 (.33)	.46 (.25)
After 24h	AM	.56 (.21)	.38 (.24)
	PM	.33 (.20)	.18 (.21)

Note. Numbers in parentheses represent standard deviations

A significant effect of Session was observed, $F(1,18)=21.01$, $p<.001$, with higher correct proportions at Baseline than after 24h (see Figure 3.5). A significant effect of the object type was observed, $F(1,18)=8.21$, $p=.010$, with Event objects having higher correct proportions compared to the Non-event objects. A significant effect of Group was found, $F(1,18)=4.55$, $p=.047$, with higher correct proportions in the AM group compared to the PM group. None of the interactions were significant, *Fs*<1.79, *ps*>.197.

Figure 3.5.

Experiment 1: Mean proportions of combined What-When-Where recalls.



Notes. Due to only Where component having an effect of Group (*p*=.047), the Group variable was not included in the figure. Error bars represent 95% confidence intervals.

3.3.3.4. Event

The Event component represents a proportion of correctly recalled events associated with the recalled event objects. On the task screen, participants were asked if any events happened to the recalled object (What). If they indicated that an event happened with the recalled object, they were asked to write it down. The written event descriptions were converted to binary (correct/incorrect) score and later into a proportion of correctly recalled events. The means and standard deviations are presented in Table 3.8.

Table 3.8.

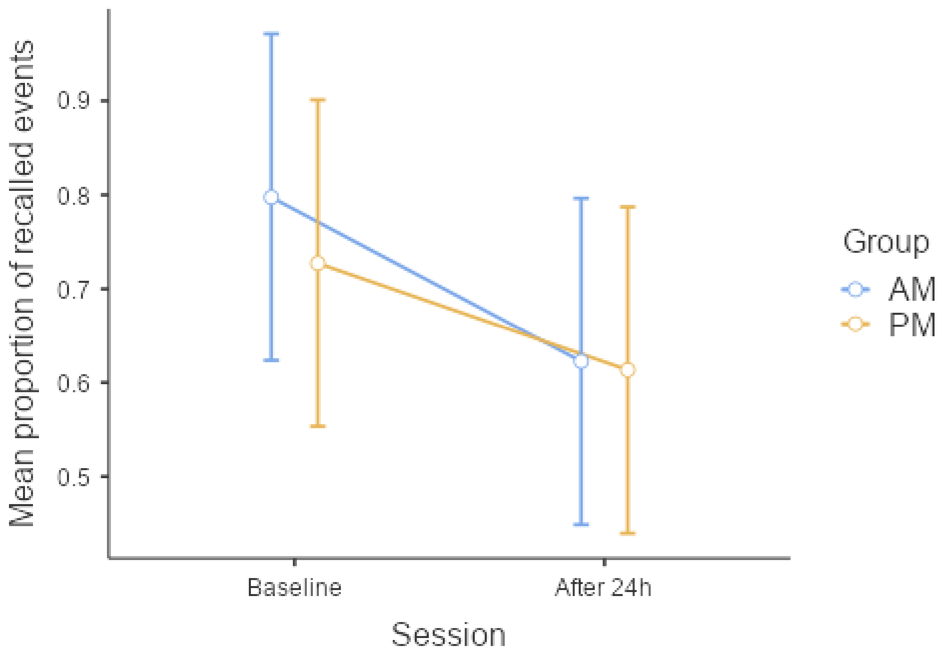
Experiment 1: Mean proportions of Event recalls.

Session	Group	Mean	SD
Baseline	AM	.80	.25
	PM	.73	.19
After 24h	AM	.62	.33
	PM	.61	.30

An effect of Session was not significant, $F(1,18)=3.85$, $p=.066$. An effect of Group was not significant, $F(1,18)=.17$, p=.684. The Session x Group interaction was not significant, $F(1,18)=.17$, $p=.683$ (see Figure 3.6).

Figure 3.6.

Experiment 1: Mean proportions of recalled events associated with the event-objects.



Note. Error bars represent 95% confidence intervals.

3.3.3.5. Detail

The Detail component represents a mean number of recalled details for one object. On the task screen, participants were asked if they could recall any perceptual detail about an object and if they could write one down. Participants were able to write down up to five details per one object. The means and standard deviations are presented in Table 3.9.

Table 3.9.

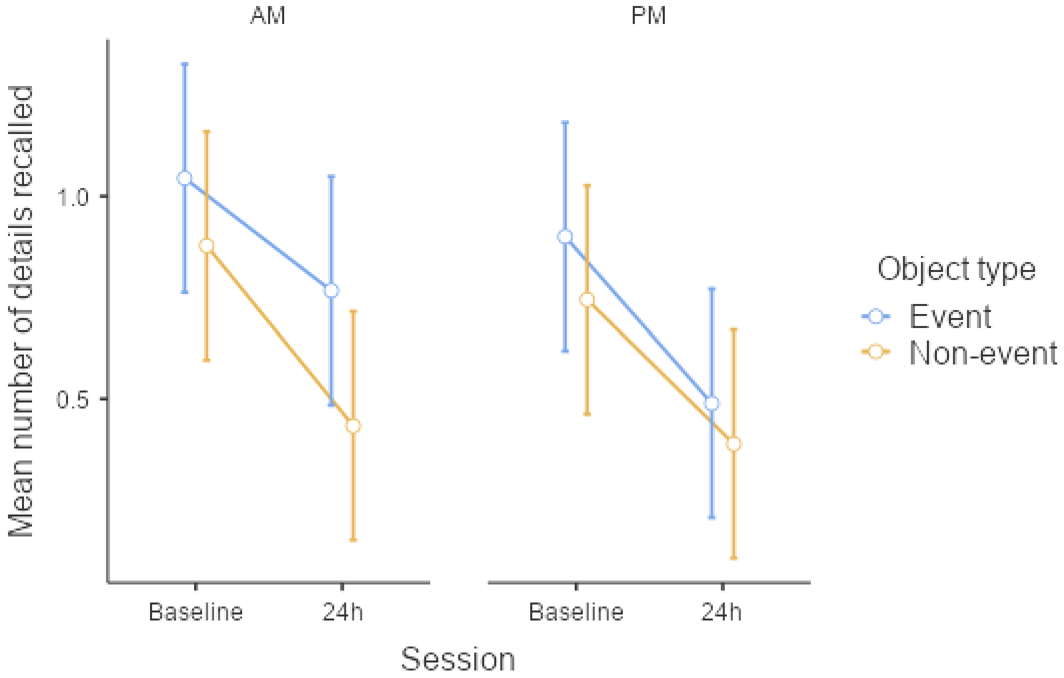
Experiment 1: Mean number of recalled perceptual details per one recalled object

		Object type	
Session	**Group**	Event	Non-event
Baseline	AM	1.04 (.36)	.88 (.43)
	PM	.90 (.62)	.74 (.22)
After 24h	AM	.77 (.48)	.43 (.28)
	PM	.49 (.57)	.39 (.41)

A significant effect of Session was observed, $F(1,18)=53.98$, $p<.001$, with more details recalled at Baseline than after 24h. A significant effect of the object type was observed, $F(1,18)=6.86$, $p=.017$, with participants recalling more details for the Event objects compared to the non-event objects. An effect of Group was not found, $F(1,18)=.88$, $p=.362$. None of the interactions were significant, $Fs<.576$, $ps>.408$ (see Figure 3.7).

Figure 3.7.

Experiment 1: Mean number of recalled perceptual details per one recalled object.



Notes. Error bars represent 95% confidence intervals.

3.3.4. Remember/know/guess judgements

After participants provided Detail and Event component information, they were asked to indicate if they Remembered, Knew or Guessed about that information. An RM ANOVA was used to analyse the Event and Detail component remember/know/guess judgements. The judgements were transformed into overall proportions using a similar method to Dewhurst, Conway, & Brandt (2009). For example, adding one participant's Remember, Know and Guess judgement proportions for the Event component at Baseline would equal 1. This transformation was undertaken so that the lower number of recalled objects on the second session would not affect the judgement data.

3.3.4.1. Event

Due to no Guess judgements for the Event objects in the 24h session, only Remember and Know judgements were analysed. The means and standard deviations of the R/K/G judgements are presented in Table 3.10.

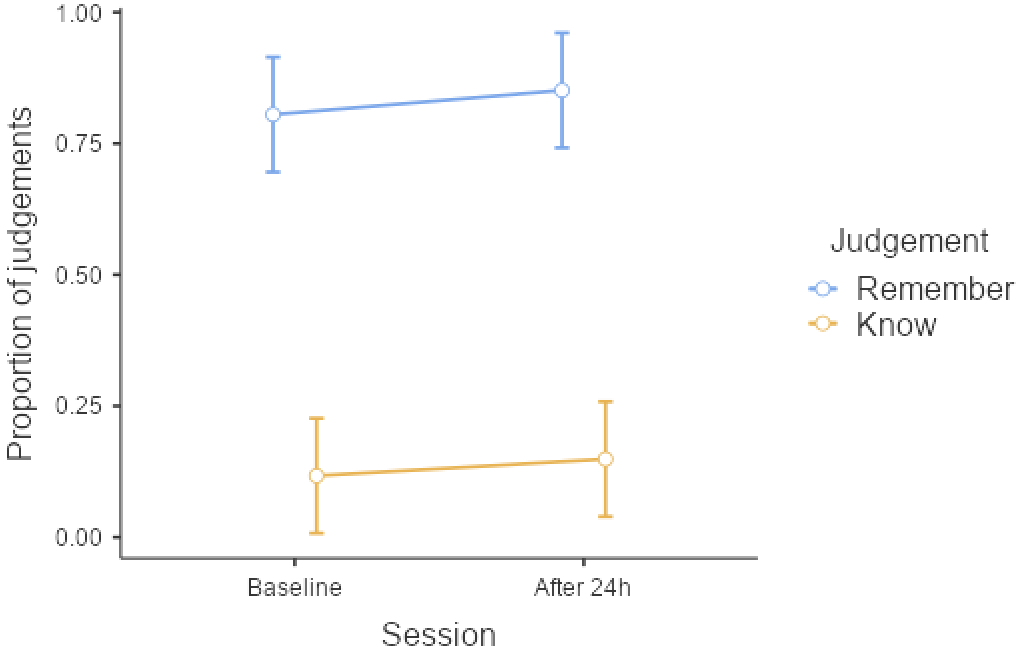
Table 3.10. Experiment 1: Mean proportions of the Remember/Know/Guess judgements given for the Event component.

Session	Group	Judgement	Mean	SD
Baseline	AM	Remember	.85	.19
		Know	.10	.15
		Guess	.05	.07
	PM	Remember	.76	.27
		Know	.13	.18
		Guess	.10	.14
After 24h	AM	Remember	.84	.22
		Know	.16	.22
		Guess	0	0
	PM	Remember	.87	.32
		Know	.13	.32
		Guess	0	0

A significant effect of Session was observed, $F(1,18)=9.80$, $p=.006$ (see Figure 3.8). This is associated with the lack of Guess judgements in the 24h session. As a result of that, the remaining Remember and Know proportions were higher at the 24h sessions as compared to the Baseline session. There was a significant effect of Judgement, $F(1,18)=42.60$, $p<.001$, with a higher proportion of Remember judgements than Know. No other interaction were significant, *Fs*<.98, *ps*>.335.

Figure 3.8.

Experiment 1: Mean proportions recalled of Remember judgements.



Notes. Error bars represent 95% confidence intervals.

3.3.4.2. Detail

Same analyses were performed as in the Event R/K/G data. The means and standard deviations are presented in Table 3.11.

Table 3.11.

Experiment 1: Mean proportions of the Remember/Know/Guess judgements given for the Event component.

			Object type	
Session	**Object type**	**Judgement**	Event	Non-event
Baseline	Event	Remember	.76 (.28)	.79 (.21)
		Know	.15 (.17)	.15 (.23)
		Guess	.12 (.17)	.15 (.22)
After 24h	Event	Remember	.63 (.33)	.60 (.34)
		Know	.21 (.25)	.22 (.25)
		Guess	.16 (.23)	.09 (.19)

Note. Numbers in parentheses represent standard deviations

The effect of Session was not significant, *F*(1,18)=3.41, *p*=.081. There was a significant effect of Judgement, *F*(2,36)=42.35, *p*<.001, with higher proportion of Remember than Know, *t*(36)=7.57, p<.001, or Guess, *t*(36)=8.32, *p*<.001, judgements. Know and Guess judgement proportion did not differ, *t*(36)=.74, *p*=1.

There was a significant Session x Judgement interaction, *F*(2,36)=6.26, *p*=.005 (see Figure 3.9). Remember judgement proportions were lower after 24h in comparison to Baseline. At both time points Remember judgement proportions were higher than both Know and Guess judgement proportions. Know and Guess judgement proportions did not differ at both time points. The pairwise comparisons are presented in Table 3.12. No other interaction were significant, *Fs*<3.46, *ps*>.079.

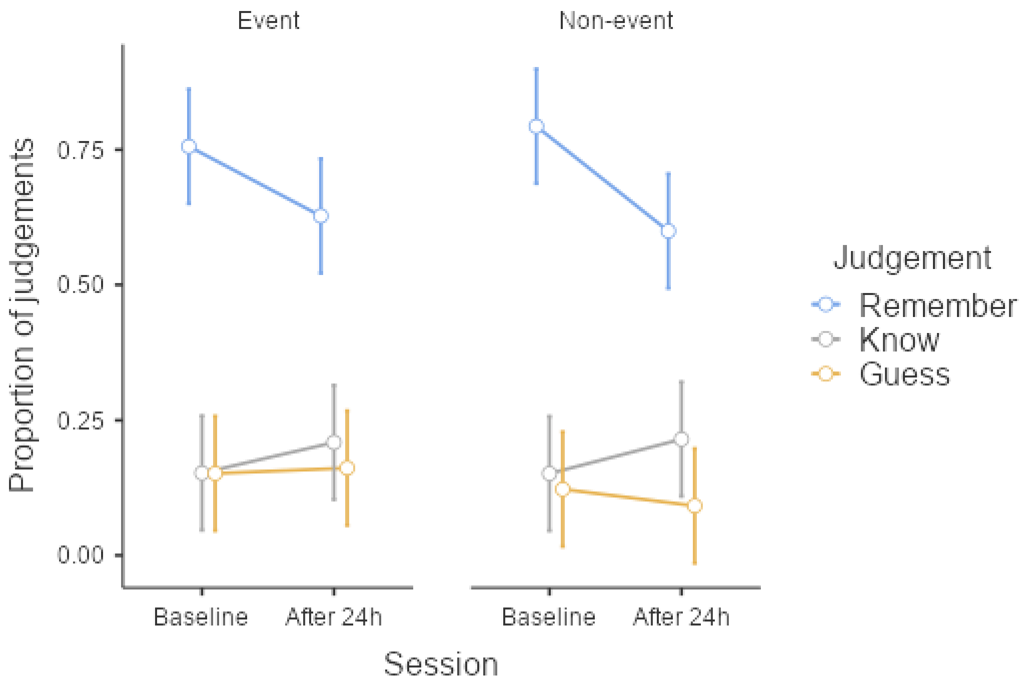
Table 3.12.

Experiment 1: Bonferroni corrected comparisons for the Remember/Know/Guess judgements given for the Detail component.

Comparison								
Session	Judgement	Session	Judgement	Mean Difference	SD	df	t	p
Baseline	Remember	Baseline	Know	.62	.33	51.30	8.33	< .001
		Baseline	Guess	.64	.33	51.30	8.53	< .001
		After 24h	Remember	.16	.19	51.50	3.84	0.005
		After 24h	Know	.56	.33	47.70	7.72	< .001
		After 24h	Guess	.65	.33	47.70	8.89	< .001
	Know	Baseline	Guess	.01	.33	51.30	.20	1
		After 24h	Remember	-.46	.33	47.70	-6.33	< .001
		After 24h	Know	-.06	.19	51.50	-1.43	1
		After 24h	Guess	.03	.33	47.70	.35	1
	Guess	After 24h	Remember	-.48	.33	47.70	-6.53	< .001
		After 24h	Know	-.08	.33	47.70	-1.03	1
		After 24h	Guess	.01	.19	51.50	.25	1
After 24h	Remember	After 24h	Know	.40	.33	51.30	5.37	< .001
		After 24h	Guess	.49	.33	51.30	6.51	< .001
	Know	After 24h	Guess	.09	.33	51.30	1.14	1

Figure 3.9.

Experiment 1: Experiment 1: Mean proportions of the Remember/Know/Guess judgements given for the Event component.



Notes. Error bars represent 95% confidence intervals.

3.3.5. Recognition – d' scores

Recognition data was converted to d' scores (Z(hit rate) − Z(false alarm rate)). The means and standard deviations are presented in Table 3.13.

Table 3.13.

Experiment 1: Mean d' object recognition scores

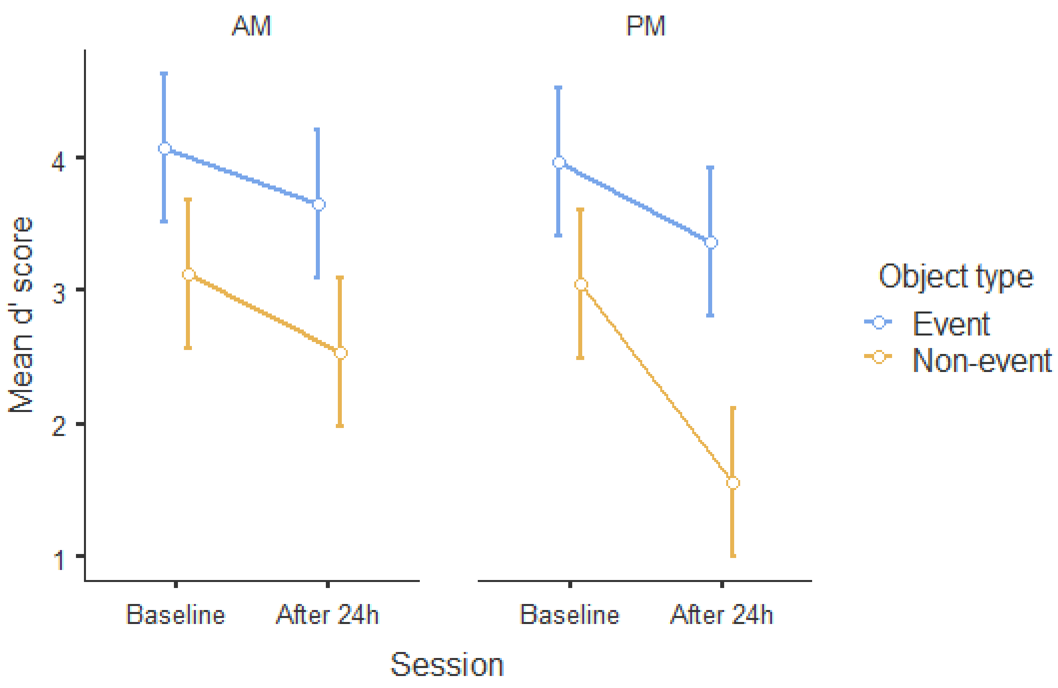
		Object type	
Session	**Group**	Event	Non-event
Baseline	AM	4.07 (.76)	3.13 (1.11)
	PM	3.98 (.75)	3.05 (1.10)
After 24h	AM	3.65 (.87)	2.54 (.76)
	PM	3.37 (.68)	1.56 (.92)

Note. Numbers in parentheses represent standard deviations

There was a significant effect of Session, $F(1,18)=25.01$, $p<.001$, with higher d' scores at Baseline than after 24h (see Figure 3.10). There was a significant effect of Object type, $F(1,18)=38.23$, $p<.001$, with Event objects having higher d' scores than non-event objects. The effect of Group was not significant, $F(1,18)=1.87$, $p=.188$. No interactions were significant, $Fs<2.74$, $p>.102$.

Figure 3.10.

Experiment 1: Mean d' object recognition scores.



Notes. Error bars represent 95% confidence intervals.

3.3.6. Recognition – Confidence ratings

The confidence ratings were ratings (from .0 to 1.0) reflecting how confident the participants felt about their recognition judgement. A confidence rating of .0 would indicate being not confident at all whereas confidence rating of 1.0 would indicate full confidence. The means and standard deviations are presented in Table 3.14.

Table 3.14.

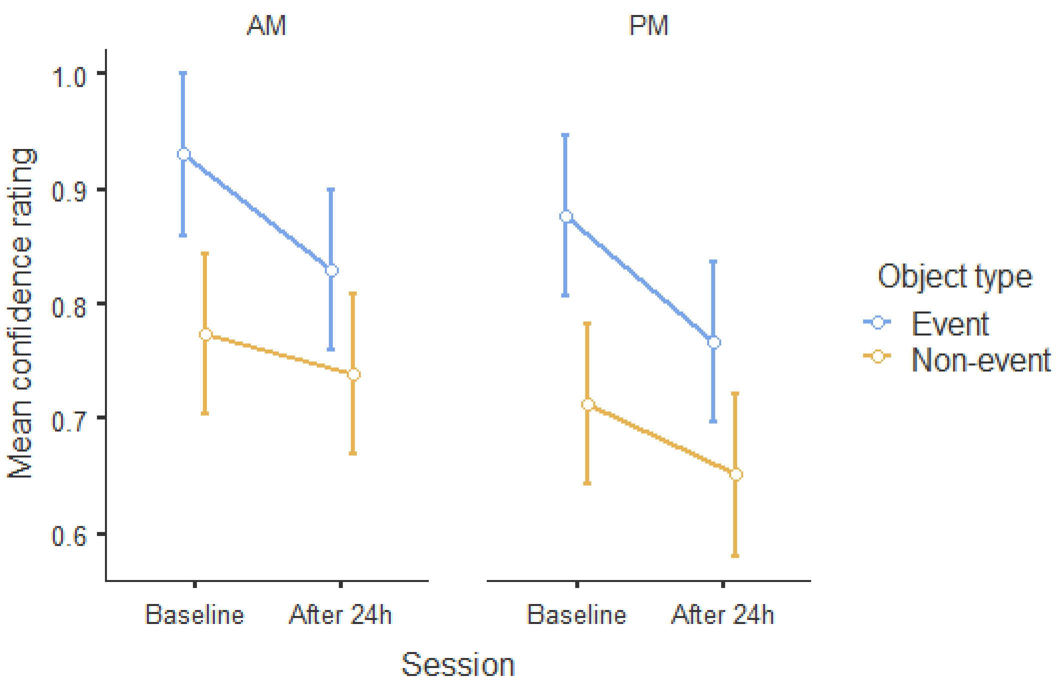
Experiment 1: Mean object recognition confidence ratings.

		Object type	
Session	**Group**	Event	Non-event
Baseline	AM	.93 (.14)	.77 (.21)
	PM	.88 (.24)	.71 (.29)
After 24h	AM	.83 (.24)	.74 (.24)
	PM	.77 (.28)	.65 (.29)

There was a significant effect of Session, *F*(1,18)=18.07, *p*<.001, with higher confidence ratings at Baseline than after 24h (see Figure 3.11). There was a significant effect of Object type, *F*(1,18)=43.13, *p*<.001, with Event objects having higher confidence ratings than non-Event objects. The effect of Group was not significant, *F*(1,18)=3.15, *p*=.093. No interactions were significant, *Fs*<2.64, *p*>.122.

Figure 3.11.

Experiment 1: Mean object recognition confidence ratings



Notes. Error bars represent 95% confidence intervals.

3.3.7. Actigraphy data

Actigraphy data from 3 participants was not retrieved due to technical problems which led to no data being recorded. For the means and standard deviations of the total time spent asleep, time spent in REM sleep and time spent in deep sleep (SWS) can be seen in Table 3.15.

Table 3.15.

Experiment 1: Mean total time spent asleep, time spent in slow-wave sleep and time spent in REM sleep.

Group	Total sleep duration	Slow-wave sleep	REM sleep
AM	424.38 (111.02)	74.58 (24.7)	104.75 (51.03)
PM	448.35 (134.32)	70.28 (45.83)	148.52 (107.15

Separate ANOVAs were performed to see if there were an effect of group on the total time spent asleep, time spent in REM sleep and time spent in SWS. Effect of Group was not significant in any of the three measures, *Fs*<.145, *ps*>.338.

The three measures showed no significant correlation with free recall, combined WWW, and recognition data (*ps*>.054) apart from the negative correlation between time spent in REM and d’ scores for Event objects (r=-0.525, *p*=.037) (see Table 3.16).

Table 3.16.

Correlation matrix for actigraphy data, free recall, combined WWW proportion and d' scores from the object recognition task (using data from the 24h session).

	Sleep duration	Time spent in slow-wave sleep	Time spent in REM sleep
Event objects	-0.06	-0.35	-0.489
Non-event objects	0.26	-0.019	-0.116
Event object details	0.06	-0.193	0.121
Non-event object details	0.2	-0.052	0.165
Combined WWW proportion for Event objects	0.16	0.125	0.126
Combined WWW proportion for Non-event objects	0.12	-0.084	0.036
d' score for Event objects	-0.26	-0.341	-0.525*
d' score for Non-event objects	0.02	-0.234	0.267

Notes: * $p < .05$, ** $p < .01$, *** $p < .001$. The p values were corrected using Bonferroni method.

3.4. Discussion

In this chapter, a novel approach was used to explore EM by exposing subjects to highly immersive, sensory-perceptual events utilising HMD-VR as a valid manipulation of EM. EM recall and recognition for both events and non-events were presented and compared on a number of different outcome measures of EM. As it is an integral part of everyday memory formation and retainment (Inostroza & Born, 2013), the effect of sleep dependant consolidation on EM via the effect of time of sleep was investigated.

One of the main aims of the present experiment was to explore the effect of sleep dependant consolidation and the effect of AM/PM testing on the EM. The prediction was that after the 24h participants in the PM (compared to the AM) group should have overall better performance in the EM tests. The present data did not support this prediction. There were no Group differences which suggest that the time of sleep did not affect the EM consolidation. This lack of difference goes against a body of literature stating that having the first part of 24h filled with sleep (PM group) would lead to better EM performance (Benson & Feinberg, 1977; Payne et al., 2008, 2012; Scullin, 2014; Talamini et al., 2008). To the contrary, in the free recall, combined WWW and separate Where component data, the AM group showed better performance than the PM group. In the rest of the measures, there was no difference between the groups. It is worth pointing out that the *p*-value for the Group effect in the free recall object data (p=.056) was close to the significance level of 0.05, with the trend visible in the recalled object figure (Figure 3.1). This indicates the possibility that perhaps with a larger sample size the effect could have been detected significantly. In general, these findings add to the body of research showing lack of effect of early sleep on EM performance (Sheth et al., 2009; Studte et al., 2015; van der Helm et al., 2011; Wilhelm et al., 2011). The present data shows that the time of encoding and the time between the encoding and nocturnal sleep does not affect the recall of EM.

The higher AM group's performance could be partly explained by circadian rhythms and arousal. Research shows that the time of day has an impact on arousal, attention and memory with higher memory performance observed in the morning compared to late-night (Baddeley et al., 1970; Barrett & Ekstrand, 1972; Folkard & Monk, 1980; May et al., 1993). For example, a study by Folkard & Monk (1980) showed that immediate memory was better in the first part of the day (9 am to 2 pm) and dropping significantly in the second part (2 pm to 11 pm). Similar results were also observed by Baddeley et al. (1970) with both papers giving human circadian rhythm and increased arousal in the morning as explanations. To better understand sleep-dependent memory consolidation and the effect of time of sleep it would be useful to have at least two sessions in one day. For example, this could be achieved by repeating a similar experiment as in the present chapter but testing EM every 12h instead of 24h (similarly to Aly & Moscovitch, 2010 or Cairney et al., 2011). This would allow exploring the differences in consolidation during both wake and sleep.

One of the other main aims of the present experiment was to investigate how EM for event objects might differ to EM for non-event objects. The prediction was that event objects would be better remembered than the non-event objects. As predicted, in all three tests (free recall, WWW, and object recognition) event objects were better recalled than the non-event objects supporting the idea that more life-like events are better recalled than static EMs. The results from the free recall task showed that event objects were overall better recalled than the non-event objects with no drop in this recall after 24h. A similar effect was observed when looking at the recalled details for the objects with more details being recalled for the event objects than the non-event objects. Interestingly, there was an effect of Session with more (both event and non-event) objects being recalled after 24h compared to Baseline, but the same effect was not found for the overall (both event and non-event) number of recalled details. This is an interesting finding as the effect of Session but no Session x Object type interaction shows that recall for both types of objects improved equally after consolidation.

Similarly to the free recall task, event objects had overall more combined WWW recalls than non-event recalls (effect of Object type). When looking at the effect of Session, participants made fewer combined WWW recalls for the event objects after 24h compared to immediately after the exploration. While this is the opposite from the object data from the free recall task it does show that it is more difficult to recall combined WWW information than the individual components as in the free recall task. However, when looking at the separate What component data (Table 3.5), it is visible that, as in the combined WWW proportions, the recalls dropped after 24h. An explanation for the difference in trends between the free recall object data and the What component data is that the free recall task was identical in both sessions. Participants were asked to recall the whole experience in both Baseline and 24H sessions whereas in the What test participants were given different sets of objects. The word count analysis provides some support for this as there was no effect of Session. This might indicate that participants in the second session (after 24h) were able to recall what they had written in the previous session and add to that leading to the practice effect (Benedict & Zgaljardic, 1998; McCabe, Langer, et al., 2011).

An important finding is the Session x Object type interaction for the combined WWW proportions. The significant difference between the two event types at Baseline but higher proportions for the event object after 24h, showing two important points. First of all, it shows that immediately after encoding, EM performance (operationalised as the combined WWW information) for event objects is at the same level as for non-event objects. It suggests that event objects were not preferentially encoded. Secondly, the difference at the 24h, indicates that event objects were preferentially consolidated over the non-event objects. This can be explained by the memory trace ‘tagging’ (Fernández & Morris, 2018) as discussed in the introduction. Memory traces associated with the event objects were ‘tagged’ as more ‘important’ for consolidation. If going with this explanation, it shows that naturalistic events are preferentially consolidated over static stimuli. Selective consolidation can also explain the enhanced memory for Event objects after the 24h period. It is possible to argue that the retention period in the present experiment, half of it being filled with sleep, led to a preferential consolidation of full EMs (combined WWW), and not separate components of the event objects. This is due to the mentioned Session x Object type interaction only observed in the combined WWW data. It is possible to speculate that event objects showed greater binding of the separate WWW components (Kessels et al., 2007) which was further strengthened by consolidation, resulting in this interaction for the combined WWW but not for separate WWW components. It is also possible to argue that event objects showed greater binding of the separate WWW components (Kessels et al., 2007) during the consolidation, resulting in this interaction for the combined WWW but not for separate WWW components. However, it is important to note that the significance level of the Session x Object type interaction was low (p=.047) and as such replication is required in the future.

When looking at the separate WWW components, the effect of object type continued to be significant with event objects having higher recalled proportions in all three components. As in the combined WWW proportions, all separate components were better recalled immediately after the test than after 24h. However, unlike in the combined WWW proportions, there were no Session x Object type interactions for any of the components. Object recognition data followed with similar finding with event objects showing higher d’ score and thus being more accurately recognised than non-event objects. The confidence rating data added to this by showing that participants were overall more confident recognising event objects. The effect of Object type on the d’ scores provides an important insight showing that more lifelike stimuli are not only better remembered (as shown by the free recall and WWW data) but also better recognised. As it has been discussed in the introduction, recognition memory while associated with EM is dependent on different brain regions and processes (Aggleton & Brown, 2006; Chen et al., 2017). The present data shows that even due to this difference, lifelike stimuli are better recalled than static stimuli. The confidence ratings add to this by showing that recalling the event objects was 'easier' compared to the non-event objects.

The additional and exploratory, actigraphy and Remember/Know/Guess measures provided mixed results. Our prediction that REM sleep would positively correlate with EM performance (Rasch & Born, 2013; Rauchs et al., 2004; Siegel, 2001) was not supported. Out of all the measures only d' scores for the event objects correlated with REM sleep with no other correlations showing significance. While this correlation does provide some support for the initial prediction, the lack of other significant correlations is concerning. One explanation for this is that the sleep-related measures were taken using a type of wrist actigraphy trackers that recently have been found to be inaccurate and particularly poor at identifying REM sleep (Cook et al., 2019). For a more accurate exploration of sleep stage effects on EM, a more reliable measuring system is needed.

While the Remember/Know/Guess data did not provide any useful insights this was not unexpected. The R/K/G judgement proportion for the Events showed no changes in the judgement proportions over the 24h period. A more interesting finding can be seen in the Detail R/K/G judgement data. While the overall effect of Session was not significant, the Session x Judgement interaction was significant. The multiple comparisons (see Table 3.10) revealed that this was due to the reduced Remember judgement proportions in the 24h session. The Know and Guess proportion showed no change. This indicates that there was a partial Remember-Know shift. However, the Remember-Know shift, as discussed in Chapter 1 (section 1.5), is usually tracked over a few week period (Dewhurst et al., 2009; Herbert & Burt, 2003) and it might not become visible after just 24h. As such, further research is needed with longer timeframes between the testing sessions.

The present experiment aimed to explore EM by exposing subjects to life-like events, the effect of sleep dependant consolidation on EM and the effect of time of sleep. The experiment provided an important insight into EM by showing that EM differs for events, arguably being more realistically encountered in our everyday interactions, compared to static experience-encoding. This finding shows the importance of ecologically valid and function-based EM testing. While the effect of time of sleep was not found in the majority of analyses, indicating that perhaps most EMs do not benefit from sleep-dependent memory consolidation over a 24 hour period, there was enhanced consolidation of events. Particularly in the combined WWW measure.

The thesis aims combined with the present findings lead to a number of questions that need to be further investigated. First of all, as pointed out in the discussion of the AM/PM results, the level of significance (*p*=.056) might indicate the possibility that the effect was missed by the present experiment. As such, the AM/PM testing needs to be repeated. Continuing with the exploration of sleep and consolidation it is important to bring back the point made while discussing the Remember/Know/Guess results that EM consolidation may not be evident over the 24 hour period. Research has shown a high drop in memory retention during the first week after encoding which then turns to a more steady linear decrease in memory accessibility (Furman et al., 2012; Moreton & Ward, 2010; Thompson, 1982; Tunney, 2010; Tunney & Bezzina, 2007). As a result, effects such as the Remember-Know shift may become more evident over a longer 30 day period than the present 24h period. In general, the increase in the retention period should provide a better understanding of life-like EM consolidation.

Chapter 4 – Experiment 2: Long term consolidation of event and non-event based episodic memory

4.1. Introduction

Experiment 1 in Chapter 3 explored the use of event objects (events) presented through HMD-VR and how episodic memory (EM) for those objects differ from static objects (non-events) commonly used in memory research. Experiment 1 employed a 24h AM/PM design, which allowed an investigation of the effect of the time period between learning and sleeping and a general effect of consolidation on EM. While the literature strongly suggested that sleep immediately following learning should facilitate EM consolidation, this was not observed in the previous experiment. To the contrary, in some cases, such as the combined WWW and the separate Where component, the opposite effect was found even if the significance level was high. The present experiment aimed to replicate the previous findings with addition of 7-day and 30-day testing sessions to investigate the effect of longer-term EM consolidation.

As it has been discussed in Chapter 1 (section 1.7), EM changes over time. Memory decay, interference and consolidation contribute to the forgetting and retention of information (see Hardt et al., 2013; Sadeh et al., 2014 for review). However, forgetting is not an all-or-nothing process. Different rates of forgetting have been identified for different aspects of memory (Bahrick, 1984; Brainerd & Reyna, 1993; Furman et al., 2012; Sekeres et al., 2016). As discussed previously, EM is especially prone to loss of detail information, with focal elements being critical to the overall coherence of an event, more likely to be retained compared to the contextual information (Dobbins et al., 2002; Fujii et al., 2004; Hayes et al., 2004; Thorndyke, 1977). While Experiment 1 provided some initial insights into the effect of one night's sleep on consolidation, the question arises as to how EM performance may change over a longer course of time.

A pattern of forgetting can be seen in a number of studies showing a high drop in memory retention during the first week which afterwards turns to a more steady linear decrease in memory accessibility (Furman et al., 2012; Moreton & Ward, 2010; Thompson, 1982; Tunney, 2010; Tunney & Bezzina, 2007). For example, Talamini & Gorree (2012) investigated how different memory elements changed over five different intervals ranging from 5 minutes to 3 months. The study found stronger forgetting of configurational components (location, detail) with respect to the featured objects. This finding goes along with the discussed research (see Chapter 1, section 1.7.1) showing that perceptual and contextual details are first to be forgotten and directly relates to the loss of Where (and When) contextual WWW components (Fujii et al., 2004; Hayes et al., 2004). The difference in forgetting developed between one week and one month after encoding. Memory for general object recognition remained highest compared to recognition for object-location associations, object-object associations or recognition for object details. This again relates to the WWW component forgetting trends, by showing that the general object recognition (related to What) shows less forgetting than contextual information (such as Where). The study showed that initially, the reduction in recognition showed a curvilinear pattern which turned to linear after around a week. However, this trend was not observed in a study by Furman, Dorfman, Hasson, Davachi, & Dudai (2007) which showed no difference between 3h and one-week tests. Furthermore, the study by Talamini & Gorree (2012) also looked at memory loss of detail and found that it progressively increased over time but, as before, the progression turned linear after a week. In addition to this, recognition for details was lowest compared to any other recognition, such as item, which is again supported by literature of time-dependant detail loss. This remained up until around a two-month mark when recognition for location information became the lowest. These trends indicate that EMs are rapidly fragmented with memory for item information (What) showing less forgetting than for the contextual information (When, Where and Detail), which relates to the loss of episodic nature of memories (Sekeres et al., 2016; Tulving, 1972)

Following the memory trace theories discussed in Chapter 1 (section 1.7.1), it is possible to explain this by selective memory trace activations. As discussed, memory traces are reactivated and replayed during sleep (Deuker et al., 2013; Peigneux et al., 2004; Skaggs & McNaughton, 1996). What is more, memory traces are reactivated and reconsolidated with each retrieval (Schwabe et al., 2014; Winters et al., 2009). These reactivations strengthen traces for the components that are being recalled and weaken components that are not. This leads to gist extraction and semantisation of memories (Dudai et al., 2015; Meeter & Murre, 2004). Using this information and the data from the Experiment 1, showing better memory performance for events than non-events after 24h, the present experiment aimed to explore how memory performance for events and non-events change over a longer period. To explore this, two additional testing sessions were added to the previous experiment – after 7 and 30 days from the initial exploration.

As discussed in the previous chapters, memory consolidation is closely linked to sleep and the reorganisation of memory patterns during post-learning sleep (Diekelmann & Born, 2010; Huber et al., 2004; Peigneux et al., 2004; Stickgold & Walker, 2007). Experiment 1 did not support the notion that a shorter period between learning and sleeping would lead to better EM performance (Gais et al., 2006; Payne et al., 2012; Scullin, 2014). One of the reasons for the lack of this effect was given as the participant tiredness levels. Another reason might have been, as discussed earlier, the short period of time (24h) for the effect to emerge. As discussed in the previous chapter, studies have shown a high drop in memory retention during the first week after encoding which then turns to a more steady linear decrease (Furman et al., 2012; Moreton & Ward, 2010). As such, the previous experiment argued that more time is needed to observe the effects such as the Remember-Know shift or the effect of AM/PM testing. It is important to point out that this explanation goes against the earlier provided research regarding the shorter period of time between learning and sleep. The earlier mentioned studies (Gais et al., 2006; Payne et al., 2012; Scullin, 2014) used stimuli that were word pairs and not, as the thesis argues, naturalistic experiences as used in the previous experiment. As discussed in Chapter 1 (section 1.8), data collection methods in EM experiments need to be as closely aligned with real-life tasks as possible and the tasks need to reflect and predict real-world tasks (Chaytor & Schmitter-Edgecombe, 2003; Ready et al., 2001; Silver, 2000). Experiment 2 continued exploring how the time of sleep post-learning affects EM for naturalistic events.

Taking into consideration results from Experiment 1 and the discussed studies it was predicted that a similar trend would be observed in the present experiment. It was predicted that the free recall, WWW (combined and separate) and Recognition scores would decrease over the 30d period. While the memory reduction after 24h should remain identical to the one observed in Experiment 1, the 24h – 7d period is not entirely clear. The prediction was that the memory performance in the present experiment should follow a Baseline > 24h > 7d > 30d pattern (Furman et al., 2012; Sekeres et al., 2016; Tunney, 2010; Tunney & Bezzina, 2007). Whereas the mentioned studies directly support the WWW and Recognition predictions, the free recall prediction can also be extrapolated from this, as in the experiments presented in this thesis, it is being used to count objects and details related to them. Thus being closely related to the What component and Detail recalls with the only difference being that the participants have to use internally generated cues to help them recall.

Some support, although again with stimuli lacking ecological validity, of a longer sleep-dependent consolidation process, can be found in studies looking into visual texture discrimination (Gais et al., 2000; Stickgold, James, et al., 2000; Stickgold, Whidbee, et al., 2000). These studies explored how sleep deprivation affected the identification of the orientation of an array of diagonal bars against a background grid of horizontal bars. As expected, a usual positive effect of sleep was observed with the discrimination positively correlating with the overnight sleep. What is important for the present experiment is that in one of the studies, no improvement was observed throughout the second-day post-learning, however, further nights of sleep did produce improvement (Stickgold, James, et al., 2000). This provides some support to the notion that memory enhancement can continue for at least 48-96 hours. In addition to a general replication of the previous experiment, the present experiment was able to investigate if the effect of the time period between learning and sleeping on EM became more pronounced after more than 24h. Following Experiment 1, the AM/PM testing remained in the present experiment, however, due to the just discussed literature the prediction was that the effect of the time period between learning and sleeping will become more visible at the 7d test (Group x Session interaction).

The present experiment also continued to explore the relationship across the time spent sleeping, sleep architecture and EM. As discussed in the previous chapters, sleep architecture affects EM consolidation (Daurat et al., 2007; Genzel et al., 2014; Plihal & Born, 1997; Scullin, 2014). While some research shows the importance of time spent in SWS for EM consolidation (Peigneux et al., 2004; Scullin, 2014), there is also a wide body of literature regarding the positive effect of REM sleep (Groch et al., 2013; Louie & Wilson, 2001; Rauchs et al., 2004). The sleep data from the previous experiment did not support this notion and showed no effect of time spent in REM or SWS sleep stages on EM performance. The present experiment tested the initial hypotheses again but this time tracking the sleep data over a longer period of time and with more reliable actigraphy bracelets (Cook et al., 2018, 2019). As previously the prediction was that time spent in SWS would positively correlate with EM performance.

While the general effect of time on EM consolidation has been explored by a number of mentioned studies, the effect of events on EM recall and memory consolidation, as used in the present experiments, is not as clear. The event stimuli used in Experiment 1 intended to explore how EM for events differed to EM for non-events. Experiment 1 showed that events were better remembered than non-events across almost all of the outcomes measures, over the 24h period. The present experiment explored how EM for events and non-events would change over a longer, 30 day period. As seen in the combined WWW proportion data in Experiment 1, no difference was found between the events and non-events at baseline but higher correctly recalled combined WWW proportions for events after 24h. A similar effect has been found in a study by Hamann et al. (1999), which explored how pleasant and aversive stimuli affected episodic recognition memory, whereby no difference in free recall and recognition between interesting and neutral stimuli after 10 minutes but better performance in both of the tasks for the interesting stimuli after four weeks. Findings with emotional stimuli show that amygdala activity during encoding is strongly correlated with memory performance after a long delay of 1-2 weeks versus immediately (Mackiewicz et al., 2006; Mickley Steinmetz et al., 2012; Ritchey et al., 2008). These findings show how additional testing time points can add to the general understanding of time-dependent consolidation and that consolidation of information can be affected by the nature of that information.

In the present experiment, the addition of the 7d and 30d time points allowed to explore how EM for events, compared to non-events, changed over longer periods. Using the data from Experiment 1 and the discussed studies, it was predicted that a similar trend would be observed at the baseline and 24h sessions with no difference between the events and non-events at the baseline with memory for events being better at the 24h session. This was also undertaken to see if the previous data would replicate adding to the robustness of the findings. The prediction for the new 7d and 30d sessions was that the events would be better remembered at both points in a similar manner to the 24h session.

Experiment 2 also explored an aspect of EM that was only touched on in Experiment 1 – the remember/know judgements and the Remember-Know shift. The R-K shift refers to the reduction of the Remember judgements and an increase in the Know judgements due to the semantisation and gist extraction of EM (Cermak, 1972;

Dewhurst et al., 2009; Herbert & Burt, 2001, 2004). This effect can be related to the memory trace theories and the sleep dependant memory consolidation. The replay and reactivation of memory traces lead to the gist extraction and make memories less dependent on the hippocampus. This is due to reduced 'binding' of the various EM components required to retrieve the main information about the episode (Ben-Yakov & Dudai, 2011; Kessels et al., 2007).

In a number of studies (Conway et al., 1997; Dewhurst et al., 2009; Harand et al., 2012; Herbert & Burt, 2003), this shift has been observed after four weeks which should be visible at the 30d session in the present experiment. In Experiment 1, the R/K judgements were only given for the object details and events and the Session x Judgement interaction was only observed for the detail judgements. As discussed, this was one of the predicted outcomes as 24h might not be enough for an observable R/K shift. The present experiment was better able to explore the R/K shift due to the additional 7d and 30d tests. In a study by Harand et al., (2012), participants had to rate valance of a series of emotional (positive or negative) and neutral pictures and memorise them. As expected, the authors found that there were fewer Remember responses after a three month period than after a three day period. What is interesting is that there were more Remember responses than Know responses even after a three day period. Participants in the present experiment had to give the R/K judgements for all of the WWW components in addition to the detail and event recalls (as in Experiment 1). This was done to test if the R/K shift would become visible over the 30d period. The general prediction was that the number of Remember judgements would reduce and the number of Know judgements would increase over the 30d period.

Lastly, the present experiment aimed to explore how the number of different EM measures used so far relate to each other. As discussed in Chapter 1 (Section 1.4.7), research shows that not all of the EM tests relate to one another (Cheke & Clayton, 2013, 2015). This suggests a contribution from multiple psychological processes and that not all of these tests necessarily test the same thing. One of the reasons for having so many measures in the experiments so far was to see what they add to the understanding of EM. Due to the use of HMD-VR and the use of event and non-event objects, a combination which so far has not been utilised in EM research, no specific predictions were made regarding the relationships between the measures.

In general, the main aim of the present experiment was to investigate how EM for events and non-events could change over a course of 30 days. This was undertaken by repeating the same procedure as in Experiment 1 but adding additional 7d and 30d testing sessions. Secondary aim followed from Experiment 1: to explore the effect of sleep dependant consolidation and the effect of AM/PM testing on the EM. Lastly, the experiment aimed to explore the different measures of EM and their validity.

4.2. Method

4.2.1. Participants

Participants in this experiment were 30 students from the Bishop Grosseteste University and members of the general public that did not participate in the previous experiment (mean age = 23.07; range = 18 - 40; female = 21). Undergraduate participants took part to obtain course credit; everyone else contributed freely. Participants were semi-randomly assigned to the two groups (n = 15): AM (mean age = 24.80; range = 19 - 40; female = 11) or PM (mean age = 21.33; range = 18 - 30; female = 10). As in Experiment 1, due to the difficulty in recruiting participants, the initial group assignment was conducted randomly (n total = 17, n in AM group = 9) with the rest of the participants being recruited on the basis of their personal availability. There were 16 students and 14 non-students. All participants had normal or corrected to normal vision. The participant screening procedure was identical to the one used in Chapter 2 (section 1.1).

4.2.2. Materials

The virtual environments (VEs) used in this experiment were the same as in Experiment 1 with the VR exploration procedure also staying identical. The general overview of the materials and procedure can be seen in Chapter 2 (sections 2.3 and 2.4).

The testing procedure remained similar to that of Experiment 1 with one major addition: participants had additional tests after 7 and 30 days. The memory tests were recreated in Unity3d game engine (Unity, 2017) and made to be available online so that the participants could carry out the 24h, 7d and 30d tests from home. Before starting each testing session, participants had to provide some additional information: the time they went to sleep the night before, the time they woke up, the current level of tiredness (on a scale from 1 – not tired at all to 10 – extremely tired) and if they had any caffeine before the test. The time of sleep and waking up was added to explore if a self-reported time spent asleep correlate with any of the EM measures. The tiredness level and the caffeine consumption was added as an exploratory measure due to tiredness being one of the discussed factors for the lack of effect of Group (AM/PM) in the previous experiment.

The previously used JawboneUp3 actigraphy bracelets were replaced with Fitbit Alta bracelets. This was done due to the new research showing their lack of accuracy at identifying sleep stages (Cook et al., 2019). As only four bracelets were available at any one time, data from four participants at a time were recorded resulting in actigraphy data from a total of 8 participants (mean age = 25.63; range = 19 - 36; female = 5).

As in Experiment 1, participants received half of all the target objects (36) at the baseline test and half of the target objects (36) at the 24h test. During the 7-day and 30-day tests, participants were given all 72 objects. As explained in Chapter 2 (section 2.5), half of the given objects in each session were lures and were not part of the explored VEs.

4.3. Design

The present experimental design was very similar to the one in Experiment 1. The experiment contained three tests (free recall, WWW and object recognition). Every test was performed at four time points: immediately after VE exploration (Baseline session), after 24h (24h session), after 7 days from the initial exploration (7d) and after 30 days from the initial exploration (30d). The exact time of testing on each of the testing days depended on the participants' group: AM – tests done at 9AM, PM - tests performed at 9PM each day. The Session (Baseline/24h/7d/30d) was the within-subject while Group (AM/PM) was the between-subjects independent variables.

There were five dependent variables in the free recall test: the number of recalled event objects, the number of recalled non-event objects, the number of recalled event object details, the number of recalled non-event object details and the total word count of the provided text.

The WWW task resulted in six dependant variables: the What, the When, the Where, the Event and combined WWW proportions and the average number of perceptual details recalled per one recalled (What) object. The Event component was a proportion of correctly recalled event associated with a particular event object.

After each Event and Detail recall participants had to provide a Remember/Know/Guess judgements regarding that information. This resulted in six dependant variables: proportions of Remember, Know and Guess judgements for the Event component and proportions of Remember, Know and Guess judgements for the Detail component.

The object recognition task resulted in two dependant variables: the d' sensitivity index and the confidence rating.

The actigraphy bracelet provided three dependant variables: total time spent asleep, time spent in REM and time spent in SWS.

4.4. Procedure

Participants were asked to come to the laboratory at either 9AM or 9PM depending on the group they were asked to be assigned to (AM or PM respectively). Participants were naïve to the study aims and were told that the study was about exploring virtual environments using virtual reality. The general procedure that followed is described in Chapter 2 (section 1.4). After the free recall, WWW and object recognition tasks, participants were given Fitbit Alta trackers which they were asked to wear every night for the next 30 days. Participants were then given a website address to perform the memory tests the next day (24h session). Participants were then able to leave the laboratory and carry out their normal daily activities. Depending on the group participants were assigned to (AM or PM), participants were asked to complete the memory tests at either 9AM or 9PM on each of the testing days.

4.5. Results

4.5.1. Free recall

In the free recall task, participants were given space to write freely about what they had experienced in the main VEs. This was done in a form of telling a story to a friend about what the participant experienced in the VEs. Free recall tests were scored in terms of the number of mentions of objects and object details. For example, “I remember seeing a radio next to a red mug and also a grey phone” would be marked as two non-event objects (radio and mug), one event object (phone), one non-event object detail (red mug) and one event object detail (grey phone).

4.5.1.1. The number of recalled objects

The number of recalled event and non-event objects was compared at each of the four testing sessions. An RM ANOVA with Session (Baseline, 24h, 7d and 30d) and object type (Event, Non-event) as within-subject variables and Group (AM, PM) as a between-subject variable was performed. The means and standard deviations are presented in Table 4.1.

Table 4.1.

Experiment 2: Mean number of recalled objects in the free recall test.

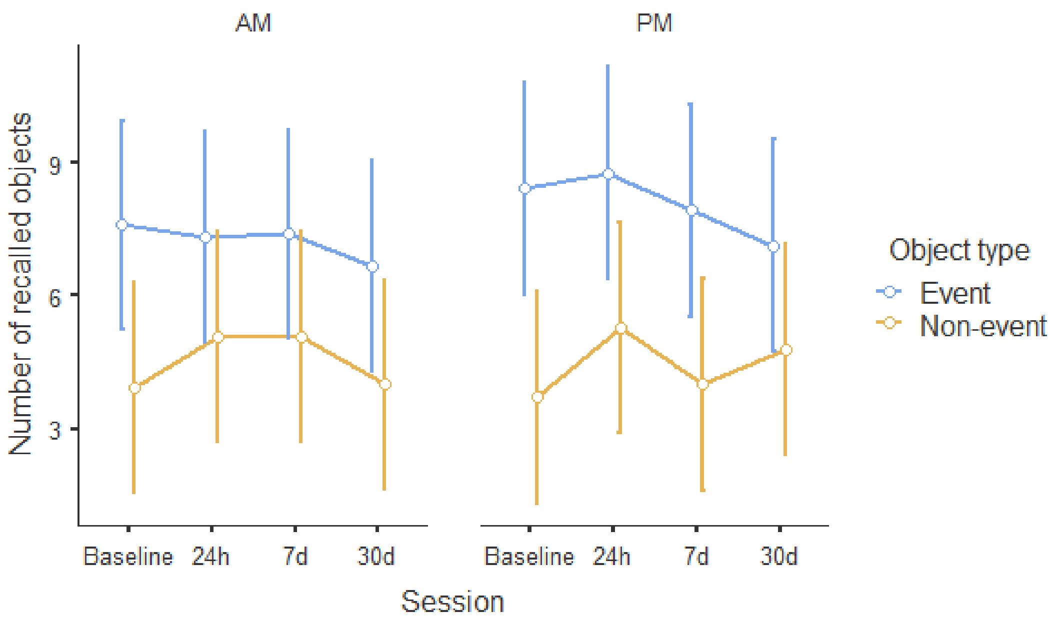
		Object type	
Group	Session	Event	Non-event
AM	Baseline	7.60 (3.94)	3.93 (3.94)
	24h	7.33 (4.51)	5.06 (4.51)
	7d	7.4 (4.65)	5.06 (4.65)
	30d	6.66 (4.68)	4.00 (4.68)
PM	Baseline	8.43 (4.85)	3.71 (3.46)
	24h	8.78 (5.07)	5.28 (4.28)
	7d	7.93 (4.73)	4.00 (3.39)
	30d	7.14 (5.27)	4.78 (3.87)

Note. Numbers in parentheses represent standard deviations

An effect of Object type was found, *F*(1,27)=41.24, *p*<.001, with Event objects being recalled better than non-event. The effect of Session was not significant, *F*(3,81)=1.38, *p*=.256, showing no difference in the number of recalled objects in all four testing sessions. The effect of Group was not significant, *F*(1,27)=.07, *p*=.797, showing no difference in the number of recalled objects between participants that were tested at 9AM or 9PM.

A significant Session x Object type interaction was observed, *F*(3,81)=3.75, *p*=.014 (see Figure 4.1). Only Object type within a session (e.g. Event object vs Non-event object at Baseline) and Object type across the sessions (e.g. Event object at Baseline vs Event object at 24h) comparisons were of interest. There were more Event objects recalled at each testing session than Non-event objects (*ts*>4.53, *ps*<.001). There was no difference in a number of recalled objects over the four testing sessions for any of the Object type (*ts*<2.19, *ps*>.693). None of the other interactions were significant, *Fs*<.75, *ps*>.479.

Figure 4.1.
Experiment 2: Mean number of recalled objects in the free recall test.



Note. Error bars represent 95% confidence intervals.

4.5.1.2. The number of recalled object details

The same analyses were performed for the number of recalled object details (Event object details, Non-event object details). The means and standard deviations are presented in Table 4.2.

Table 4.2.

Experiment 2: Mean number of details recalled in the free recall test.

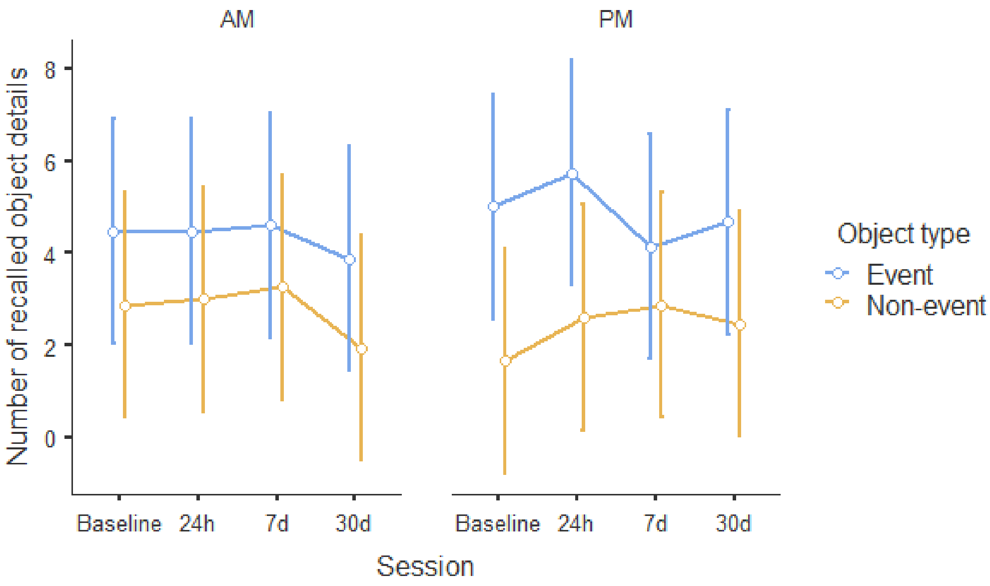
		Object type	
Group	**Session**	Event	Non-event
AM	Baseline	4.47 (5.72)	2.87 (3.93)
	24h	4.47 (4.76)	3.00 (3.07)
	7d	4.60 (5.32)	3.27 (3.69)
	30d	3.87 (5.60)	1.93 (3.01)
PM	Baseline	5.00 (6.05)	1.67 (2.19)
	24h	5.73 (6.83)	2.60 (2.61)
	7d	4.13 (6.71)	2.87 (3.31)
	30d	4.67 (5.89)	2.47 (2.29)

Note. Numbers in parentheses represent standard deviations

An effect of Object type was found, $F(1,28)=10.11$, $p=.004$, with overall more Event object details being recalled than Non-event (see Figure 4.2). The effect of Group was not significant, $F(1,28)=.003$, $p=.955$. None of the other effects or interactions were significant, $Fs<1.55$, $ps>.290$.

Figure 4.2.

Experiment 2: Mean number of recalled object details in the free recall test.



Notes. Error bars represent 95% confidence intervals.

4.5.1.3. Overall word count

An overall word count produced during the task was analysed using RM ANOVA with Session (Baseline, 24h, 7d and 30d) being the within-participant and Group (AM/PM) between-participant variables. The means and standard deviations are presented in Table 4.3.

Table 4.3.

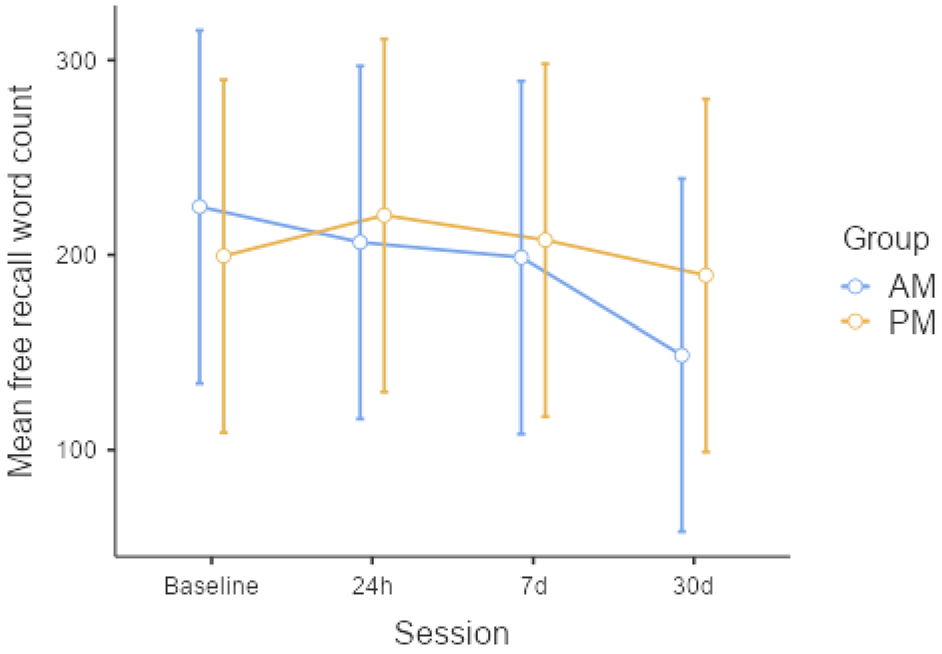
Experiment 2: Mean number of words written in the free recall test.

Group	Session	Mean	SD
AM	Baseline	196	198
	24h	202	187
	7d	210	210
	30d	171	196
PM	Baseline	228	129
	24h	225	163
	7d	197	146
	30d	167	134

An effect of Session was found, $F(3,84)=2.94$, $p=.038$ but after performing Bonferroni corrected pairwise comparisons, none of the analyses were significant, $ts<2.59$, $ps>.068$. None of the other effects or interactions were significant, $Fs<1.26$, $ps>.292$ (see Figure 4.3).

Figure 4.3.

Experiment 2: Mean number of words written in the free recall test.



Notes. Error bars represent 95% confidence intervals.

4.5.2. Combined What-When-Where components

When a participant correctly recalled the object (What), the room number the object was in (When) and where in the room the object was (Where) it was said that the participant recalled the full WWW information regarding that object. The combined WWW score is the proportion of correctly recalled WWW information out of all given objects (ranges from 0 to 1). For example, a score of 0.5 would mean that a participant recalled combined WWW information for 18 objects out of 36 possible targets. The means and standard deviations are presented in Table 4.4.

Table 4.4.

Experiment 2: Mean proportions of combined What-When-Where recalls.

		Object type	
Group	**Session**	Event	Non-event
AM	Baseline	.68 (.25)	.63 (.25)
	24h	.68 (.31)	.58 (.28)
	7d	.66 (.27)	.56 (.26)
	30d	.69 (.25)	.51 (.25)
PM	Baseline	.74 (.26)	.56 (.26)
	24h	.63 (.23)	.46 (.19)
	7d	.65 (.23)	.57 (.23)
	30d	.63 (.24)	.51 (.21)

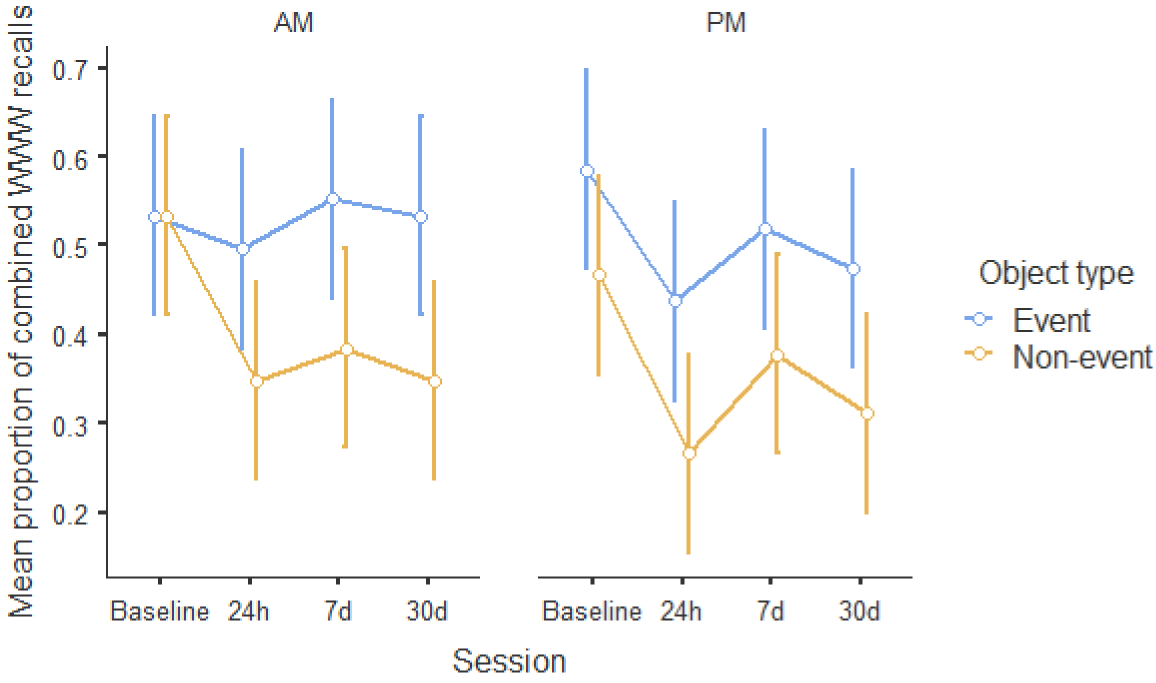
Note. Numbers in parentheses represent standard deviations

A significant effect of the object type was observed, $F(3,84)=30.41$, $p<.001$, with Event objects having higher recall compared to the Non-event objects, $t(28)=5.51$, $p<.001$. A significant effect of Session was observed, $F(3,84)=8.55$, $p<.001$, with higher proportion correctly recalled at Baseline than after 24h, $t(84)=4.77$, $p<.001$, and 30d, $t(84)=3.78$, $p=.002$, but not after 7d, $t(84)=2.38$, $p=.116$. No other comparisons were significant, *ps*>.116. The effect of Group was not statistically significant, $F(1,28)=.38$, $p=.544$.

A significant Session x Object type interaction was detected, $F(3,84)=2.78$, $p=.046$, however after the Greenhouse-Geisser correction the statistical significance disappeared, F(2.13,59.60)=2.78, $p=.067$ (see Figure 4.4). This was the only analysis that required this correction. None of the other interactions were significant, *Fs*< .72, *ps*> .326.

Figure 4.4.

Experiment 2: Mean proportions of combined What-When-Where recalls.



Notes. The combined WWW score is the proportion of correctly recalled What-When-Where information for all of the possible objects (ranges from 0 to 1). Error bars represent 95% confidence intervals.

4.5.3. Separate What-Where-When components

4.5.3.1. What

The What component represents a recall of an object. On the task screen, it is worded as "Do you recall X?" where X was a name of an object. Similarly to the combined WWW, the What component was measured as a proportion of correctly recalled objects. The means and standard deviations are presented in Table 4.5.

Table 4.5.

Experiment 2: Mean proportions of What (object) recalls

		Object type	
Group	**Session**	Event	Non-event
AM	Baseline	.80 (.12)	.84 (.12)
	24h	.71 (.17)	.56 (.21)
	7d	.83 (.12)	.68 (.14)
	30d	.78 (.12)	.67 (.17)
PM	Baseline	.76 (.17)	.79 (.20)
	24h	.70 (.13)	.59 (.16)
	7d	.80 (.15)	.68 (.21)
	30d	.75 (.18)	.64 (.23)

Note. Numbers in parentheses represent standard deviations

A significant effect of the Object type was observed, $F(1,28)=16.93$, $p<.001$, with Event objects having higher correct proportions compared to the non-Event objects, $t(28)=4.11$, $p<.001$. A significant effect of Session was observed, $F(3,84)=9.81$, $p<.001$, with higher proportions at Baseline than after 24h, $t(84)=5.25$, $p<.001$, and 30d, $t(84)=2.90$, $p=.029$, but not after 7d, $t(84)=1.62$, $p=.658$. Proportions at the 24h test were significantly lower than at the 7d test, $t(84)=-3.63$, $p=.003$, but did not differ to 30d, $t(84)=-2.35$, $p=.127$. Proportions between 7d and 30d tests did not differ, $t(84)=1.28$, $p=1$. An effect of Group was not found, $F(1,28)=.34$, $p=.565$.

A significant Session x Object type interaction was detected (see Figure 4.5), $F(3,84)=9.79$, $p<.001$. Multiple comparisons revealed no Object type difference at Baseline, $t(91.9)=-1.20$, $p=1$, but statistically significant difference at the remaining tests; 24h - $t(91.9)=4.33$, $p<.001$, 7d - $t(91.9)=4.39$, $p<.001$, 30d - $t(91.9)=3.55$, $p=.017$. Event object What recalls did not differ over the four sessions, $ts<2.08$, $ps=1$. Non-event object What recalls did not differ between Baseline and 7d, $t(139.3)=6.87$ $p<.001$, and Baseline and 30d, $t(91.9)=4.53$, $p<.001$, sessions. None of the other interactions were significant, $Fs<.23$, $ps>.757$.

4.5.3.2. When

The When component represents a correct recall of a room number in which the object from the What task was seen. On the task screen, it is worded as “In which room you have seen X?" where X is the name of an object from the previous task. Similarly to the combined WWW, the When component was measured as a proportion of correct room recalls out of all possible. The means and standard deviations are presented in Table 4.6.

Table 4.6.

Experiment 2: Mean proportions of When (temporal) recalls

		Object type	
Group	**Session**	Event	Non-event
AM	Baseline	.58 (.21)	.67 (.24)
	24h	.59 (.21)	.44 (.24)
	7d	.62 (.24)	.48 (.23)
	30d	.60 (.18)	.49 (.22)
PM	Baseline	.65 (.24)	.59 (.27)
	24h	.53 (.18)	.41 (.13)
	7d	.64 (.21)	.52 (.21)
	30d	.60 (.24)	.45 (.21)

Note. Numbers in parentheses represent standard deviations

A significant effect of Session was observed, $F(3,84)=5.66$, $p=.001$, with higher proportions at Baseline than after 24h, $t(84)=4.00$, $p<.001$, but not after 7d, $t(84)=1.72$, $p=.540$, or 30d, $t(84)=2.66$, $p=.056$. No other comparisons were significant, *ps*>.145.

A significant effect of the object type was observed, $F(1,28)=14.75$, $p<.001$, with Event objects having higher correct proportions compared to the non-Event objects, $t(28)=3.84$, $p<.001$. An effect of Group was not found, $F(1,28)=.32$, $p=.860$.

A significant Session x Object type interaction was detected (see Figure 4.4), $F(3,84)=9.79$, $p<.001$. Multiple comparisons revealed no Object type difference at Baseline, $t(86.6)=-.31$, $p=1$, but statistically significant difference at the remaining tests; 24h - $t(86.6)=3.58$, $p=.016$, 7d - $t(86.6)=3.79$, p=008, 30d - $t(86.6)=3.63$, $p=.013$. Event object proportions did not differed across the four sessions, *ts*<1.53, *ps*=1. Non-event object proportions were higher at Baseline than after 24h, $t(144.9)=5.17$, $p<.001$, 7d, $t(144.9)=3.35$, $p=.029$, and 30d, $t(144.9)=4.07$, $p=.002$. None of the other interactions were significant, *Fs*<1.89, *ps*>.480.

4.5.3.3. Where

The Where component represented correctly recalling an object's location in the virtual room. On the task screen, participants had to use a mouse and point on a top-down map of the room where they thought the object was located. This provided a distance – how far away the participant's guess was from the object's real location. Using a method explained in Chapter 2 (section 2.7) the distance was converted to a binary correct/incorrect outcome. Similarly to the combined WWW, the Where component was measured as a proportion of correct location recalls out of all possible. A high correlation was observed (as a measure of validity) between the Where pointing errors and the Where proportions (r = -.928, n = 240, *p*<.001). The means and standard deviations are presented in Table 4.7.

Table 4.7.

Experiment 2: Mean proportions of Where (spatial) recalls.

		Object type	
Group	**Session**	Event	Non-event
AM	Baseline	.71 (.20)	.64 (.17)
	24h	.58 (.11)	.41 (.13)
	7d	.69 (.22)	.50 (.19)
	30d	.68 (.24)	.46 (.15)
PM	Baseline	.66 (.22)	.61 (.17)
	24h	.59 (.24)	.39 (.17)
	7d	.62 (.19)	.47 (.21)
	30d	.59 (.18)	.48 (.17)

Note. Numbers in parentheses represent standard deviations

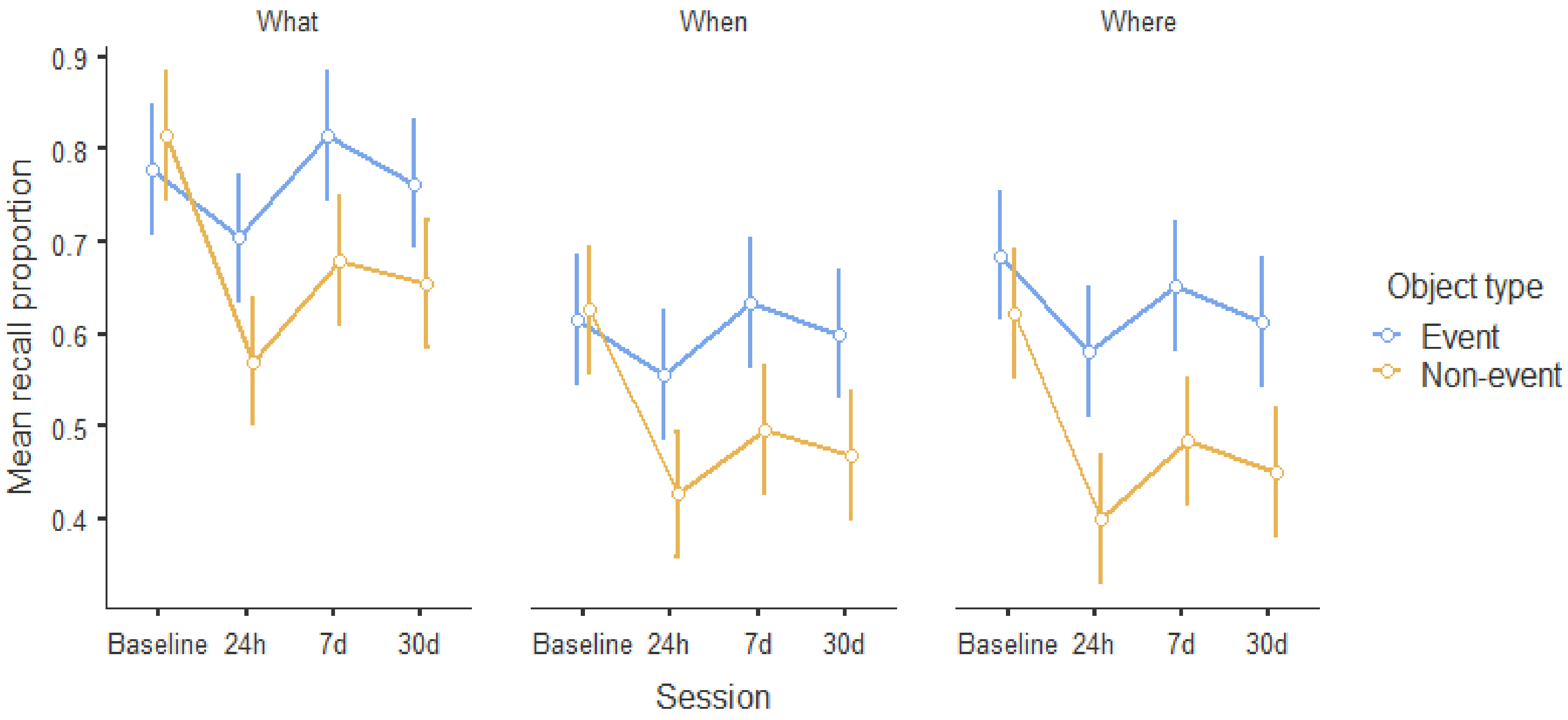
A significant effect of Session was observed, *F*(3,84)=10.32, *p*<.001, with higher proportions at Baseline than after 24h, *t*(84)=5.34, p<.001, 7d, *t*(84)=2.82, *p*=.036, and 30d, *t*(84)=4.00, *p*<.001. No other comparisons were significant, *ts*<1.18, *ps*>.083. A significant effect of the Object type was observed, *F*(1,28)=38.82, *p*<.001, with Event objects having higher proportions compared to the Non-event objects. An effect of Group was not found, *F*(1,28)=.54, *p*=.470.

A significant Session x Object type interaction was detected (see Figure 4.5), *F*(3,84)=3.16, *p*=.029. Multiple comparisons revealed no Object type difference at

Baseline, $t(95.5)=1.79$, $p=1$, but statistically significant difference at the remaining tests; 24h - $t(95.5)=5.15$, $p<.001$, 7d - $t(95.5)=4.78$, $p<001$, 30d - $t(95.5)=4.62$, $p<.001$. Event object proportions did not differed across the four sessions, $ts<2.77$, $ps>.178$. Non-event object proportions were higher at Baseline than after 24h, $t(151.7)=5.93$, $p<.001$, 7d, $t(144.9)=3.71$, $p=.008$, and 30d, $t(144.9)=4.60$, $p<.001$. None of the interactions were significant, $Fs<.22$, $ps>.766$.

Figure 4.5.

Experiment 2: Mean correct proportions recalled for the separate What-When-Where components.



Notes. To make the figure more readable and due to the lack of the effect, the data presented here was not split by the Group (AM/PM). Error bars represent 95% confidence intervals

4.5.3.4. Event

The Event component represents a proportion of correctly recalled events associated with the recalled event objects. On the task screen, participants were asked if any events happened to the recalled object (What). If they indicated that an event happened with the recalled object, they were asked to write it down. The written event descriptions were converted to binary (correct/incorrect) score and later into a proportion of correctly recalled events. The means and standard deviations are presented in Table 4.8.

Table 4.8.

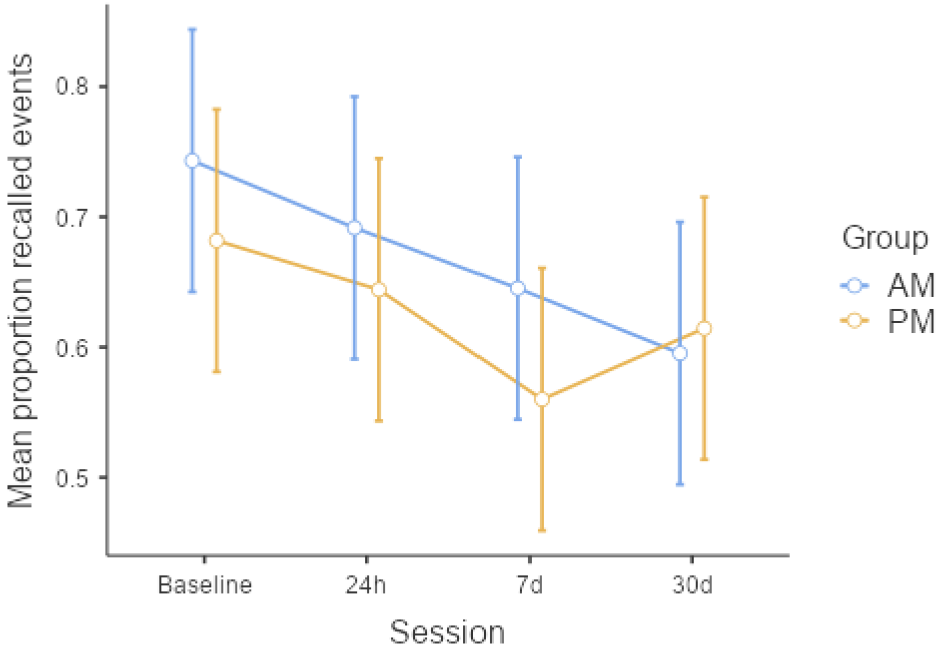
Experiment 2: Mean proportions of Event recalls.

Group	Session	Mean	SD
AM	Baseline	.74	.19
	24h	.69	.22
	7d	.65	.14
	30d	.60	.15
PM	Baseline	.68	.27
	24h	.64	.24
	7d	.56	.16
	30d	.61	.16

An effect of Session significant, $F(3,84)=3.13$, $p=.030$. After performing Bonferroni correction, none of the pairwise comparisons were significant, $ts<2.59$, $ps>.068$. An effect of Group was not significant, $F(1,28)=.79$, $p=.383$. The Session x Group interaction was not significant, $F(3,84)=.56$, $p=.646$ (see Figure 4.6).

Figure 4.6.

Experiment 2: Mean proportions of recalled events associated with the event-objects.



Notes. Error bars represent 95% confidence intervals.

4.5.3.5. Detail

The Detail component represented the mean number of recalled details for one object. On the task screen, participants were asked if they could recall any perceptual detail about an object and if they could write one down. Participants were able to write down up to five details per one object. The means and standard deviations are presented in Table 4.9.

Table 4.9.

Experiment 2: Mean number of object details recalled per one recalled object.

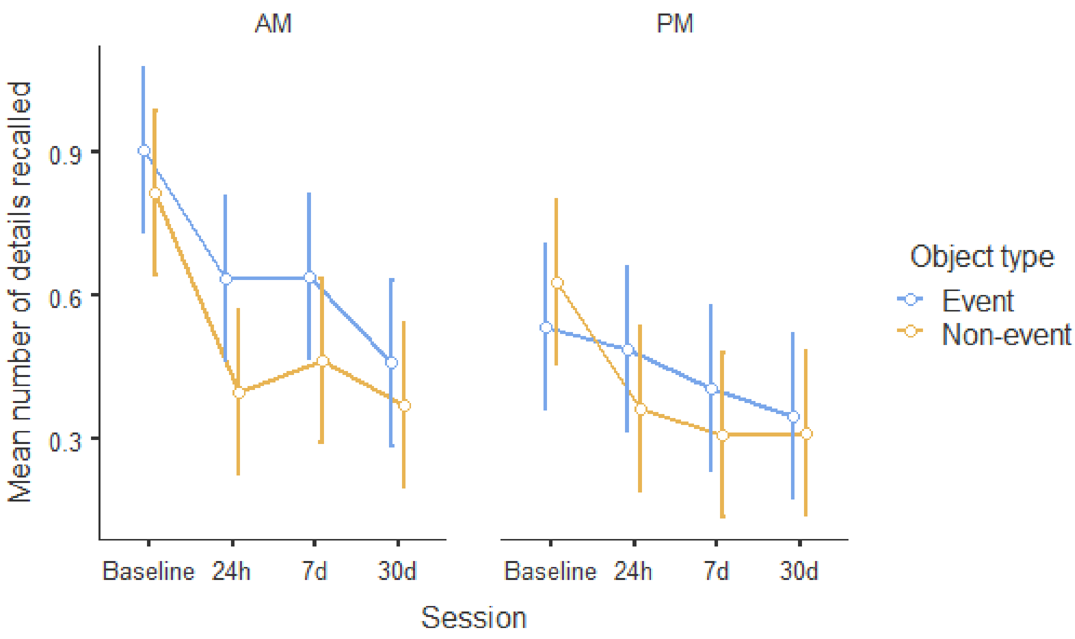
		Object type	
Group	**Session**	Event	Non-event
AM	Baseline	.90 (.49)	.82 (.38)
	24h	.68 (.42)	.40 (.47)
	7d	.64 (.39)	.47 (.25)
	30d	.46 (.59)	.37 (.37)
PM	Baseline	.53 (.46)	.63 (.26)
	24h	.49 (.44)	.36 (.33)
	7d	.41 (.33)	.31 (.27)
	30d	.35 (.47)	.32 (.29)

Note. Numbers in parentheses represent standard deviations

A significant effect of Session was observed, $F(3,84)=18.63$, $p<.001$, with more details recalled at Baseline than after 24h, $t(84)=5.06$, $p<.001$, 7d, $t(84)=5.38$, $p<.001$, and 30d, $t(84)=7.08$, $p<.001$ (see Figure 4.7). No other comparisons were significant, *ps*>.279. A significant effect of the Object type was observed, $F(1,28)=6.80$, $p=.014$, with more details recalled for Event objects compared to the Non-event objects. An effect of Group was not found, $F(1,28)=3.25$, p=.082. None of the interactions were significant, *Fs*<2.64, *ps*>.055.

Figure 4.7.

Experiment 2: Mean number of object details recalled per one recalled object.



Notes. Error bars represent 95% confidence intervals.

4.5.4. Remember/know/guess judgements

Separate repeated measure ANOVAs were used to analyse the remember/know/guess judgements for the What, When, Where, Event and Detail components. The judgements were transformed into overall proportions using a similar method to Dewhurst, Conway, & Brandt (2009). For example, adding one participant's Remember, Know and Guess judgement proportions at Baseline would equal 1. This transformation was undertaken so that the lower number of recalled objects on the second session would not affect the judgement data. The means and standard deviations of the R/K/G judgements are presented in Table 4.10.

Table 4.10.

Experiment 2: Mean proportions of the Remember/Know/Guess judgements given in the WWW task

		Object type	
Judgement	**Session**	Event	Non-event
Remember	Baseline	.63 (.30)	.61 (.25)
	24h	.63 (.32)	.57 (.30)
	7d	.61 (.31)	.56 (.27)
	30d	.65 (.30)	.60 (.29)
Know	Baseline	.27 (.28)	.26 (.24)
	24h	.27 (.26)	.25 (.26)
	7d	.29 (.29)	.31 (.25)
	30d	.25 (.28)	.27 (.26)
Guess	Baseline	.10 (.12)	.13 (.11)
	24h	.10 (.12)	.18 (.16)
	7d	.10 (.09)	.14 (.10)
	30d	.09 (.09)	.13 (.09)

Notes. Due to the lack of effect, Group (AM/PM) was not included in the table. Numbers in parentheses represent standard deviations

Remember

There was a significant effect of Object type, $F(1,28)=7.16$, $p=.012$, with Event objects having higher proportions of remember responses than Non-event objects (see Figure 4.8). The effect of Session was not significant, $F(3,84)=.73$, $p=.538$. The effect of Group was not significant, $F(1,28)=.17$, $p=.686$. No interactions were significant, $Fs<2.40$, $p>.074$.

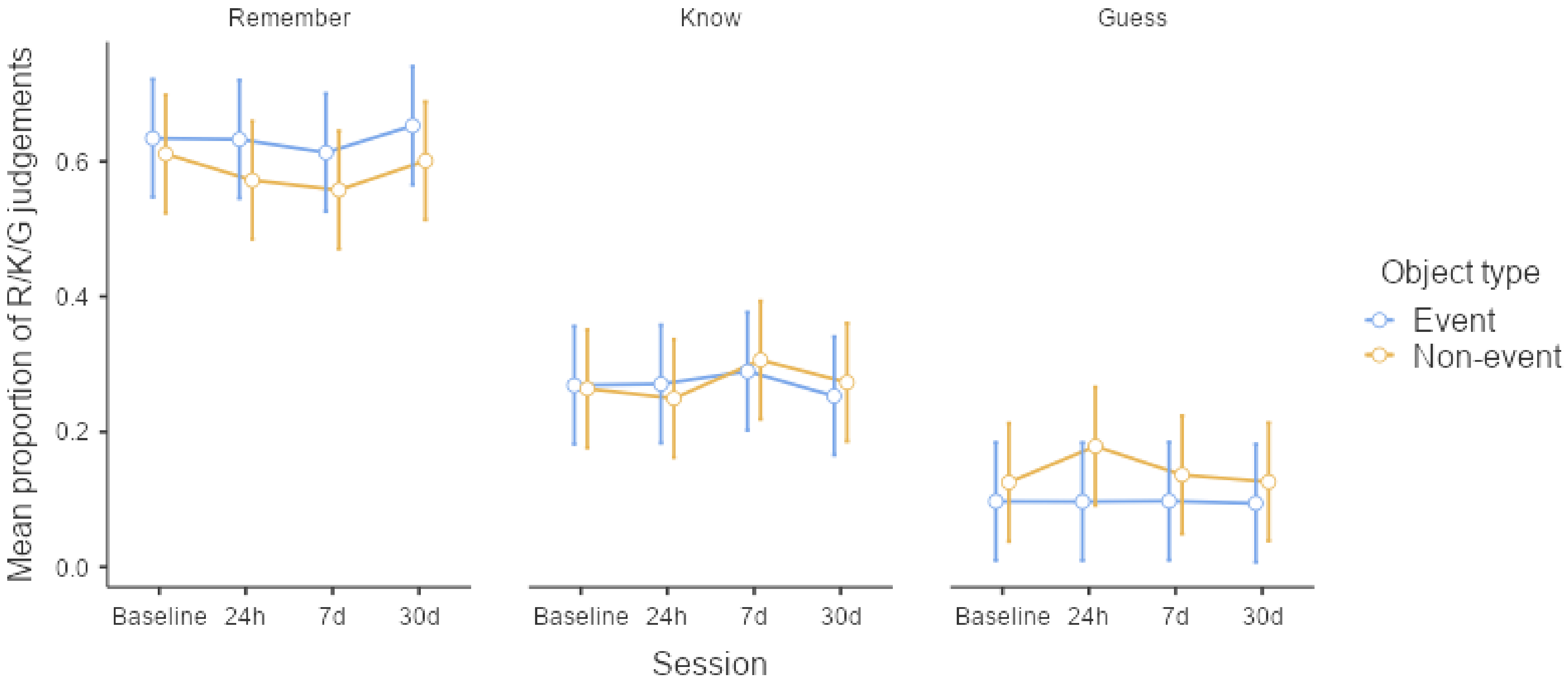
Know

The effect of Object type was not significant, $F(1,28)=.03$, $p=.874$. The effect of Session was not significant, $F(3,84)=.66$, $p=.577$. The effect of Group was not significant, $F(1,28)=.14$, $p=.710$. No interactions were significant, $Fs<1.39$, $p>.249$.

Guess

There was a significant effect of Object type, $F(1,28)=12.29$, $p=.002$, with Event objects having lower proportions than Non-event objects. The effect of Session was not significant, $F(3,84)=1.31$, $p=.276$. The effect of Group was not significant, $F(1,28)=.04$, $p=.852$. No interactions were significant, $Fs<1.73$, $p>.166$.

Figure 4.8.
Experiment 2: Mean total proportions for Remember, Know and Guess judgements provided in the WWW task.



Notes. Error bars represent 95% confidence intervals

4.5.5. Recognition – d' scores

Recognition data was converted to d' scores (Z(hit rate) − Z(false alarm rate)). The means and standard deviations are presented in Table 4.11.

Table 4.11.

Experiment 2: Mean d' object recognition scores.

		Object type	
Group	**Session**	Event	Non-event
AM	Baseline	2.89 (.45)	2.63 (.54)
	24h	2.58 (.66)	1.96 (.78)
	7d	2.47 (.58)	1.77 (.64)
	30d	2.37 (.70)	1.76 (.70)
PM	Baseline	2.89 (.50)	2.63 (.68)
	24h	2.30 (.71)	1.78 (.61)
	7d	2.48 (.68)	1.99 (.78)
	30d	2.19 (.60)	1.60 (.62)

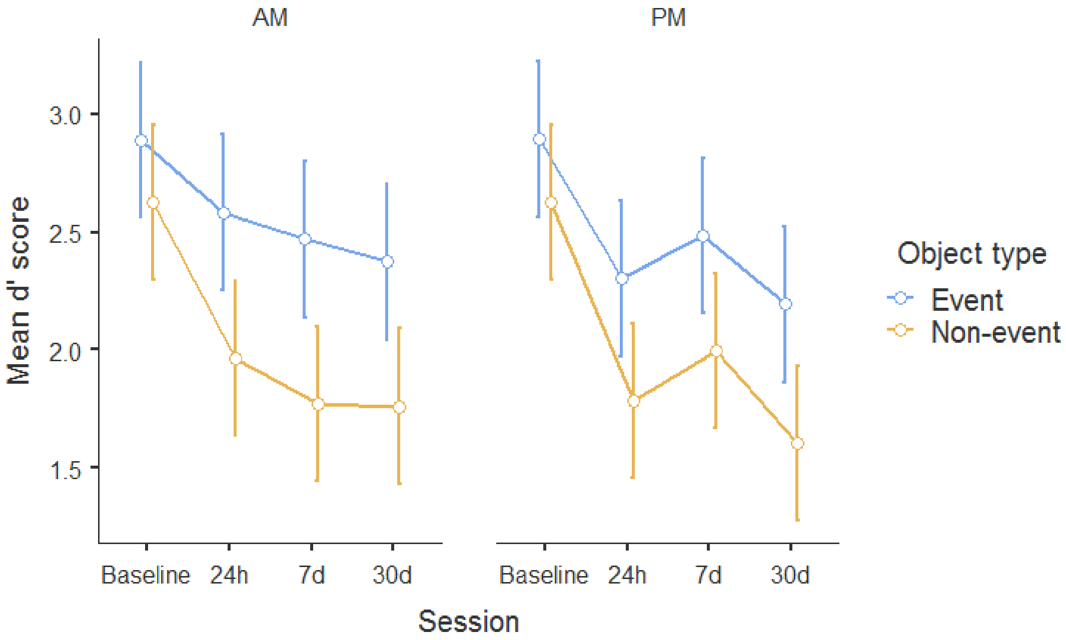
Note. Numbers in parentheses represent standard deviations

There was a significant effect of Session, $F(3,84)=14.20$, $p<.001$, with higher d' scores at Baseline than after 24h, $t(84)=4.74$, $p<.001$, 7d, $t(84)=4.57$, $p<.001$, and 30d, $t(84)=6.12$, $p<.001$. No other comparisons were significant, *ts*<1.56, *ps*>.741. There was a significant effect of Object type, $F(1,28)=59.02$, $p<.001$, with Event objects having higher d' scores than Non-event objects. The effect of Group was not significant, $F(1,28)=.22$, $p=.645$.

There was a significant Session x Object type interaction, $F(3,84)=4.24$, $p=.008$ (see Figure 4.9). The d' scores did not differ between the two object types at the Baseline, $t(86.5)=2.80$, $p=.177$, but were higher for the Event objects at the remaining three sessions, *ts*>6.01, *ps*<.001. Event object d' scores were higher at the baseline compared to 24h, $t(115)=3.24$, $p=.044$, and 30d, $t(115)=4.40$, $p<.001$, but not to 7d, $t(115)=3.00$, $p=.092$. Non-event object d' scores were higher at the baseline compared to 24h, *t*(115)=5.44, *p*<.001, 7d, *t*(115)=5.37, *p*<.001 and 30d, *t*(115)=6.82, *p*<.001, 7d, *t*(115)=3.00, *p*=.092. No other interactions were significant (*Fs*<.789, *p*>.503).

Figure 4.9.

Experiment 2: Mean d' object recognition scores.



Notes. Error bars represent 95% confidence intervals

4.5.6. Recognition – confidence ratings

The confidence ratings were ratings (from .0 to 1.0) reflecting how confident the participants felt about their recognition judgement. A confidence rating of .0 would indicate being not confident at all whereas confidence rating of 1.0 would indicate full confidence. The means and standard deviations are presented in Table 4.12.

Table 4.12.

Experiment 2: Mean object recognition confidence ratings.

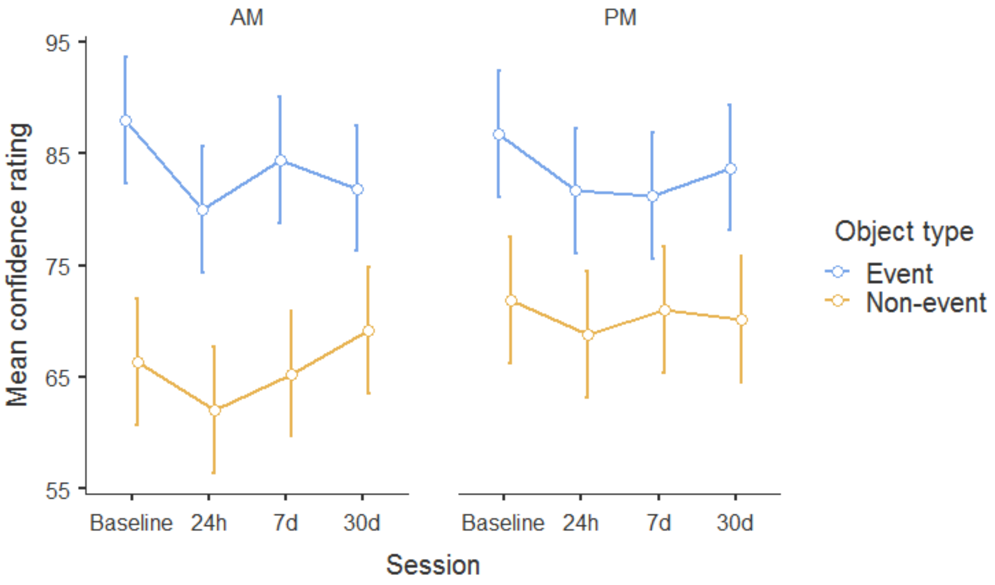
		Object type	
Group	**Session**	Event	Non-event
AM	Baseline	87.96 (31.15)	66.34 (29.70)
	24h	80.01 (22.25)	62.01 (24.05)
	7d	84.37 (32.00)	65.28 (25.34)
	30d	81.88 (28.22)	69.17 (26.80)
PM	Baseline	86.71 (28.78)	71.89 (39.19)
	24h	81.65 (24.48)	68.75 (26.94)
	7d	81.20 (36.95)	71.01 (27.76)
	30d	83.72 (25.55)	70.12 (24.79)

Note. Numbers in parentheses represent standard deviations

There was a significant effect of Object type, $F(1,28)=90.45$, $p<.001$, with Event objects having higher confidence ratings than non-Event objects. There was a significant effect of Session, $F(3,84)=5.48$, $p=.002$, with higher confidence ratings at Baseline than after 24h, $t(84)=4.01$, $p<.001$, but not after 7d, $t(84)=2.16$, $p=.203$, or 30d, $t(84)=1.57$, $p=.726$ (see Figure 4.10). No other comparisons were significant, *ts*<-.59, *ps*>.407. The effect of Group was not significant, $F(1,28)=.60$, $p=.445$. No interactions were significant, *Fs*<2.38, $p>.134$.

Figure 4.10.

Experiment 2: Mean object recognition confidence ratings for the two object types over the four testing sessions.



Notes. Error bars represent 95% confidence intervals

4.5.7. Actigraphy data

Due to software errors data from the actigraphy bracelets was lost. Every participant's data was overwritten by the data from the next participant that used the same bracelet. After extracting the remaining data it was found that in all cases there were a lot of missing data (numerous nights without any recorded sleep data). Due to this, it was decided that the actigraphy data would not be analysed.

4.5.8. The effect of tiredness and sleep time

One of the explanations for the lack of effect of Group (AM/PM) in Experiment 1 was tiredness. As an exploratory measure in the present experiment, participants were asked to indicate how tired they were (on a scale from 1 to 10). Additionally, participants were asked to indicate what time they went to sleep the day before and the time they woke up on the day of the test which resulted in a subjective time spent asleep. This was added as a precaution, in case of any technical difficulties with the actigraphy bracelets. The means and standard deviations are presented in Table 4.13.

Table 4.13.

Experiment 2: Mean total time spent asleep and tiredness level.

Group	Session	Time spent asleep (min)	Tiredness level
AM	Baseline	455 (61.90)	3.20 (2.31)
	24h	448 (103)	4.60 (2.26)
PM	Baseline	587 (174)	4.60 (2.41)
	24h	480 (73.90)	5.53 (2.70)

Note. Numbers in parentheses represent standard deviations

Both the subjective time spent asleep and tiredness levels were higher in the PM group than the AM group - $F(1,118)=15.10$, $p<.001$ and $F(1,118)=6.87$, $p=.010$, respectively.

To explore if there were any relationships between the level of tiredness and self-reported time spent asleep a Pearson's r correlation table was performed with the mentioned measures and the EM measures analysed in the previous sections. However, the correlations were only done with data from the Baseline and 24h sessions. This was done as only these two sessions were on two consecutive days.

The level of tiredness showed negative correlations with the WWW What component recalls, $r(120) = -.22$, $p=.016$, the mean number of details from the WWW task $r(120) = -.23$, $p=.013$ and confidence ratings from the object recognition task, $r(120) = -.23$, $p=.013$. Time spent asleep did not correlate with any of the EM measures, *ps*>.216 (see Table 4.14).

Table 4.14.

Experiment 2: Correlation table showing relationships between the subjective time spent asleep and tiredness level.

	Time spent asleep (min)	**Tiredness level**
Free recall: Objects	.06	.12
Free recall: Details	.05	-.00
Combined WWW	.02	-.13
WWW What component	.01	-.22*
WWW When component	.02	-.11
WWW Where component	-.01	-.17
WWW Details	-.04	-.23*
WWW Event	-.10	.10
Remember judgements	-.11	.06
Know judgements	.07	-.14
Guess judgements	.11	.16
d' index	.08	-.04
Confidence rating	.10	-.23*

Note. * p < .05, ** p < .01, *** p < .001

4.5.9. Comparison of measures

One of the aims of the present experiment, and the thesis as a whole, was to explore the different measures of EM and the relationships among them. As a result, a correlation matrix was created with all of the measures analysed above (see Table 4.15). The only measure that was not added was the total word count from the free recall task. This was due to the fact that it could not be split by the Object type (Event, non-event).

Table 4.15.

Experiment 2: Correlation matrix for the used episodic memory measures.

Measure	1	2	3	4	5	6	7	8	9	10	11	12
1. Free recall – Nr. of recalled objects	—											
2. Free recall – Nr. of recalled object details	.76***	—										
3. Combined WWW	.53***	.46***	—									
4. WWW -What component	.26***	.26***	.61***	—								
5. WWW -When component	.41***	.38***	.89***	.70***	—							
6. WWW -Where component	.49***	.41***	.84***	.74***	.67***	—						
7. WWW - Mean number of details recalled per one object	.41***	.50***	.33***	.48***	.33***	.44***	—					
8. Total proportion of Remember judgement	.14*	.16*	.24***	.18**	.20**	.29***	.24***	—				
9. Total proportion of Know judgement	-0.04	-0.04	-0.12	-0.09	-0.11	-.13*	-.16*	-.92***	—			
10. Total proportion of Guess judgement	-.27***	-.29***	-.33***	-.26***	-.26***	-.42***	-.25***	-.38***	-0.01	—		
11. D' sensitivity index	.29***	.16*	.46***	.41***	.39***	.52***	.13*	.14*	-0.06	-.21***	—	
12. Confidence rating	.32***	.32***	.33***	.36***	.30***	.42***	.27***	.18**	-0.01	-.45***	.41***	—

Notes. $^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$.

4.6. Discussion

The present chapter explored how EM for events and non-events changed over a course of 30 days. This was achieved by repeating the same procedure as in Experiment 1 but adding additional 7d and 30d testing sessions, to elongate the time course of EM retrieval. The experiment also aimed to explore the effect of sleep dependant consolidation and the effect of AM/PM testing on the EM. Lastly, the experiment followed the general aim of the thesis to explore the validity of different measures of EM.

The main aim of the present experiment was to investigate how EM for events and non-events change over a course of 30 days. The prediction was that there would not be any difference between the two object types at the Baseline but better recall for the event objects in the remaining sessions with a general trend of memory performance following a Baseline > 24h > 7d > 30d pattern.

As predicted, and following the results of Experiment 1, Event objects were recalled better than the Non-event objects in all three memory tests (free recall, WWW, and object recognition). In the combined WWW (see Figure 4.4) and object recognition data (see Figure 4.9), memory performance for Event objects did not differ at the Baseline, with the difference becoming visible after 24h and persisting throughout the remaining two (7d and 30d) testing sessions. The free recall results showed a difference between the two object types at all four testing sessions (see Figure 4.1), with events being consistently better retrieved than non-events. The present experiment again provided supporting data showing that EMs for experiences that are more event-like and as argued by the present thesis more life-like are better remembered when compared to EMs for static objects, the latter of which are regularly used in EM research. According to the present data, this effect remained over a 30d period. However, after looking at the recall trends, there was a lack of general forgetting between the 7d and 30d session in all of the tests. Interestingly, improved memory performance was observed in two of the measures, the separate What and Where component, during the 7d session as compared to the 24h session (see Figure 4.5). One explanation for these trends is the effect of retesting.

The lack of reduction in memory performance over the 7d and 30d sessions and the increase in the recall at the 7d session can be attributed to the effect of retesting. All of the stimuli at the 7d test were tested at least once at either Baseline or 24h conditions. Likewise, the 30d test was identical to the 7d test which means that only the 24h test might show true forgetting. Additionally, participants performed the free recall task during each testing session which meant that they were asked to recall all the objects regardless if those objects part of the random sample of objects given during the WWW or object recognitions tests. Furthermore, the object recognition task followed the WWW test in which participants were shown pictures of objects from the environments. This means that participants have seen how every object looked like before the 7d session. This may have resulted in participants relearning the information and thus slowing down the forgetting.

Research has shown that retesting improves memory (Baddeley et al., 2019; Karpicke & Roediger, 2008; Roediger & Butler, 2011; Soderstrom et al., 2016). Indeed, a recent and methodologically closely related study by Baddeley et al. (2019) using the Four Doors Test and the Crimes Test showed that when participants were tested after one day, one week and one month (the Four Doors Test) they showed significantly less or no forgetting at all, compared to when they were tested only immediately and after a month (the Crimes Test). Indeed their observed trend of forgetting was very similar to the combined WWW from the present experiment. Due to the use of the same material at the 7d and 30d tests, the experiment did not show the full effect of forgetting but unintentionally provided evidence for the effect of retesting on EM. This can also be seen in the d’ object recognition scores where they dropped after 24h but maintained the same level at both 7d and 30d tests. While, the memory trends observed in the present experiment support the hypothesis that repeated recall tests can help to maintain memory performance (Jansari et al., 2010; Karpicke & Roediger, 2008; Tulving, 1967) without a control condition that would only contain Baseline and 30d tests such conclusion is difficult to reach.

Another aim of the present experiment and the thesis as a whole was to explore the different measures of EM. This was the reason for the high number of measures used in the experiments so far. While Experiment 1 was more focused on the initial exploration of the VEs and the use of HMD-VR, the present experiment was more focused on the EM and its measures. As such a correlation matrix was created that included all of the measures used in the present experiment (see Table 4.15). The correlation data provided a number of interesting insights. First of all, the combined

WWW proportions positively correlated with almost all (excluding the Know and Guess judgement proportions) of the EM measures. This indicates that directly or indirectly all of those measures might relate to EM. However, free recall object data and the d' sensitivity index showed the highest correlations ($r = .53$ and $r = .46$ respectively). Taking into consideration that the present thesis argues that the combined WWW information is the main representation of an EM, the mentioned correlations provide additional validity to the two common measures of EM.

A number of interesting relationships can be seen between the combined WWW and the R/K/G judgements. The Remember judgements positively correlated with the combined WWW data which is expected as the Remember judgement is related to the autonoetic consciousness which is an important aspect of EM (Tulving, 2002, 2004; also see Klein, 2013). What is interesting is that statistically significant correlation was not found between the combined WWW data and the Know judgement but a statistically significant negative correlation was observed regarding the Guess judgements. This was a noteworthy finding as in the main R/K/G data there was no difference between the Know and Guess judgements. This leads to an important point. It shows that there is a relationship between the 'episodic-ness' of memory and guessing and this judgement is a better indication of the 'episodic-ness' than knowing. The use of Guess judgements in the R/K testing has been shown to 'purify' the Know judgements and leave them so that participants would not use them when they were just guessing (see Gardiner et al., 2002 for a review; but also see Migo et al., 2012).

Looking at the relationships between the separate WWW components it is visible that the What component shows a lower correlation with the combined WWW than the separate When and Where components. This difference in correlations between the WWW components shows that contextual information is important in creating a full EM. Using the memory trace theories (Moscovitch & Nadel, 1998; Yassa & Reagh, 2013) this could be compared to an activation of a traces related to contextual information having a higher chance of retrieving the whole episode than just the activation of a trace related to the item. It is important to point out that all the discussed relationships are based only on correlations and thus inferences need to be made with caution. Nevertheless, these relationships show the importance of using a number of measures while exploring EM as the different measures can provide different insights in EM (Cheke & Clayton, 2013, 2015).

It is important to look closer to the results of the separate tests. For example, the difference between events and non-events at Baseline in the free recall object data but not in the combined WWW data., which demonstrates that event objects were more likely to be recalled freely than non-event objects. When the object information was cued, as in the What part of the WWW task, the difference between events and non-events disappeared. This can be interpreted as a memory trace reactivation (Nadel et al., 2000; Yassa & Reagh, 2013). During the free recall, one needs to internally search their memory for information whereas, in a cued recall (as in a paired associates test, see Wilson et al., 1982), the cue can reactivate the memory trace resulting in better memory recall. While keeping in mind the observed effect of retesting, results such as these, show the importance of combining different memory tests to explore memory (Cheke & Clayton, 2013, 2015).

When looking at the WWW data, it is important to note the drop in the recallability of Non-event objects observed during the 24h session. This can be seen in both combined (even if the Object type x Session interaction was not significant) and separate component data. The same drop in the recall is not present for the Event objects, with around the same levels of recall observed in all sessions. This shows that regardless of the discussed retesting and cueing, Event objects were still better recalled over the 30d period with memory for Non-event objects dropping after 24h. This indicates that EM for life-like events is better retained and as such preferentially consolidated, over EM for static objects. This is further supported by the data from the Session x Object type interaction. While the interaction was not significant, a closer look at the descriptive statistics (Table 4.4), and the data from Experiment 1, demonstrates that Event objects were better recalled not due to enhanced encoding but due to enhanced consolidation over the initial 24h. It is important to point out that these data are difficult to explain using forgetting or memory interference as none of the measures showed a significant effect of Group (AM/PM).

An interesting insight into EM was provided by the free recall and WWW perceptual detail data. The free recall detail data showed an effect of the Object type, with more Event object detail being recalled, but no effect of Session, showing that on average the same amount of details was recalled at each of the four sessions (see Figure 4.2). The WWW detail data showed more Event object details being recalled than the non-event, and the number of object details being higher at the Baseline compared to the other three sessions (see Figure 4.7). Looking at both sets of data it is possible to say that while in general the details for the event objects were better recalled, the detail data did not follow the general trend observed in the other EM measures. As such the additional detail may have been superfluous to the core EM. Furthermore, the lack of Object type x Session interaction might indicate that perceptual object details might not be as important part of EM as, for example, the When and Where components.

The d' object recognition data showed almost identical trends of forgetting as in the other tests: no difference at the Baseline but higher scores at the remaining three sessions (see Figure 4.9). Interestingly, the d' scores were statistically lower (unlike the combined WWW scores) for both Event and Non-event objects at the 24h session and then remained on the same level for the remaining sessions. While it is possible to argue that the 24h drop is due to the fact that the objects at the 24h session were not tested at the baseline, this does not explain why such effect is not visible in the combined WWW data. An identical drop in memory at 24h session is only observed for the Non-event objects but not Event objects. It is important to mention a related study by Harand et al. (2012) in which they have found that d' scores did not change between three-day and three-month time-points. This shows that it is possible to maintain similar levels of memory over longer periods of time.

The confidence ratings given during the object recognition task provide another interesting insight into EM. The data shows, as in the other measures, that the ratings were higher at the baseline than at the other three sessions (see Figure 4.10). However, the main interest is in the lack of Object type x Session interaction. The confidence ratings were lower for the Non-event objects starting from the Baseline and throughout the whole experiment. The statistically significant difference between the event and non-event objects at the Baseline but lack of similar differences in the other measures (apart from the free recall object data) shows that both types of objects are encoded at a similar level even if the subjective confidence in memory differs. One explanation for this is based on a theory that confidence ratings represent the ease of access to the memories (Burke et al., 1991; Busey et al., 2000). This would indicate that non-event objects were more difficult to recall than event objects which would go against the earlier discussed hypothesis that both types of objects are encoded at the same level.

The Remember/Know results from the present experiment did not support the initial prediction that an R/K shift would be observed at the 7d or 30d session. The data showed that there were overall more Remember and fewer Guess judgements for the event objects (see Figure 4.8). However, the lack of effect of Session (or any related interactions) shows that the proportions did not differ throughout the whole experiment. This finding is quite difficult to explain as it indicates that both types of objects were remembered, with remembering being linked to episodic retrieval, throughout the whole month-long experiment. These results also go against the combined WWW trends, which arguably represent an EM, which showed a drop in recall after 24h. One possible explanation would be that the R/K judgements in the present experiment were not indicating episodic retrieval but were more related to the confidence of memory retrieval. By some researchers, the R/K procedure has been interpreted as a measure of confidence in or strength of recognition memory (Dunn, 2004; Migo et al., 2012; Wais et al., 2008). However, this explanation would go against the confidence rating data seen in the present experiment.

Looking back at all of the measures used in the present experiment it is imperative to note the complete lack of effect of Group (AM/PM). One of the aims of the experiment was to explore the effect of time of sleep following learning. The prediction was (and followed Experiment 1) that having a shorter period between learning and sleep (PM group) should lead to better EM performance. The current data did not provide any support for this prediction. This is quite an interesting result as there is a body of literature showing that such effect does exist (Benson & Feinberg, 1977; Gais et al., 2006; Payne et al., 2012; Scullin, 2014; Talamini et al., 2008). The lack of effect of Group shows that interference (Ellenbogen et al., 2006) as discussed in Chapter 1 (section 1.7.2), did not affect the consolidation of EM. The present results, in addition to results from Experiment 1, do not provide support to the hypothesis that less time between learning and sleep should lead to better EM performance. Due to the fact that two consecutive experiments failed to support this hypothesis, the AM/PM design will not be used in further experiments.

However, it is important to mention the data regarding the subjective time spent asleep and the levels of tiredness. The time spent asleep was explored as a precaution in case there would be problems with the actigraphy bracelets. The tiredness levels were added as an exploratory measure as one of the explanations for the lack of AM/PM differences given in Experiment 1 was circadian rhythms and tiredness (Baddeley et al., 1970; Barrett & Ekstrand, 1972; Folkard & Monk, 1980; May et al., 1993). Interestingly while PM group showed significantly more time spent it did not correlate with any of the EM measures. This finding in addition to the lack of correlations found in the actigraphy data from Experiment 1 (see Table 3.15) and the technical problems encountered in the present experiment (see Section 4.5.7) shows the need for similar experiments to employ polysomnography instead of actigraphy bracelets. Using polysomnography in HMD-VR study should reveal more accurate relationships between sleep and EM for life-like experiences. Looking at the tiredness levels, there was a number of negative correlations such as the WWW What component, WWW details and recognition confidence ratings. While this does indicate that tiredness has an effect on some aspects of memory it did not affect the main EM measure – combined WWW proportions.

In general, the present experiment added data supporting the hypothesis that event objects, which well represent everyday episodes, are recalled better than static objects, which in turn are common in the EM research. This effect mostly becomes visible after 24h from the initial learning and remains for at least 30 days arguably due to selective long-term memory consolidation. Such consolidation may be sleep-dependent, however, the present experiment failed to show that a shorter period between learning and sleep would lead to enhanced memory performance. The consolidation may therefore require multiple iterations of sleep, or may require time rather than sleep. Additionally, the experiment provided data showing the importance of using more than one measure when exploring EM as some information can only available in certain measures or only if certain measures are used together.

Both Experiments 1 and 2 have shown that HMD-VR can be a useful tool in EM research. However, by continuing with the thesis aim of investigating EM in an ecological fashion it is important to go back to the point made in Chapter 1 (section 1.9.5), that HMD-VR is a more ecologically valid medium compared to the more common Desktop-VR. As Desktop-VR is widespread in EM research (King et al., 2002; Plancher et al., 2008; Selzer et al., 2019; Spiers, Burgess, Hartley, et al., 2001; Spiers, Burgess, Maguire, et al., 2001) it is imperative to know the cost/effectiveness of using HMD-VR over Desktop-VR. Such comparison between the VR types will be presented in Experiment 3.

Chapter 5 – Experiment 3: Investigating differences between event and non-event memory in HMD-VR and Desktop-VR

5.1. Introduction

The previous two experiments (Chapter 3 and Chapter 4) used HMD-VR to explore EM for life-like events over time. Experiment 1 was the initial exploration of novel VEs presented through HMD-VR with room-scale navigation to investigate memory for events and non-events over a 24 hour period. The experiment showed that, in most cases, EM for events did not differ from non-events straight after experiencing them but was better recalled after the 24 hour period. Experiment 2 focused on how the EM changes over a longer 30 day period and showed similar results to Experiment 1. One of the underlying ideas of these experiments was that HMD-VR use should lead to more life-like memory representations and thus ecologically valid data. The experiment in the present chapter aimed to explore how HMD-VR compared to a more conventional and more widely used Desktop-VR while using identical VEs and tasks, to determine the validity of HMD-VR in more certain situations.

Research comparing HMD-VR and Desktop-VR in the field of psychology is scarce with even less research regarding (episodic) memory. As discussed in Chapter 1 (section 1.9), VR use in memory research started mainly in the field of spatial memory (e.g. Hamilton & Sutherland, 1999; Leplow et al., 1998; Maguire, Burgess, et al., 1998). Studies have shown that navigation in complex VEs presented through Desktop-VR led to the creation of cognitive maps that were comparable to those acquired in real-life environments (Ruddle et al., 1997). In addition to this, research showed that learning spatial representations in Desktop-VR led to those representations being transferred to real-world knowledge when navigating matching real-life environments (Arthur et al., 1997; Waller et al., 1998; Wesley Regian & Yadrick, 1994; Witmer et al., 1996). It was also found that in spatial memory tasks, Desktop-VR elicited a stronger sense of presence compared to real-life table-top experiments (Held, 1992). However, it is worth pointing out the year the mentioned table-top study was published and their use of Desktop-VR. Due to that, their findings might not be valid for modern-day HMD-VR studies. Nevertheless, ‘presence’ in research employing any type of VR is defined as the sense of mental transportation to the virtual environment which lead to the increase in the spatio-temporal accuracy, improve the conditions for encoding, and/or improve attention and focus on the task the one is being ‘present’ in (Lessiter et al., 2001). In regards to EM, presence is related to attentional engagement. It is suggested that while using any type of VR system, one’s attention is always divided between the virtual environment and the real-world (Witmer & Singer, 1998). The level of presence is argued to reflect how much attention is paid to the virtual environment (Darken et al., 1999). This notion is supported by neuroimaging, with presence leading to higher activity in fronto-parietal regions which are associated with the allocation of attentional resources (Kober & Neuper, 2012). Burgess et al. (2001) also found that events experienced in Desktop-VR led to brain activations that were not observed in conventional neuropsychological tests (Burgess et al., 2001), thus reflecting either greater cognitive engagement in the VR task or better reflecting real-life behaviour. However, another possibility is that these activations reflected engagement with a novel medium – Desktop-VR. Numerous studies have described Desktop-VR memory performance as more life-like and ecologically valid than the more traditional EM tests (Jebara et al., 2014; Picard et al., 2017; Plancher et al., 2010, 2013, 2008; Spiers, Burgess, Maguire, et al., 2001).

As discussed in Chapter 1 (section 1.9.2), with the emergence of consumer-available HMD-VR, studies started to utilise it in memory research (Cárdenas-Delgado et al., 2017; Davison et al., 2018; Krokos et al., 2019; Ouellet et al., 2018; Srivastava et al., 2019). The results from a number of the studies showed that measures obtained in HMD-VR positively correlated with traditional memory tests such as the Wechsler Memory Scale or the California Verbal Learning Test (Davison et al., 2018; Parsons & McMahan, 2017) or were better at predicting cognitive decline (Corriveau-Lecavalier et al., 2018; Ouellet et al., 2018). The main argument for utilising HMD-VR, and the one used in the present thesis, is the suggested higher ecological validity of EM encoding as compared to Desktop-VR. Studies employing both HMD-VR and Desktop-VR have shown that higher levels of sensory immersion in HMD-VR (objective level of sensory fidelity) promote better EM performance (Gamberini, 2000; Ruddle et al., 2011). Features provided by HMD-VR, such as head and hand tracking, has been shown to result in increased item and scene recognition (Ruddle et al., 2011). However, it is important to point out that in these studies, higher levels of sensory immersion were not measured but assumed due to fact that HMD-VR allowed participants to physically move around. What is lacking in the literature, and is the main focus of the present experiment, is the comparison of the two types of VR in relation to EM testing.

Both Desktop-VR and HMD-VR allow exploring memory for life-like situations using virtual environments and both types of VR show positive correlations with traditional psychological memory tests (Davison et al., 2018; Ouellet et al., 2018). However, the earlier mentioned findings of levels of presence and immersion and correlations with traditional tests observed while using HMD-VR should be used carefully as an argument of its improvement over Desktop-VR. Overall there is a lack of research comparing the two types of VR purely in relation to EM. The research employing both types of VR that is there is somehow mixed with some studies showing no difference or Desktop-VR having higher performance (Polcar & Horejsi, 2015; Sousa Santos et al., 2009; Srivastava et al., 2019), some showing better performance in HMD-VR (Harman et al., 2017; Krokos et al., 2019; Mania et al., 2003) and some showing mixed findings (Mania & Chalmers, 2001; Stevens et al., 2015).

For example, in a study by Polcar & Horejsi (2015), participants received a tour of a 3D virtual production plant using HMD-VR, Desktop-VR or CAVE system. After the tour participants were examined on their knowledge of the plant floor, such as the spatial layout or the shape of the manufacturing belt, and were asked to recall what product was being assembled. The results showed that participants in the HMD-VR group received only two-thirds of the score (points were given for each successful answer) that participants obtained in the Desktop-VR group. While this study did provide some insight into the differences between the two VR systems the study was more focused on the levels of motion-sickness and the memory measures were more focused on spatial memory than EM. An opposite trend was found in a study by Krokos et al. (2019). It was found that HMD-VR condition led to superior recall when compared to Desktop-VR. In this study, participants were asked to explore a virtual town that had faces of well-known people or characters distributed throughout it. The task was to memorise the faces and the locations in which the faces appeared. Unlike in the previous study, the main focus of this study was memory, but it again fell short of exploring EM. Additionally, participants were only allowed to rotate their view but not translate which could have affected their levels of presence or immersion. Lastly, in a study by Mania & Chalmers (2001) the difference between the two VR types depended on the measure. Participants in the study were presented with a 15-minute seminar in either real-life, Desktop-VR or HMD-VR. After the seminar participants were tested on their memory recall for factual information received in the seminar and the spatial awareness of the environment they were in. The study reported that Desktop-VR had better recall of factual information than HMD-VR but the effect only approached significance ($p<.06$). Additionally, the confidence levels were higher for the Desktop-VR than HMD-VR but the tendency to give remember judgements, which is associated with EM, was statistically higher for the HMD-VR condition.

As discussed previously, a number of studies that compared HMD-VR and Desktop-VR used old VR technology that was inferior to the one currently available. For example, the HMD-VR system used in Sousa Santos et al. (2009) study only provided a 800x600 pixel resolution with 26° field of view. As a comparison, human vision is around 220°. The HMD-VR system used in the present thesis provides 1080×1200 pixel resolution per eye with a 110° field of view. In the more recent study by Polcar & Horejsi (2015) a more advanced model of HMD-VR equipment was used that provided around 95° field of view and 960×1080 pixel resolution. Research has shown that a higher field of view leads to higher levels of presence and immersion (Bowman et al., 2009; Buttussi & Chittaro, 2018; Lin et al., 2002). More importantly, as in a number of other studies (Krokos et al., 2019; Mania & Chalmers, 2001), and unlike in the present thesis, participants were not able to walk physically around and explore the VEs.

The HMD-VR system used in the present thesis utilised room-scale tracking which provided participants the ability to use locomotion. This allows the participants to physically walk around the VEs. The ability of locomotion has been shown to improve memory performance (Chrastil & Warren, 2013; Murcia-López & Steed, 2016; van der Ham et al., 2015). For example, in a study by Murcia-López & Steed (2016) participants were asked to look at a number of virtual objects presented in VEs using HMD-VR with locomotion or Desktop-VR. In the recall stage, participants had to place real object counterparts in a real room as they remembered them from the VEs. The results showed that participants in the HMD-VR condition performed better and showed lower placement errors than participants in Desktop-VR condition. When the same HMD-VR system was used, but without the locomotion, no difference was found between Desktop-VR and HMD-VR as shown in a study by Srivastava et al. (2019) in which participants had to sketch maps of VEs they have explored. In general, research has shown that using self-movement and vestibular (balance) cues can lead to better spatial updating of egocentric representations (Frances Wang & Simons, 1999; Frances Wang & Spelke, 2002; Mou et al., 2004) and the use of self-motion cues can lead to realistic responses and behaviour to situations and events (Slater, 2009; Usoh et al., 1999). This also translates to increase in levels of immersion which positively affects learning and recall performance (Bowman et al., 2009; Dehn et al., 2018; Gamberini, 2000; Ragan, 2010; Ruddle et al., 2011; Schöne et al., 2019; Waller et al., 1998).

As HMD-VR is being used in memory research more widely, the discussed studies show an important gap in the literature: the lack of studies fully utilising HMD-VR to compare it to Desktop-VR with the focus being on EM. As such, the present experiment aimed to fill that gap. The main aim of the present experiment was to investigate if there was an effect of Group (HMD-VR versus Desktop-VR) on the free recall, WWW and object recognition tasks. Such investigation was done to obtain a greater understanding of the potential benefits of HMD-VR on EM testing. In addition to that, the experiment continued exploring the differences between events and non-events by using identical VEs and tasks as in the previous experiments.

Using the findings from the discussed literature, it was predicted that memory performance in free recall, WWW and object recognition tasks in the HMD-VR group would be higher than in the Desktop-VR group. The reasoning behind this prediction was based on the technological advances in HMD-VR systems which translate to higher levels of immersion and presence. Due to the lack of HMD-VR research focusing on EM, task-specific predictions were difficult to make. However, from the discussed literature it is visible that HMD-VR is especially useful in the field of spatial memory (Cárdenas-Delgado et al., 2017; Helbing et al., 2020; Kim et al., 2018; Murcia-López & Steed, 2016) which relates to the WWW Where component. As such, it was predicted that in the Where component data, there would be an effect of Group, with the higher memory performance in the HMD-VR.

It is important to point out that unlike in the previous experiments, the present experiment only had VR type (HMD-VR/Desktop-VR) as the between-participant variable and not time of testing (AM/PM) as before. This move was due to both Experiments 1 and 2 showing no effect of the time of testing. As the present experiment was focused on the differences in the VR types the time of sleep was not explored at all. The only time-related variable that remained was the effect of Session (Baseline/24h). By maintaining the two testing sessions the experiment was able to explore if and how different types of VR affected EM consolidation over a day-long period. As at the time of writing such investigation had not been conducted before, no predictions were made regarding the Group x Session interaction.

In general, the present study had one main objective: to investigate if and to what degree EM differed between HMD-VR and Desktop-VR. The secondary aims that were continuing through the thesis were to explore the differences between events and non-events, and how EMs for them may be affected by time-based consolidation.

5.2. Method

5.2.1. Participants

Participants in this experiment were 20 students from Bishop Grosseteste University and members of general public (mean age = 23.65; range = 18 - 41; female = 11). There was 1 student and 19 non-students. The student participant took part to obtain course credit; everyone else contributed freely. The other participants were students from different universities (n = 7) and members of the general public (n = 12). All participants had normal or corrected to normal vision. The screening procedure is described in Chapter 2 (section 1.1).

To compare the differences between HMD-VR and Desktop-VR conditions, it was decided to sample 20 participants from Experiment 2 and use them as the HMD-VR group. Due to the participants in the present experiment having their first testing session in the first half of the day (average time of testing = 11am, range = 9am – 14pm), all of the AM group (n=15) and 5 PM group participants were used from Experiment 2 as the HMD-VR group. The reason behind the reuse of the data for the HMD-VR group was the lack of available participants that did not participate in any of the previous experiments.

5.2.2. Materials

The virtual environments (VEs) and software used in this experiment were the same as in Experiment 2 with the VR exploration procedure also staying identical. The general overview of the materials and procedure can be seen in Chapter 3, sections 3.3 and 3.4.

The testing procedure remained similar to the one in Experiment 2 with one major change: participants were able to participate at any time of the day and not just at either 9am or 9pm. Participants in the present experiment were able to have their Baseline session at any time with the only requirement being that the second session was completed 24h following Baseline.

Participants in the Desktop-VR group explored identical VEs as in the previous experiments. The only difference was that the VEs were presented on a desktop computer screen. Participants used a mouse and a keyboard for navigation in the VEs.

As in Experiments 1 and 2, participants received half of all the objects (36) at the Baseline test and half of the objects (36) at the 24h test. As explained in Chapter 2, (section 2.5), half of the given objects in each session were lures and were not part of the explored VEs.

5.2.3. Design

The present experimental design was very similar to the one in Experiment 2. The experiment contained three tests (free recall, WWW and object recognition). Every test was performed at two time points: immediately after VE exploration (Baseline session) and after 24h (24h session). The exact time of testing on each of the testing varied depending on the participants' preference. The Session (Baseline/24h) was the within-subject while Group (HMD-VR/Desktop-VR) was the between-subjects independent variables.

There were five dependent variables in the free recall test: the number of recalled event objects, the number of recalled non-event objects, the number of recalled event object details, the number of recalled non-event object details and the total word count of the provided text.

The WWW task resulted in six dependant variables: the What, the When, the Where, the Event and combined WWW proportions and the average number of perceptual details recalled per one recalled (What) object. The Event component was a proportion of correctly recalled event associated with a particular event object.

After each Event and Detail recall participants had to provide a Remember/Know/Guess judgements regarding that information. This resulted in six dependant variables: proportions of Remember, Know and Guess judgements for the Event component and proportions of Remember, Know and Guess judgements for the Detail component.

The object recognition task resulted in two dependant variables: the d' sensitivity index and the confidence rating.

5.2.4. Procedure

The general procedure that followed is described in Chapter 2 (section 1.4). There were three main differences to the previous experiments. First of all, participants were able to take part in the experiment at any time of the day and not just at 9am or 9pm. Participants were asked to perform the second testing session 24h after their initial baseline test. Secondly, VR exploration only employed Desktop-VR as the HMD-VR data was taken from Experiment 2). The VEs that had been used in the previous two experiments were presented in the present study on a computer screen, and participants had to use a mouse and a keyboard rather than hand controls to perform an identical exploration task.

After the free recall, WWW and object recognition tasks, participants were given a website address to perform the memory tests after 24h (24h session). Participants were then able to leave the laboratory and carry out their normal daily activities.

5.2.5. Data analysis

The information regarding data processing and analysis can be seen in Chapter 2 (section 1.7). The data analyses were identical to the ones in Experiment 2 with two main differences. First of all, as described before, participant data from Experiment 2 was used as the HMD-VR group while the newly collected data was used as the Desktop-VR group. Secondly, only the Baseline and 24h testing session data were analysed from the HMD-VR group. The 7d and 30d session data were discarded as participants in the present experiment only had the Baseline and 24h sessions.

5.3. Results

5.3.1. Free recall

In the free recall task, participants were given space to write freely about what they had experienced in the main VEs. This was done in a form of telling a story to a friend about what the participant experienced in the VEs. Free recall tests were scored in terms of number of mentions of objects and object details. For example, “I remember seeing a radio next to a red mug and also a grey phone” would be marked as two non-event objects (radio and mug), one event object (phone), one non-event object detail (red mug) and one event object detail (grey phone).

5.3.1.1. The number of recalled objects

The number of recalled event and non-event objects was compared at each of the four testing sessions. An RM ANOVA with Session (Baseline, 24h) and object type (Event, Non-event) as within-subject variables and Group (HMD-VR, Desktop-VR) as a between-subject variable was performed. The means and standard deviations are presented in Table 5.1.

Table 5.1.

Experiment 3: Mean number of recalled objects in the free recall test

		Object type	
Group	**Session**	Event	Non-event
Desktop-VR	Baseline	8.95 (4.08)	3.45 (3.65)
	24h	8.50 (2.89)	3.90 (2.92)
HMD-VR	Baseline	7.55 (4.08)	3.50 (4.96)
	24h	7.20 (3.62)	4.55 (4.21)

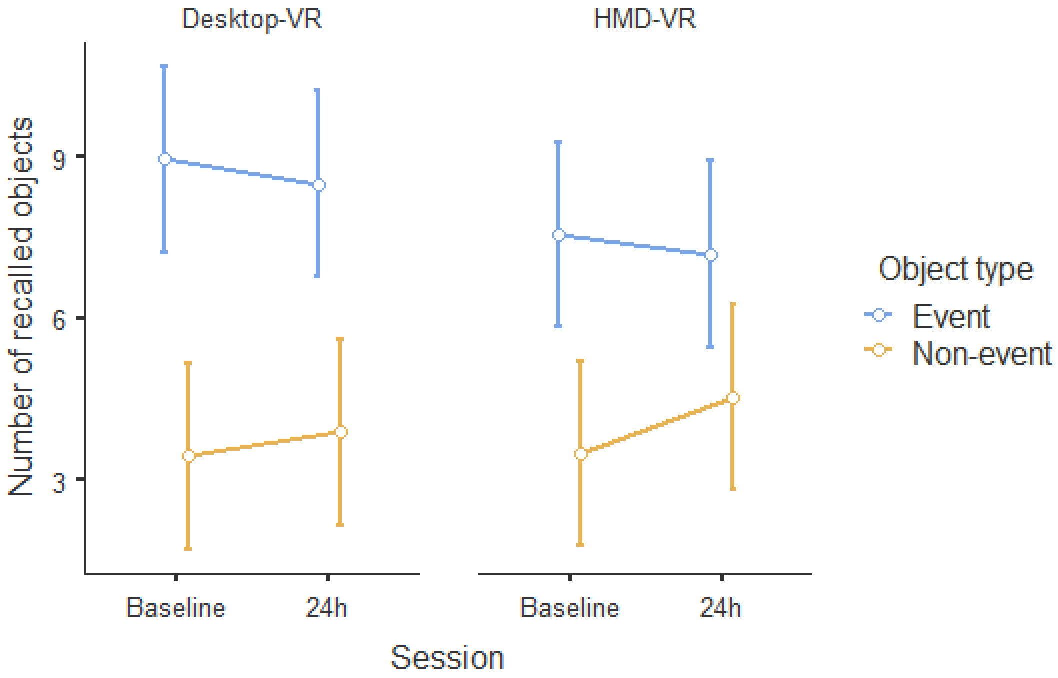
Note. Numbers in parentheses represent standard deviations

The effect of Group was not significant, $F(1,38)=.23$, $p=.632$. An effect of Object type was significant, $F(1,38)=98.04$, $p<.001$, with more Event objects being recalled than Non-event. An effect of Session was not significant, $F(1,38)=.03$, $p=.864$.

A significant Session x Object type interaction was observed, $F(1,38)=7.06$, $p=.011$ (see Figure 5.1). Pairwise comparisons showed that there were more Event objects recalled than Non-events at both Baseline, $t(56.5)=10.03$, $p<.001$, and 24h sessions, $t(56.5)=7.61$, $p<.001$. The number of recalled event objects did not differ between the Baseline and 24h sessions, $t(55.7)=.82$, $p=1$. The number of recalled non-event objects did not differ between the Baseline and 24h sessions, $t(55.7)=-1.54$, $p=.773$ None of the other interactions were significant, $Fs<4.02$, $ps>.052$.

Figure 5.1.

Experiment 3: Number of Event and Non-event objects recalled at Baseline and 24h sessions.



Notes. Error bars represent 95% confidence intervals.

5.3.1.2. The number of recalled object details

The same analyses were performed for the number of recalled object details (Event object details, Non-event object details). The means and standard deviations are presented in Table 5.2.

Table 5.2.

Experiment 3: Mean number of details recalled in the free recall test

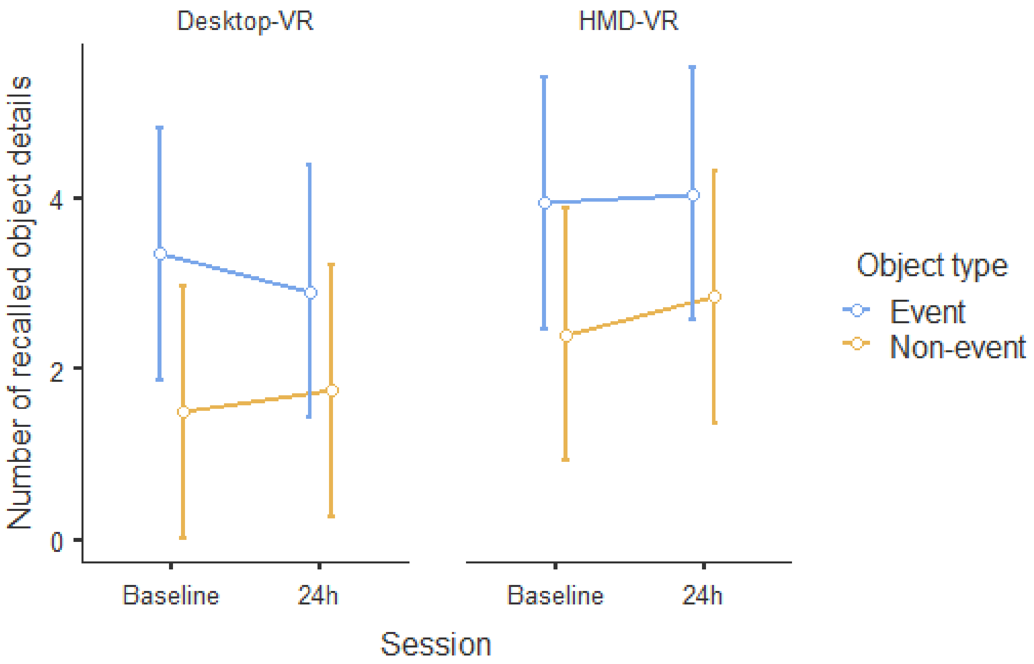
		Object type	
Group	**Session**	Event	Non-event
Desktop-VR	Baseline	3.35 (2.23)	1.50 (2.77)
	24h	2.90 (1.85)	1.75 (2.22)
HMD-VR	Baseline	3.95 (5.10)	2.4 (3.62)
	24h	4.05 (3.50)	2.85 (4.21)

Note. Numbers in parentheses represent standard deviations

An effect of Group was not significant, $F(1,38)=1.08$, $p=.305$. An effect of Object type was found, $F(1,38)=20.88$, $p<.001$, with more Event object details being recalled than Non-event (see Figure 5.2). An effect of Session was not significant, $F(1,38)=.06$, $p=.810$. None of the other effects or interactions were significant, $Fs<.65$, $ps>.425$.

Figure 5.2.

Experiment 3: Mean number of recalled object details in the free recall test.



Notes. Error bars represent 95% confidence intervals.

5.3.2. Combined What-Where-When

When a participant correctly recalled the object (What), the room number the object was in (When) and where in the room the object was (Where) it was said that the participant recalled the full WWW information regarding that object. The combined WWW score is the proportion of correctly recalled WWW information out of all given objects (ranges from 0 to 1). For example, a score of 0.5 would mean that a participant recalled combined WWW information for 18 objects out of 36 possible. The means and standard deviations are presented in Table 5.3.

Table 5.3.

Experiment 3: Mean proportions of combined What-When-Where recalls.

		Object type	
Group	**Session**	Event	Non-event
Desktop-VR	Baseline	.53 (.50)	.45 (.50)
	24h	.42 (.50)	.24 (.43)
HMD-VR	Baseline	.53 (.50)	.53 (.50)
	24h	.46 (.50)	.31 (.46)

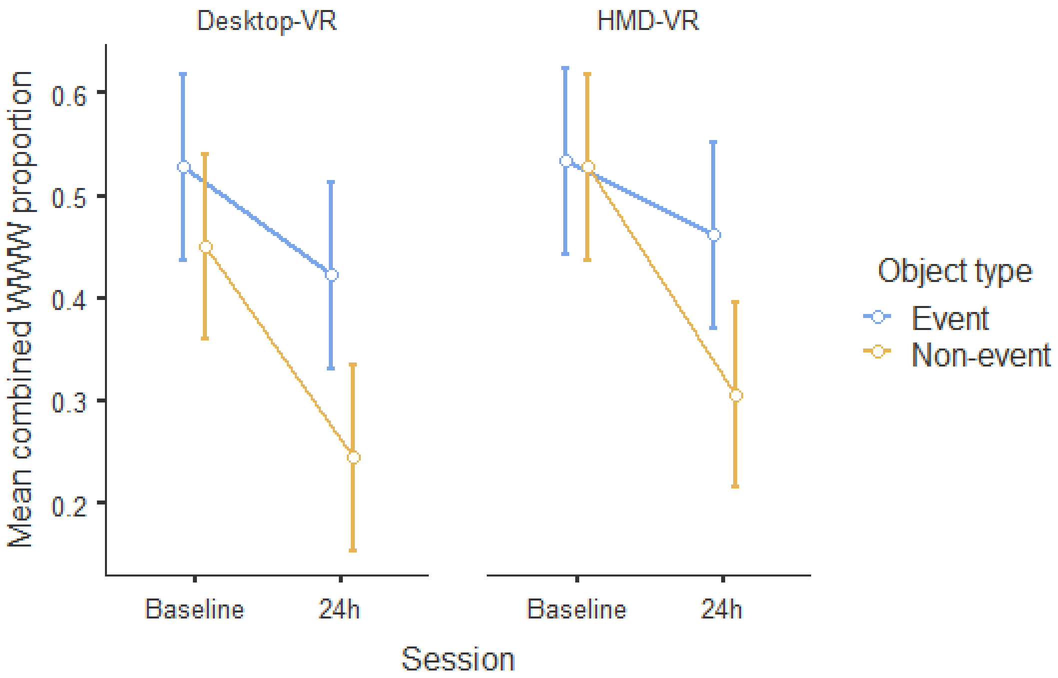
Note. Numbers in parentheses represent standard deviations. Descriptives for the separate WWW components can be seen in Table 5.4.

The effect of Group was not significant, $F(1,38)=1.04$, $p=.313$. A significant effect of the object type was observed, $F(1,38)=18.61$, $p<.001$, with Event objects having higher WWW proportions compared to the Non-event objects (see Figure 5.3). A significant effect of Session was observed, $F(1,38)=22.14$, $p<.001$, with higher correct proportions at Baseline than after 24h.

A significant Session x Object type interaction emerged, $F(1,38)=7.33$, $p=.001$. Multiple comparisons revealed Event objects having higher correct WWW proportions in the 24h Session compared to the Non-event objects, $t(75.87)=4.99$, $p<.001$, lower Non-event proportions at the 24h Session compared to Baseline, $t(68.9)=5.40$, $p<.001$, and higher Event object proportions at Baseline compared to Non-event object proportions after 24h, $t(70.5)=6.35$, $p<.001$. There were no differences between Event and Non-event proportions in the Baseline session, $t(75.8)=1.25$, $p=1$. No other interactions were significant, $Fs<.96$, $ps>.334$. Overall, the interaction revealed that mean combined WWW proportions did not differ for Event and Non-event objects at the Baseline but were higher for the Event objects after 24h. Proportions for the Event object did not differ between the two testing sessions while proportions for the Non-event objects were lower after 24h.

Figure 5.3.

Experiment 3: Mean proportions of recalled combined What-When-Where information.



Notes. The combined WWW score is the proportion of correctly recalled What-When-Where information for all of the possible objects (ranges from 0 to 1). Error bars represent 95% confidence intervals.

5.3.3. Separate What-Where-When components

5.3.3.1. What

The What component represents a recall of an object. On the task screen, it is worded as "Do you recall X?" where X was a name of an object. Similarly to the combined WWW, the What component was measured as a proportion of correctly recalled objects. The means and standard deviations are presented in Table 5.4.

Table 5.4.

Experiment 3: Mean proportions of the recalled What, When and Where components

			Object type	
Group	**Session**	**Component**	Event	Non-event
Desktop-VR	Baseline	What	.79 (.41)	.78 (.41)
		When	.62 (.49)	.58 (.50)
		Where	.67 (.47)	.59 (.49)
	24h	What	.68 (.47)	.52 (.50)
		When	.53 (.50)	.36 (.48)
		Where	.55 (.50)	.37 (.49)
HMD-VR	Baseline	What	.80 (.40)	.82 (.38)
		When	.58 (.49)	.66 (.48)
		Where	.70 (.46)	.63 (.48)
	24h	What	.69 (.46)	.53 (.50)
		When	.56 (.50)	.42 (.50)
		Where	.56 (.50)	.36 (.48)

Note. Numbers in parentheses represent standard deviations.

An effect of Group was not found, $F(1,38)=.41$, $p=.524$. A significant effect of Session was observed, $F(1,38)=59.61$, $p<.001$, with higher proportions at Baseline than after 24h. A significant effect of the Object type was observed, $F(1,38)=11.98$, $p=.001$, with Event objects having higher proportions compared to the Non-event objects.

A significant Session x Object type interaction was detected, $F(1,38)=13.30$, $p<.001$ (see Figure 5.4). Multiple comparisons revealed Event objects having higher What proportions in the 24h Session compared to the Non-event objects, $t(76)=5.03$, $p<.001$, lower Non-event proportions at the 24h Session compared to Baseline, $t(75.5)=8.15$, $p<.001$, and higher Event object proportions at Baseline compared to Non-event object proportions after 24h, $t(75.3)=3.40$, $p<.001$. There was no difference between Event and Non-event proportions in the Baseline session, $t(76)=-.17$, $p=1$. None of the other interactions were significant, $Fs<1.45$, $ps>.244$.

5.3.3.2. When

The When component represents a correct recall of a room number in which the object from the What task was seen. On the task screen, it is worded as “In which room you have seen X?” where X is a name of an object from the previous task. Similarly to the combined WWW, the When component was measured as a proportion of correct room recalls out of all possible. The means and standard deviations are presented in Table 5.4.

An effect of Group was not found, $F(1,38)=.58$, $p=.452$. A significant effect of Session was observed, $F(1,38)=21.34$, $p<.001$, with higher proportions at Baseline than after 24h. A significant effect of the object type was observed, $F(1,38)=8.23$, $p=.007$, with Event objects having higher proportions compared to the Non-event objects.

A significant Session x Object type interaction was detected, $F(1,38)=13.30$, $p<.001$ (see Figure 5.4). Multiple comparisons revealed Event objects having higher When proportions in the 24h Session compared to the Non-event objects, $t(74.6)=4.85$, $p<.001$, lower Non-event proportions at the 24h Session compared to Baseline, $t(67.3)=6.11$, $p<.001$, and higher Event object proportions at Baseline compared to Non-event object proportions after 24h, $t(72)=5.40$, $p<.001$. There was no difference between Event and Non-event proportions in the Baseline session, $t(74.6)=-.52$, $p=1$. None of the other interactions were significant, $Fs<2.23$, $ps>.144$.

5.3.3.3. Where

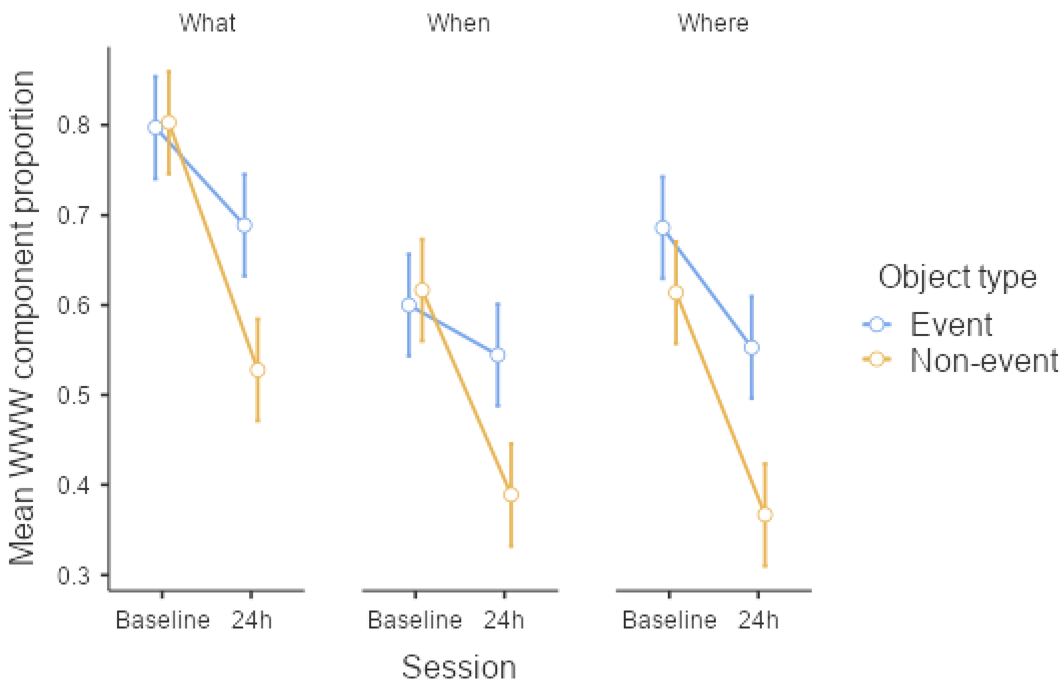
The Where component represents correctly recalling object’s location in the virtual room. On the task screen, participants had to use a mouse and point on a top-down map of the room where they thought the object was located. This provided a distance – how far away participant’s guess was from the object’s real location. Using a method explained in Chapter 2 (section 2.7) the distance was converted to a binary correct/incorrect outcome. Similarly to the combined WWW, the Where component was measured as a proportion of correct location recalls out of all possible. A high correlation was observed between the Where pointing errors and the Where proportions ($r = -.911$, $n = 80$, $p<.001$). The means and standard deviations are presented in Table 5.4.

Effect of Group was not found, *F*(1,38)=.17, *p*=.681. A significant effect of Session was observed, *F*(1,38)=53.78, *p*<.001, with higher proportions at Baseline than after 24h (see Table 5.4). A significant effect of the object type was observed, *F*(1,38)=29.27, *p*<.001, with Event objects having higher proportions compared to the Non-event objects.

A significant Session x Object type interaction was detected, *F*(1,38)=4.22, *p*=.047 (see Figure 5.4). Multiple comparisons revealed Event objects having higher Where proportions in the 24h Session compared to the Non-event objects, *t*(74.4)=5.09, *p*<.001, lower Non-event proportions at the 24h Session compared to Baseline, *t*(75.7)=6.51, *p*<.001, and higher Event object proportions at Baseline compared to Non-event object proportions after 24h, *t*(75.5)=9.06, *p*<.001. There was no difference between Event and Non-event proportions in the Baseline session, *t*(74.4)=1.97, *p*=.312. None of the other interactions were significant, *Fs*<.48, *ps*>.491.

Figure 5.4.

Experiment 3: Mean correct proportions recalled for the separate What-When-Where components.



Notes. To make the figure more readable and due to the lack of the effect, the data presented here was not split by the Group (HMD-VR/Desktop-VR). Error bars represent 95% confidence intervals

5.3.4. Detail

The Detail component represents a mean number of recalled details for one object. On the task screen, participants were asked if they could recall any perceptual detail about an object and if they could write one down. Participants were able to write down up to five details per one object. The means and standard deviations are presented in Table 5.5.

Table 5.5.

Experiment 3: Mean number of object details recalled per one recalled object

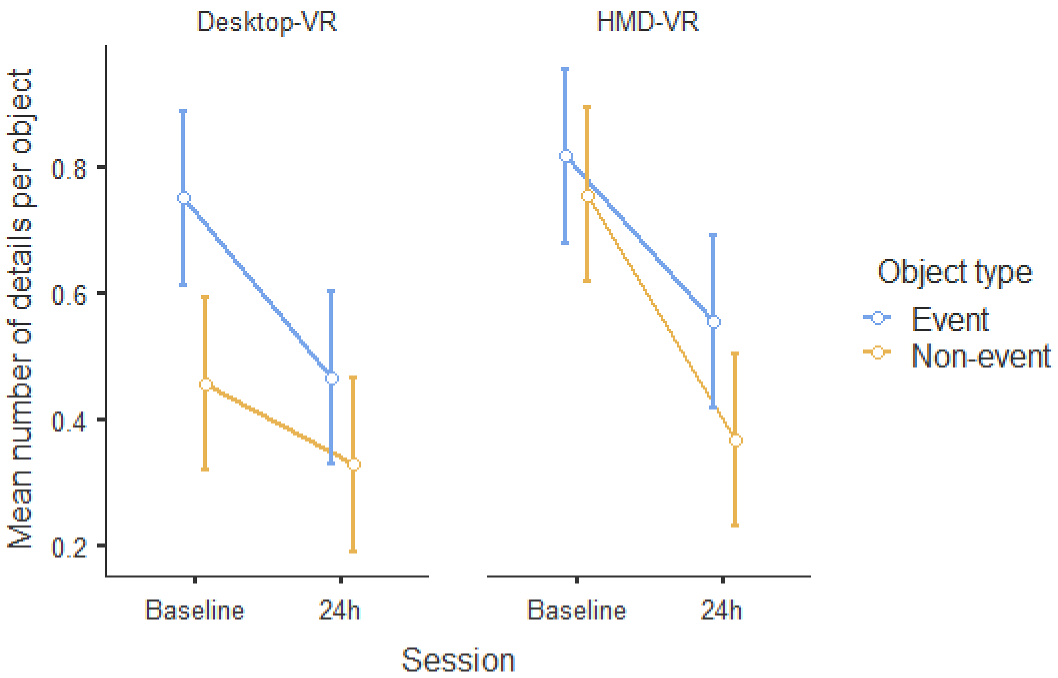
		Object type	
Group	**Session**	Event	Non-event
Desktop-VR	Baseline	.75 (.90)	.46 (.61)
	24h	.47 (.63)	.33 (.55)
HMD-VR	Baseline	.82 (.93)	.76 (.89)
	24h	.56 (.79)	.37 (.67)

Note. Numbers in parentheses represent standard deviations.

An effect of Group was not found, $F(1,38)=3.10$, $p=.086$. A significant effect of Session was observed, $F(1,38)=57.31$, $p<.001$, with more details recalled at Baseline than after 24h. A significant effect of the object type was observed, $F(1,38)=22.40$, $p<.001$, with participants recalling more details for the Event objects compared to the Non-event objects (see Figure 5.5). None of the interactions were significant, $Fs<2.91$, $ps>.096$.

Figure 5.5.

Experiment 3: Mean number of object details recalled per one recalled object



Notes. Error bars represent 95% confidence intervals

5.3.1. Remember/know/guess judgements

Separate RM ANOVAs were used to analyse the remember/know/guess judgements for the What, When, Where, Event and Detail components. The judgements were transformed into overall proportions in a similar fashion to Dewhurst, Conway, & Brandt (2009). For example, adding one participant's Remember, Know and Guess judgement proportions at Baseline would equal to 1. This transformation was undertaken so that the lower number of recalled objects on the second session would not affect the judgement data. The means and standard deviations of the R/K/G judgements are presented in Table 5.6.

Table 5.6.

Experiment 3: Mean proportions of the Remember/Know/Guess judgements given in the WWW task

		Object type			
		Event		Non-event	
		Group		Group	
Judgement	**Session**	Desktop-VR	HMD-VR	Desktop-VR	HMD-VR
Remember	Baseline	.71 (.16)	.54 (.30)	.64 (.12)	.56 (.30)
	24h	.68 (.17)	.56 (.34)	.67 (.22)	.49 (.31)
Know	Baseline	.21 (.14)	.35 (.30)	.24 (.11)	.31 (.27)
	24h	.21 (.15)	.32 (.29)	.18 (.13)	.29 (.31)
Guess	Baseline	.08 (.06)	.12 (.12)	.12 (.09)	.13 (.11)
	24h	.10 (.11)	.12 (.13)	.14 (.11)	.22 (.18)

Note. Numbers in parentheses represent standard deviations.

Remember

The effect of Group was significant, $F(1,38)=4.74$, $p=.036$, with higher Remember proportions being in Desktop-VR group compared to HMD-VR group (see Figure 5.6). No other effects or interactions were significant, $Fs<1.24$, $ps>.645$.

Know

The effect of Group was not significant, $F(1,38)=2.99$, $p=.092$. No other effects or interactions were significant, $Fs<1.89$, $p>.177$.

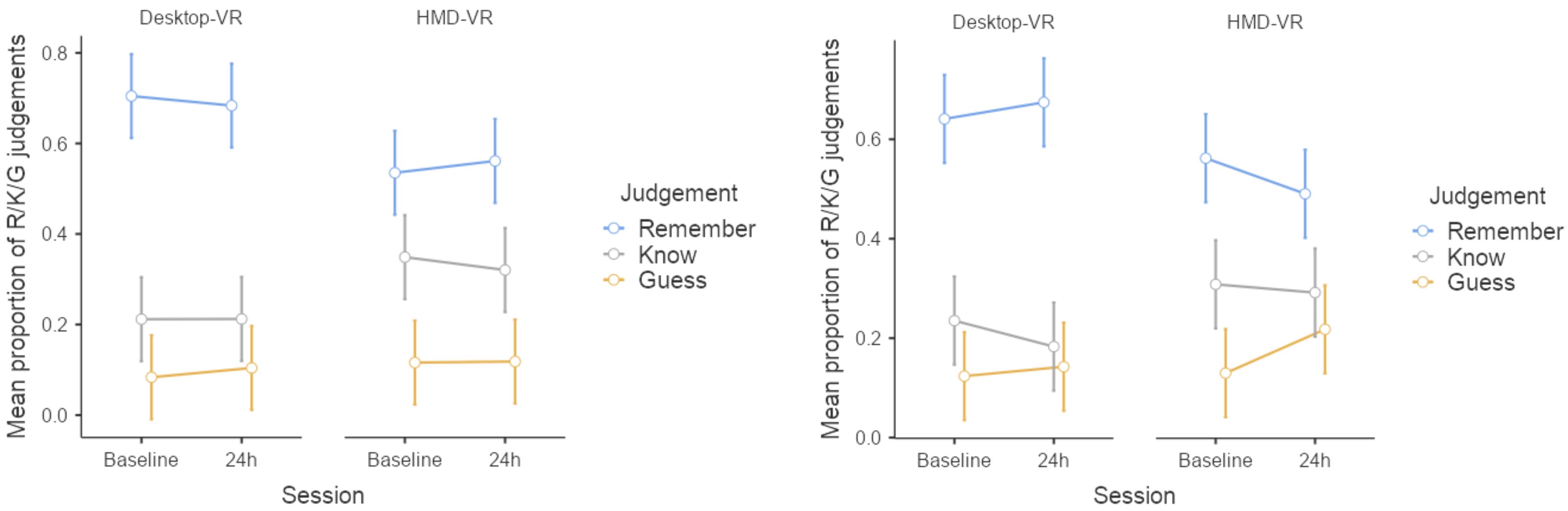
Guess

The effect of Group was not significant, $F(1,38)=1.27$, $p=.267$. The effect of Session was significant, $F(1,38)=6.10$, $p=.018$, with higher Guess proportions at the 24h session compared to the Baseline session. There was a significant effect of Object type,

$F(1,38)=8.47$, $p=.006$, with Event objects having lower proportions than Non-event objects. No interactions were significant, $Fs<3.71$, $p>.062$.

Figure 5.6.

Experiment 3: Mean total proportions for Remember, Know and Guess judgements provided in the WWW task.



Notes. The left figure represents Event object data while the right figure represents the Non-event object data. Error bars represent 95% confidence intervals

5.3.2. Recognition – d’ scores

Recognition data was converted to d’ scores (Z(hit rate) − Z(false alarm rate)). The means and standard deviations are presented in Table 5.7.

Table 5.7.

Experiment 3: Mean d’ object recognition scores

		Object type	
Group	**Session**	Event	Non-event
Desktop-VR	Baseline	2.81 (.51)	2.35 (.60)
	24h	2.35 (.60)	1.73 (.71)
HMD-VR	Baseline	2.91 (.51)	2.63 (.56)
	24h	2.41 (.71)	1.84 (.76)

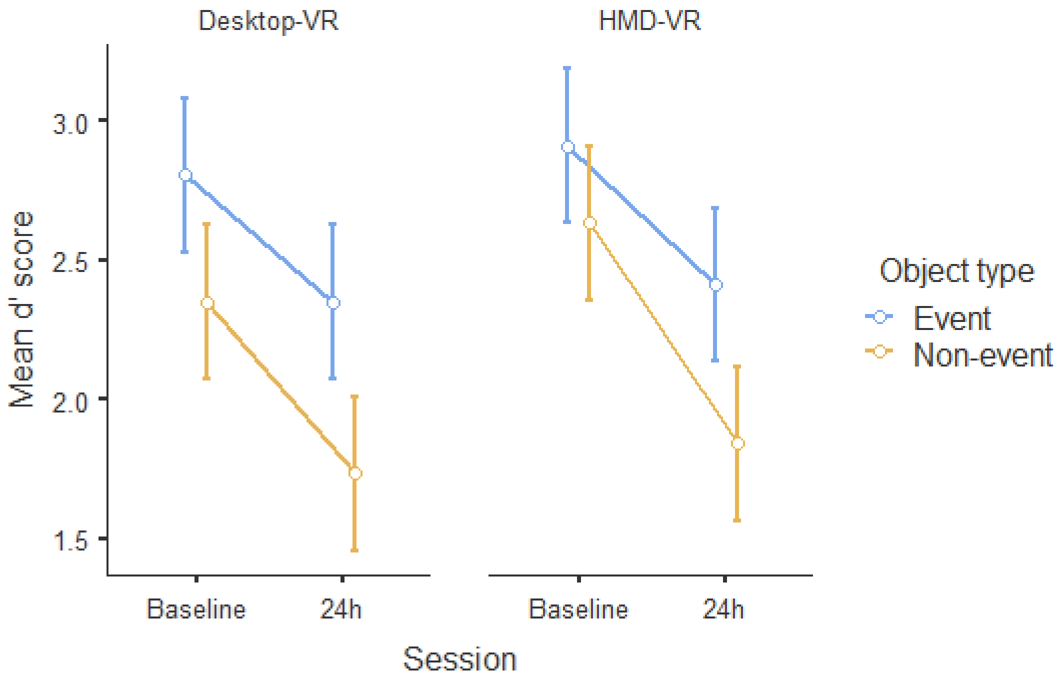
Note. Numbers in parentheses represent standard deviations.

The effect of Group was not significant, $F(1,38)=.88$, $p=.355$). There was a significant effect of Session, $F(1,38)=35.19$, $p<.001$, with higher d’ scores at Baseline than after 24h. There was a significant effect of Object type, $F(1,38)=53.66$, $p<.001$, with Event objects having higher d’ scores than Non-event objects.

A significant Session x Object type interaction was detected, $F(1,38)=5.60$, $p=.023$ (see Figure 5.7). Multiple comparisons revealed Event objects having higher mean d’ scores in the 24h Session compared to the Non-event objects, $t(69.2)=7.32$, $p<.001$, lower Non-event scores at the 24h Session compared to Baseline, $t(54.4)=6.37$, $p<.001$, and higher Event object scores at Baseline compared to Non-event object proportions after 24h, $t(65.8)=8.99$, $p<.001$. Event objects had higher d’ scores at Baseline than the Non-event objects, $t(69.2)=4.55$, $p<.001$. None of the other interactions were significant, $Fs<.73$, $ps>.400$.

Figure 5.7.

Experiment 3: Mean d' object recognition scores



Notes. Error bars represent 95% confidence intervals

5.3.3. Recognition – Confidence ratings

The confidence ratings were ratings (from .0 to 1.0) reflecting how confident the participants felt about their recognition judgement. A confidence rating of .0 would indicate being not confident at all whereas confidence rating of 1.0 would indicate full confidence. The means and standard deviations are presented in Table 5.8.

Table 5.8.

Experiment 3: Mean object recognition confidence ratings and standard deviations

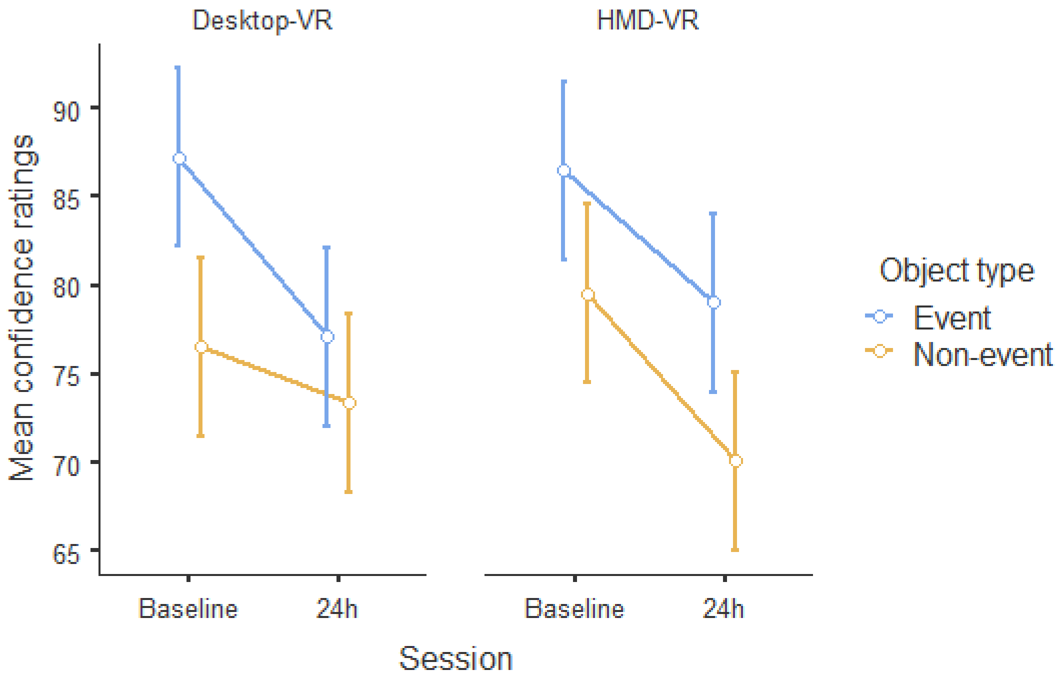
		Object type	
Group	**Session**	Event	Non-event
Desktop-VR	Baseline	87.22 (7.86)	76.53 (11.75)
	24h	86.47 (8.53)	79.52 (10.49)
HMD-VR	Baseline	77.12 (10.7)	73.36 (16.97)
	24h	79.03 (12.02)	70.11 (9.84)

Note. Numbers in parentheses represent standard deviations.

The effect of Group was not significant, $F(1,38)=.01$, $p=.932$. There was a significant effect of Session, $F(1,38)=25.75$, $p<.001$, with higher confidence ratings at Baseline than after 24h. There was a significant effect of Object type, $F(1,38)=42.31$, $p<.001$, with Event objects having higher confidence ratings than Non-event objects. No interactions were significant, $Fs<2.04$, $p>.162$.

Figure 5.8.

Experiment 3: Mean object recognition confidence ratings.



Notes. Error bars represent 95% confidence intervals

5.3.4. The effect of tiredness and sleep time

One of the potential explanations for the lack of effect of Group (AM/PM) in Experiment 1 was tiredness. As an exploratory measure in Experiment 2, participants were asked to indicate how tired they were (on a scale from 1 to 10). Additionally, participants were asked to indicate what time they went to sleep the day before and the time they woke up on the day of the test. This was added as a precaution, in case of any technical difficulties with the actigraphy bracelets. To explore if there were any relationships between the level of tiredness and self-reported time spent asleep a Pearson's r correlations were performed between the mentioned measures and the EM measures. The same analyses were done in the present experiment to explore if tiredness and self-reported time spent asleep correlated with any of the EM measures (see Table 5.10). The means and standard deviations are presented in Table 5.9.

Table 5.9.

Experiment 3: Subjective time spent asleep

Group	Session	Time asleep (mins)	Tiredness level
Desktop-VR	Baseline	454 (46.40)	3.20 (2.12)
	24h	464 (45.80)	4.70 (2.23)
HMD-VR	Baseline	467 (74.70)	3.45 (2.35)
	24h	458 (98.00)	5.05 (2.11)

Note. Numbers in parentheses represent standard deviations.

The level of tiredness and self-reported time spent asleep did not show any significant correlations any of the EM measures, *ps*>.125.

Table 5.10.

Experiment 3: Correlation table between tiredness levels and self-reported time spent asleep

	Tiredness	Time asleep
Free recall: event objects	-0.013	-0.023
Free recall: non-event objects	-0.03	0.081
WWW event objects	-0.107	0.044
WWW non-event objects	0.229	0.146
d' event objects	0.243	0.089
d' non-event objects	0.377	0.06

Note. * $p < .05$, ** $p < .01$, *** $p < .001$

5.4. Discussion

The present chapter aimed to investigate if and to what degree EM differed between HMD-VR and Desktop-VR. The secondary aims that were continuing through the thesis were to explore the differences between events and non-events and how EMs for them are affected by memory consolidation. Participants performed the same tasks as in the previous two experiments with the main difference being that the exploration was undertaken using Desktop-VR. To compare memory performance between the two types of VR, the data from the present study were compared to the HMD-VR data obtained from Experiment 2.

The general prediction was that EM performance in the free recall, WWW and object recognition tests would be higher in the HMD-VR group, compared to the Desktop-VR. The prediction was based on the research showing that the increased levels of presence and immersion obtained in the HMD-VR lead to better memory performance when compared to Desktop-VR (Cárdenas-Delgado et al., 2017; Harman et al., 2017; Krokos et al., 2019; Mania et al., 2003; Repetto et al., 2016). Surprisingly, the memory performance did not differ between the two VR types in almost all of the EM tests and measures.

Out of all the measures (free recall, WWW and object recognition), the effect of Group (Desktop-VR/HMD-VR) was only significant in the WWW Remember judgement proportions, showing higher proportions in the Desktop-VR group (it is worth pointing out the relatively high p-value, p=.036, perhaps indicating that this could have occurred by chance). This result not only goes against the main prediction of the experiment but also against the findings of one of the previously presented studies in which more remember judgements were found in the HMD-VR condition when compared to Desktop-VR (Mania & Chalmers, 2001). However, after a closer look at the full R/K/G data (see Figure 5.6) some other interesting trends can be seen. While there were fewer Remember judgements in the HMD-VR group, on average there were more Know judgements even if the difference was not statistically significant (p=.092). These data suggest that while the overall memory performance did not differ between the VR groups the way it was retrieved was based more on familiarity than recollection (Migo et al., 2012; Yonelinas, 2002). However, this is contradictory as familiarity is associated with semantic memory whereas recollection is based on EM. If participants did rely more on semantic than EM this should have been reflected in the combined WWW measure but it did not. It is important to remember that these are just speculations as the statistical difference was observed only in the Remember but not Know judgements. As such, this finding is closer to being a statistical anomaly than a notable trend.

One of the more specific predictions concerned the spatial Where component was that there would be a higher memory performance in the HMD-VR group compared to Desktop-VR. This was based on the research showing higher memory performance while using HMD-VR (Cárdenas-Delgado et al., 2017; Helbing et al., 2020; Kim et al., 2018; Murcia-López &

Steed, 2016) and the positive effect of locomotion (Chrastil & Warren, 2013; Murcia-López & Steed, 2016; van der Ham et al., 2015). This was not the case as performance in both groups did not differ. This suggests that spatial representations were created equally well regardless of the VR type. Indeed performance was comparable across the two modes of presentation for almost all tests.

There are a number of explanations for the lack of difference between the groups. First of all, it is possible that the encoding of EM in the present life-like VEs was not affected by the higher levels of presence. This would lead to two interesting conclusions. First of all, it would suggest that EM can be equally well encoded regardless if the observer is surrounded by the environment in which the event happens or just observing it on a screen. This means that EMs can be formed regardless of how the VEs are presented and the EM encoding is based more on the events that make up the EMs than how those events are presented. This leads to the second conclusion, that HMD-VR might not be necessary to obtain ecologically valid data and instead Desktop-VR can be used equally well. Interestingly, such a conclusion would go against the main premise of the present thesis that HMD-VR should lead to a more ecologically valid EM data.

Another possible explanation for the lack of the effect of Group could be due to the novelty of the HMD-VR system. As discussed in the introduction of Chapter 3, there are no studies, at the time of writing, that have used a task that is similar to the one used in the present VEs. The discussed higher levels of immersion and presence in the HMD-VR group and its predicted positive effect on memory might have been overshadowed by the novelty VR system itself. As discussed in Chapter 1 (section 1.9.3), HMD-VR is a relatively new system with only recently becoming fully consumer-available. Due to this, not a lot of people have experienced it making it a novel encounter. Indeed, an argument regarding HMD-VR novelty was also used by Polcar & Horejsi (2015). The participants in the HMD-VR group were usually very surprised by the VEs. In addition to the general surprise and novelty of the VR headset, the VR system used in the present thesis has also utilised the room-scale headset tracking that comes with the newer versions of the HMD-VR systems. This allowed participants to employ locomotion and physically walk around the VEs. These features might have led participants in the present experiment to pay less attention to the task as their attention was divided by the use of HMD-VR system itself and research has shown that attention modulates memory (Buckner et al., 2000; Chun & Turk-Browne, 2007; Iidaka et al., 2000). Continuing with this explanation it is also possible to argue that desktop computers and thus Desktop-VR are more well-known which might have led participants to not pay as much attention to the VR system and concentrate on the task. These two opposing effects might have led to the lack of differences between the two VR systems.

The effect of novelty opens up an interesting avenue of research and a problem that can be seen in a lot of HMD-VR research – the effect of habituation to HMD-VR. In a number of discussed studies, participants were not given a lot of time to get used to HMD-VR or such information is not provided (Krokos et al., 2019; Mania & Chalmers, 2001; Polcar & Horejsi, 2015; Sousa Santos et al., 2009). The HMD-VR group participants in the present and past experiments spent only around two minutes in the training VE before starting the main exploration. This lack of habituation to a novel activity might have led to lower memory performance.

Another explanation for these findings and related to technical differences in VR systems is the observed differences in participant behaviour. During the VE exploration, it was noted that exploratory behaviour between the two groups of participants differed. Participants in the Desktop-VR group tended to explore less and spend more time 'standing' still and just use the mouse to look around the VEs. This led to the experimenter providing more prompts for the participants to continue exploring the VEs so that they could trigger the Event objects. On the other hand, in the HMD-VR group, participants tended to walk around and explore a lot more instead of standing still. This difference in exploration might have allowed participants in the Desktop-VR group to encode more information due to spending more time looking at wider areas containing more objects. The differences in exploratory behaviour between the two types of VR has been noted in an earlier mentioned study by Murcia-López & Steed (2016). Interestingly the study found an opposite behaviour between the two types of VR; less exploration and more time spent stationary in the HMD-VR as compared to Desktop-VR. Nevertheless, the study showed better memory performance in the HMD-VR group. These findings and observations show the need to track the behaviour of participants when comparing different types of VR. It is likely that if participants are given enough time to get used to the HMD-VR system the hypothesised increase in performance might become more visible.

Lastly, it is worth noting the continued difference between the event and non-event objects and the differences between the testing sessions. Looking at the results it is visible that the effects and trends observed in the previous two experiments that used HMD-VR were also visible in the present experiment. Event objects were better recalled in every measure. Additionally, the Object type x Session interaction was also visible in most of the measures which show that even in Desktop-VR, event and non-event objects were encoded at the same level and only after 24h and through selective memory consolidation event objects were later recalled better. This provides evidence that the more life-like events are recalled better than static objects, regardless of the VR type.

In general, the present experiment provided data showing no differences between Desktop-VR and HMD-VR when measuring EM. This leads to an important conclusion that Desktop-VR may be as useful as HMD-VR for the study of EM. However, more research is needed in which participants' behaviour in VR is controlled for. In terms of EM, the present experiment again showed that EM for events is preferentially consolidated over EM for non-events leading to better memory. In conclusion, the present study showed that it is possible to obtain ecologically valid EM data using Desktop-VR.

Due to this conclusion, it was important to explore the ecological validity of HMD-VR and Desktop-VR even further and compare EM obtained in the two VR systems to EM obtained in a real-world setting. While literature shows that knowledge is can be easily transferred between Desktop-VR and the real-life and HMD-VR and the real-life (for reviews see Brooks & Rose, 2003; Smith, 2019), there is lack of research that compares all three conditions while focusing on EM. Furthermore, measures of immersion and presence in the VR states could offer insight into the ways in which encoding is experienced, which may increase ecological validity, even if memory performance at test is unaffected.

Chapter 6 – Experiment 4: Episodic memory differences among real-life, HMD-VR and Desktop-VR

6.1. Introduction

The previous experiment explored episodic memory (EM) differences between HMD-VR and Desktop-VR. The results showed that memory performance did not differ between HMD-VR and Desktop-VR conditions in terms of free recall, WWW or object recognition tasks. These findings did not support the hypothesis that HMD-VR should lead to more ecologically valid EM testing and better memory performance. Instead, the results suggest EM can be explored equally well through the more commonly used Desktop-VR. However, two important points were raised that the similar performance might have been due to the nature of the task (free exploration) and potential differences in the levels of presence between HMD-VR and Desktop-VR. In order to address this, the present experiment aimed to explore these points by using a changing the main experimental task and measuring participant's levels of immersion and presence. Additionally, the present experiment aimed to explore the ecological validity of EM testing by introducing a laboratory-based real-life condition. This was included to gain a better insight into the differences between the HMD-VR and Desktop-VR when compared to the real-life.

As has been discussed, EM testing lacks ecological validity in both clinical and laboratory settings (Parsons, 2015; Sbordone, 2008; Silver, 2000). Experiments in the present thesis so far have explored how HMD-VR can be used to present more life-like tasks (in comparison to word or picture lists) that involve observing unique events in distinct environments. While the experiments showed comparable results to findings from other literature such as item (What) memory being higher than the memory for temporal (When) or spatial (Where) information (Experiments 1, 2 and 3) (e.g. Plancher et al., 2008) or the effect of retesting (Experiment 2) (e.g. Baddeley et al., 2019). As one of the main premises of the present thesis was that HMD-VR should provide close to real-life experiences, it is important to explore how EM derived from HMD-VR (and Desktop-VR) compares to EM from an equivalent real-life experience.

As discussed in Chapter 1 (section 1.9.3), research shows that knowledge is easily transferable between Desktop-VR and real-life and HMD-VR and real-life (for reviews see Brooks & Rose, 2003; Smith, 2019). However, in the field of EM, there is a lack of research comparing all three conditions. In a number of studies that did compare memory performance between real-life and HMD-VR, the results have shown better source memory performance (Hoffman et al., 2001) and spatial memory (Waller et al., 1998). For example, in a study by Hoffman et al., (2001), participants were asked to touch a number of objects in real-life and in HMD-VR. The location of their hands were tracked in HMD-VR but lacked any tactile feedback. After a week, participants were given names of the real and virtual objects intermixed with some new lure objects and were asked to indicate if they recognised those objects from before and how confident they were with their decisions. Both item recognition performance and the confidence ratings were higher for the real items compared to virtual. In light of the results from Experiment 3, the comparison between HMD-VR and real-life is especially important. Not only would it provide additional insights into the memory performance differences between HMD-VR and Desktop-VR but also a reference point (the real-life condition) for a better understanding of EM.

In general, literature shows that memory performance in a real-life condition should be higher than in any virtual condition (Flannery & Walles, 2003; Hoffman et al., 2001; Waller et al., 1998; for a review see Smith, 2019). While it is commonly assumed that HMD-VR, should produce data that is more representative of everyday behaviour, performance between HMD-VR and Desktop-VR has been shown not to differ by a number of studies (Jensen & Konradsen, 2018; Mania & Chalmers, 2001; Sousa Santos et al., 2009). The results from Experiment 3 also reflect this. This suggests that the assumed higher levels of presence and immersion in HMD-VR do not lead to better EM performance and that it is possible to acquire equally ecologically valid data using Desktop-VR. The present experiment further explores differences among the VR applications. If EM performance does not differ between HMD-VR and Desktop-VR, it will provide evidence that Desktop-VR, which is more affordable and accessible, is as good a medium for memory testing and producing life-like experiences as HMD-VR.

As discussed in the previous chapter, the lack of differences between the two VR conditions could have stemmed from the free exploration leading to behavioural differences, and the way visual information is provided in the two VR conditions. As discussed, participants' behaviour differed between the two VR conditions with the participants in Desktop-VR group tending to stand in one place and needing more 'encouragement' for the exploration. It was argued that by just standing and looking around participants were better able to observe and encode the objects and their locations. However, it is important to point out the research showing that locomotion and orientation in HMD-VR lead to better (spatial) memory (Chrastil & Warren, 2013; Murcia-López & Steed, 2016; van der Ham et al., 2015). It is especially important to discuss the exploration results in a study by Murcia-López & Steed (2016). As discussed in the previous chapter, participants were asked to look at a number of virtual objects presented in VEs using HMD-VR with locomotion, Desktop-VR or real-life. In the recall stage, participants had to place real object counterparts in a real room as they remembered them from the VEs. Participants' movements in the VEs were recorded and mapped. After plotting their movements it was shown that participants using Desktop-VR tended to spend more time outside the object placing area when compared to HMD-VR and real-life. The explanation for this difference was that when learning object locations in less immersive systems such as Desktop-VR, participants tend to navigate toward the boundaries of the environment to obtain a more global view of the scene. This explanation can be related to the behaviour seen in Experiment 3. Instead of actively exploring the environment participants tended to look around to get a view of the scene. This highlights how different VR systems can lead to different behavioural interaction with said systems. Additionally, the novelty of the life-like 360° presentation in HMD-VR and the headset itself might have taken attentional resources from the tasks itself reducing the episodic encoding compared to the 2D Desktop-VR presentation (see Polcar & Horejsi (2015) for a similar observation). The present experiment aimed to address both of these problems.

The behaviour problem regarding participants being stationary can be addressed by employing a goal-based task which will force (encourage) interaction and 'locomotion' (albeit virtual). An active task involving interaction with objects should prevent participants from passively standing still and observing the environments. This could stop participants paying attention to the headset and the general novelty of HMD-VR which in turn should allow exploration of the true differences in EM. While it is possible to argue that interaction happens both in HMD-VR and Desktop-VR, the difference is that in HMD-VR, the interactions are more life-like due to the locomotion and physical enactment. Research shows that interactivity such as handling of objects that later one would be tested on positively affects memory due to the motor information facilitating both encoding and retrieval of memory (Mohr et al., 1989; Russ et al., 2003; Zimmer & Engelkamp, 1989). More importantly, this effect has been observed even when an action is pantomimed (Nilsson, 2000) which is arguably what happens in HMD-VR.

The suggested changes to the present experiment so far have mostly focused on the methodological side of the study. It is also important to continue exploring and improving the theoretical side of EM research. With the changes to the main task and the inclusion of the real-life condition, it was imperative to ensure that the collected data truly represents EM. Pause et al. (2013) proposed seven criteria for a good test of EM. All of the testing should happen in a controlled laboratory setting i) without any explicit instruction to memorise any information; ii) an unusual (thus arousing) task iii) should be based on one-trial learning events iv) producing the needed WWW information about the events v). Finally, the memory test should be unexpected vi) and should test relatively long-term memory vii). The fulfilment of the criteria should increase the robustness of both theoretical findings regarding EM and, in the case of the present experiment, the methodological differences between the VR settings.

To better fulfil the criteria and to overcome the earlier discussed issues, the present experiment differed considerably from the previous experiments. First of all, the cued WWW recall, as used in all previous experiments, was redone to be a guided free recall. This was done to combine the free recall and the WWW tasks. The free recall task was used in all of the previous experiments as it has been argued in the literature that the internal cueing should lead to a more episodic recall of information and recall that is more reflective of real-life behaviour (Tulving, 1972). As such, the What part of the WWW test did not contain any cues, such as the name of the object, which meant that participants had to use internal cues to recall information (in a similar fashion to Mazurek et al., 2015).

Due to the addition of the real-life condition, a new avenue of exploration was included in the present experiment – the temporal component of EM. As briefly discussed in Chapter 1 (section 1.4.1), both When and Which can be used as the temporal components of EM. While When is used more for temporal dating of information, Which is about the spatio-temporal context in which the event happened (Easton & Eacott, 2008; Friedman, 1993). Experiments in the present thesis so far have only looked at the serial order When (e.g. "was the object in the first, second or third room?"), as it closely followed Tulving's original definition of EM and due to all environments being virtual. It has been argued that the context information is more useful for the recall of an episode than the recall of serial information. This is due to context information providing more cues for the recall of the event than the recall of the temporal order in which the event happened (Easton & Eacott, 2008). Additionally, When can be answered using semantic knowledge and reasoning. For example, if a participant from Experiments 1, 2 or 3 remembered that the first room was a bedroom and the last room was a kitchen they could deduce that the second room had to be the workroom. Such reasoning regarding the second room would not involve EM (Clayton et al., 2003). Following this logic, the presently used Which question would be identical to the source memory test as discussed in Chapter 1 (section 1.4.5). The source memory test is based on data showing that memory for focal elements and memory for contexts differ and are due to focal factual (item) information being more important and thus more likely to be encoded than the source information independently (Johnson & Raye, 1981; Schacter et al., 1984; Shimamura & Squire, 1987). The inclusion of the real-life setting in the present experiment gave rise to an interesting question: which setting is better integrated into an episode – real-life, HMD-VR or Desktop-VR? As such, the present experiment explored which of the three settings would lead to better integration of contextual (Which) information. This was achieved by asking participants to provide both When and Which information resulting in a final What-When-Where-Which combination.

There were two main predictions. First, it was predicted that there will be an overall higher proportion of recalled combined WWWW information when compared to the traditional combined WWW information (as used in the previous experiments). Second, it was predicted that memory performance measured through WWWW, object recognition and the detail task will be highest in the real-life condition following HMD-VR, with the Desktop-VR having the lowest performance. The reasoning behind the first prediction was based on the earlier mentioned research suggesting that Which information provides more useful information for the retrieval of EM (Easton & Eacott, 2008). Higher WWWW performance, compared to the traditional WWW, would indicate that when the full WWW information is recalled about an episode, the knowledge in which context the episode occurred is also automatically recalled. On the other hand, better WWW performance, compared to WWWW, would suggest that the recall of When information does not lead to the automatic recall of the specific context and that Which information is not as easily retrieved. The second prediction is based on literature showing that memory in a real-world setting is higher than in HMD-VR or Desktop-VR (Flannery & Walles, 2003; Hoffman et al., 2001; Waller et al., 1998). While Experiment 3 showed no differences in EM performance between HMD-VR and Desktop-VR, it was predicted that the combined WWWW proportions, number of Remember judgements and d’ object recognition scores will be higher in HMD-VR compared to Desktop-VR setting. This prediction was based on the discussed changes in tasks in the present experiment and the literature discussed in the previous chapters showing better performance in the HMD-VR compared to Desktop-VR (Harman et al., 2017; Krokos et al., 2019; Mania et al., 2003).

The present experiment also continued using the Remember/Know/Guess (R/K/G) paradigm in the exploration of EM. Experiment 3 data did not support the prediction that HMD-VR would have higher proportion of Remember judgements over Desktop-VR. On the opposite, the data showed more Remember judgements in Desktop-VR. The result was mainly explained as a statistical chance due to the high p-value (p=.036) and the literature showing support to the initial prediction (Mania & Chalmers, 2001). Due to this and the multiple changes to the experimental design, the R/K/G judgements were also recorded in the present experiment. The prediction regarding the HMD-VR and Desktop-VR differences remained the same with more Remember judgements predicted to be observed in the HMD-VR setting. However, the prediction regarding the added real-life condition is more difficult. Due to the lack of research exploring R/K/G judgements between the three settings, the prediction was based on the general research showing better memory performance in real-life over any VR settings (Flannery & Walles, 2003; Hoffman et al., 2001; Waller et al., 1998; for a review see Smith, 2019). As such it was predicted that there would be higher proportion of Remember judgements than HMD-VR or Desktop-VR.

The present experiment also looked at the vividness ratings of participants' recalled information. As discussed in Chapter 1 (section 1.5) EM is characterised as being sensory-perceptual in nature (Conway, 2001, p. 1375). As such, perceptual details play an important role in EM. The Remember judgements are associated with episodic retrieval of information which should lead to detailed 're-experiencing' of the episode (Cassel et al., 2012; Markowitsch & Staniloiu, 2011; Suddendorf & Corballis, 2007; Tulving, 1985). What is more, hippocampal damage, which negatively affects EM, also makes subjective ratings of vividness either not consistent with objective scores of vividness (Kwan et al., 2010) or even correlates negatively with them (Addis et al., 2007). Due to these links, the present experiment also aimed to explore the relationship between the R/K/G judgements and vividness. Participants were asked to rate on a scale from 1 to 5 how vividly they re-experienced the WWWW information. Such rating is based on the Vividness of Visual Imagery Questionnaire (Marks, 1973).

The design used in the present experiment is very similar to one employed by Smulders and colleagues (Mazurek, Bhoopathy, Read, Gallagher, & Smulders, 2015; Smulders et al., 2017a, also see Holland & Smulders, 2011). In their research, participants performed a real-life task hiding eight objects on two occasions, in eight different locations around a room. This resulted in information about unique and arousing episodes and their spatio-temporal contexts which can easily be compared to the event objects used so far in the present thesis. The present experiment used both the task and the real-life condition from Mazurek, Bhoopathy, Read, Gallagher, & Smulders, 2015; Smulders et al., 2017a and Holland & Smulders, 2011, while also maintaining the HMD-VR and Desktop-VR conditions.

As mentioned before, the underlying idea for the thesis was that experiences in HMD-VR are closer to real-life than Desktop-VR. This 'closeness' relating to the discussed ability to use locomotion, the effect of enactment, the level of immersiveness and the sense of presence in HMD-VR. With the introduction of the real-life setting, it is possible to test this hypothesis. As discussed both here and in the previous chapter, it has been found that HMD-VR leads to a better transfer of knowledge and better memory retrieval than Desktop-VR. One of the arguments for this is HMD-VR having higher levels of immersion and sense of presence (Amin et al., 2016; Boyd, 1997; Gutiérrez et al., 2007; for a review see Mestre & Vercher, 2011 and Smith, 2019). The present experiment aimed to explore the relationships between the levels of presence and EM. While the majority of research shows that HMD-VR leads to higher levels of presence than Desktop-VR (Buttussi & Chittaro, 2018; Seibert & Shafer, 2018; Shu, Huang, Chang, & Chen, 2019; but see Mania & Chalmers, 2001) there is lack of research in its impact on EM. It was predicted that higher the level of presence would positively correlate with the WWWW EM performance with HMD-VR having higher overall presence score than Desktop-VR.

One aim that was part of all of the previous experiments was a consolidation period containing night's sleep. In all three of the experiments, there was a 24h consolidation period filled with sleep. However, Experiments 1 and 2 showed that there was no effect of time between learning and sleep on EM and Experiment 3 showed no significant interactions between HMD-VR and Desktop-VR and sleep on EM. Additionally, as discussed in Chapter 1 (section 1.7) and as one of the underlying notions behind the present thesis, it is important to test memory after some time and not just straight after encoding. As such, and following the earlier mentioned criteria (vi) by Pause et al. (2013) of testing long-term (>60min) memory, the present experiment had memory tests taking place after an hour-long retention period. As a result, the present experiment went back to the main aim of the thesis – exploring the viability of HMD-VR in EM research.

To summarise, the present study had one main objective: to investigate how EM performance differed in HMD-VR, Desktop-VR and real-life settings. The secondary aims were to explore if source information obtained from the Which component can be used to improve understanding and testing of EM and to explore the relationship between the level of presence and the EM performance.

6.2. Method

6.2.1. Participants

Participants in this experiment were 25 students from Bishop Grosseteste University and members of the general public (mean age = 25.20; SD = 7.46; range = 19-50; female = 14). All participants received a £10 Amazon voucher for participating in the experiment. Undergraduate psychology participants also received course credit. There were 2 students and 23 non-students. The non-students included members from the general population and students from other universities. All participants had normal or corrected to normal vision. The participant screening procedure was identical to the one used in Chapter 2 (section 1.1).

6.2.2. Materials

Virtual environments (VEs) were created using Unreal Engine software. VE for HMD-VR were presented through an HTC Vive system while the Desktop-VR environment was presented on a computer screen (size – 23in; resolution – 1980x1080). For the real-life environment, a 2.5m x 3m room was used in the Bishop Grosseteste University. A more in-depth description of the equipment and VEs is provided in Chapter 2 (section 1.2 and 1.3).

6.2.3. Design

The experiment contained three tests: WWWW, object recognition and detail. The tests were performed 1 hour after finishing all the tasks in real-life, HMD-VR and Desktop-VR settings. The setting was a within-subject independent variable.

There were five dependent variables in the WWWW test: the What, the Where, the Which, the When and the combined WWWW proportion. After every component, participants were asked to indicate if they remembered, knew or were just guessing regarding that information. This resulted in three dependent variables: the proportion of Remember, Know and Guess judgements. Finally, participants had to provide an overall rating of vividness (on a scale from 1-5) associated with the recall of the particular object or location. See Figure 6.1 for the task screen that was visible to the participants.

The object recognition task resulted in four dependant variables: the d' sensitivity index and the confidence rating, the proportion of correctly recalled When and the proportion of correctly recalled Which.

6.2.4. Procedure

All participants were met in a room separate to those used all learning environments (Real-life, HMD-VR, Desktop-VR). Participants were told that they will be hiding objects in three different rooms and then completing a number of questionnaires afterwards. Depending on the given order, participants were taken to the real-world experimental room or the laboratory with the HMD-VR/Desktop-VR equipment. Before starting the main task in HMD-VR and Desktop-VR settings, participants had to perform similar tasks in the specially created training VEs. For more information regarding the training rooms see Chapter 2 (section 2.3.1), or for a more in-depth description of the procedure see Chapter 2 (section 2.4).

After finishing the training rooms, participants were shown the main experimental rooms. The four objects that had to be hidden were presented in the middle of each room mixed with two additional objects that were not asked to be hidden. After hiding the four objects, the two remaining objects were taken away and another group of objects were presented resulting in eight objects from two groups hidden in each of the three rooms. After finishing each environment participants had a two-minute break.

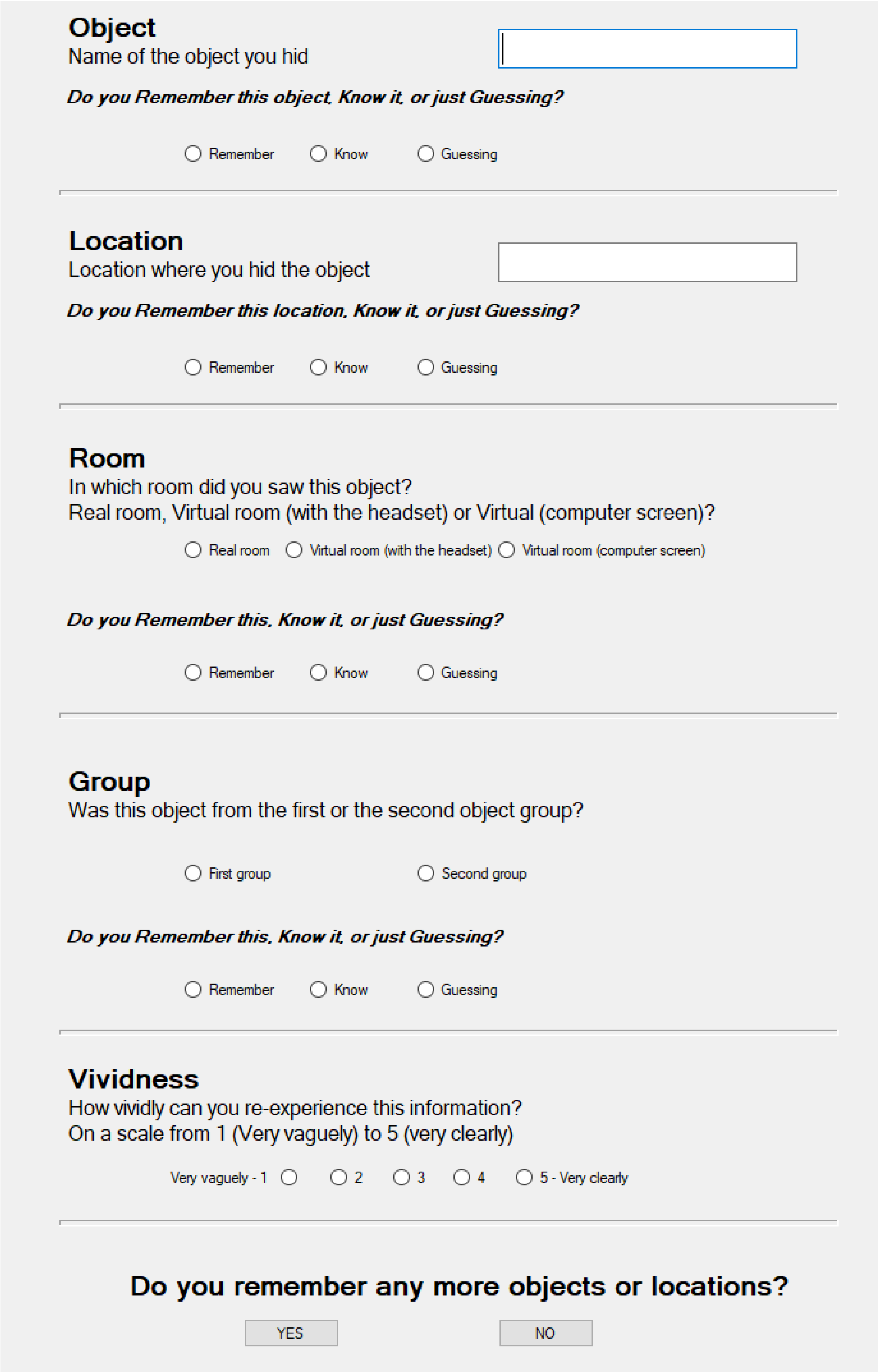
After finishing all three environments, participants were taken to the room they were initially greeted in and were asked to fill in the Igroup Presence Questionnaire (IPQ) (Igroup Project Consortium, 2015) and Presence Questionnaire (PQ) (Witmer & Singer, 1998) questionnaires. Following this, participants were told that they can leave the room and to come back to the same room after one hour. After participants came back they were asked to complete the WWWW, object recognition and detail tests. All of the tests were performed on a computer in the room where the participants were initially met.

6.2.4.1. The WWWW test

Participants were asked to recall as many objects and the location in which they had to hide them in (see Figure 6.1 for the task screen that was visible to the participants). Participants were told that if they can only remember an object but not the location or vice versa, they should still write down the object or the location and leave the other field blank. The participants were then asked to identify in which setting and in which group the object, the location or both belonged to. If an object, location or both were recalled, participants had to indicate if they remembered, knew or were guessing regarding the object, the location, the group and the room. After filling this information, participants were asked to give an overall rating (on a scale of 1 to 5) on how vividly they recalled the information. Participants repeated this process until they could no longer recall any more objects or locations. If they indicated that they could recall more objects or locations, the screen was cleared of the previous information and participants were able to repeat the procedure.

Figure 6.1.

Experiment 4: The WWWW task screen.



6.2.4.2. The Recognition task

The recognition task followed a similar procedure as in the previous experiments (See Chapter 2, section 1.5.3). Participants were presented with 72 pictures of objects, one at a time. Thirty-six of those objects were objects that were presented in the middle of each room throughout the experiment: 24 of which were the objects that had to be hidden and 12 objects which were not interacted with. The remaining 36 objects were lures and were not in any of the rooms (the list of all the objects can be found in Appendix B). In addition to the yes/no recognition and confidence ratings, participants had to indicate in which group and in which room the recognised object was in. This was done only if the participant indicated that they have recognised the object. Additionally, participants had to provide remember/know/guess judgements for the object recognition, group and room information.

6.2.4.3. Detail test

Participants were given a Microsoft Word file with three fields – one for each room. The task was to provide any objects and/or details that they have not mentioned in the WWW task. However, in the interests of brevity, the Detail task was omitted as it was not considered useful in exploring EM.

6.3. Results

6.3.1. Combined What-Where-Which-When

When a participant correctly recalled the object (What), the location in which they had to hide that object (Where), the setting in which the object was being hidden (Which) and in which of the two groups the object was in (When) it was said that the participant recalled the full combined WWWW information. The combined WWWW score is the proportion of correctly recalled WWWW information out of all objects that had to be hidden (ranging from 0 to 1). For example, a score of 0.5 (i.e., 50%) would mean that a participant recalled combined WWWW information for 12 objects out of 24 possible. The means and standard deviations for each of the settings are presented in Table 6.1.

Table 6.1.

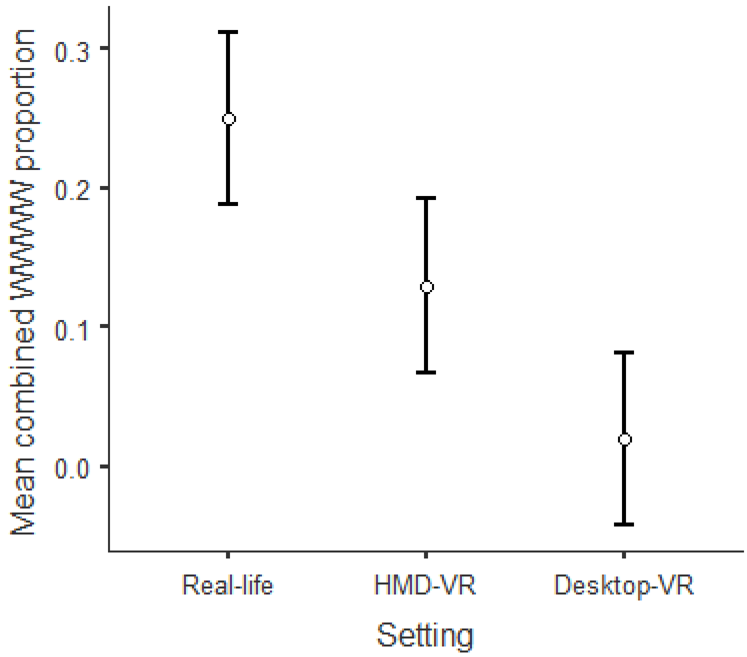
Experiment 4: Mean proportions of recalled combined WWWW information.

Setting	Mean	SD
Real-life	.25	.18
HMD-VR	.13	.20
Desktop-VR	.02	.05

A statistically significant difference in the performance on the combined WWWW measure was observed between the three settings, $F(2,48)=16.48$, $p<.001$ (see Figure 6.2). The WWWW performance was higher in the real-life setting than in the HMD-VR, $t(48)=2.99$, $p=.013$, or Desktop-VR settings, $t(48)=5.74$, p<.001, and the performance was higher in the HMD-VR setting than in the Desktop-VR setting, $t(48)=2.75$, $p=.025$.

Figure 6.2.

Experiment 4: Mean proportions of recalled combined WWWW information.



Notes. Error bars represent 95% confidence intervals.

6.3.2. Separate What-Where-Which-When components

As in the previous chapters, the separate WWWW components were analysed. As previously, the data was transformed into proportions (ranging from 0 to 1). For example, a score of 0.5 for What component would mean that a participant recalled What information for 12 objects out of 24 possible. The means and standard deviations for each of the settings are presented in Table 6.2.

Table 6.2.

Experiment 4: Mean proportions of recalled separate WWWW components in the real-world, HMD-VR and Desktop-VR settings.

	Setting		
Component	Real-life	HMD-VR	Desktop-VR
What	0.45 (.50)	0.3 (.46)	0.19 (.39)
When	0.32 (.49)	0.25 (.45)	0.13 (.34
Where	0.45 (.50)	0.21 (.41)	0.14 (.35)
Which	0.57 (.50)	0.31 (.46)	0.15 (.35)

Note. Numbers in parentheses represent standard deviations.

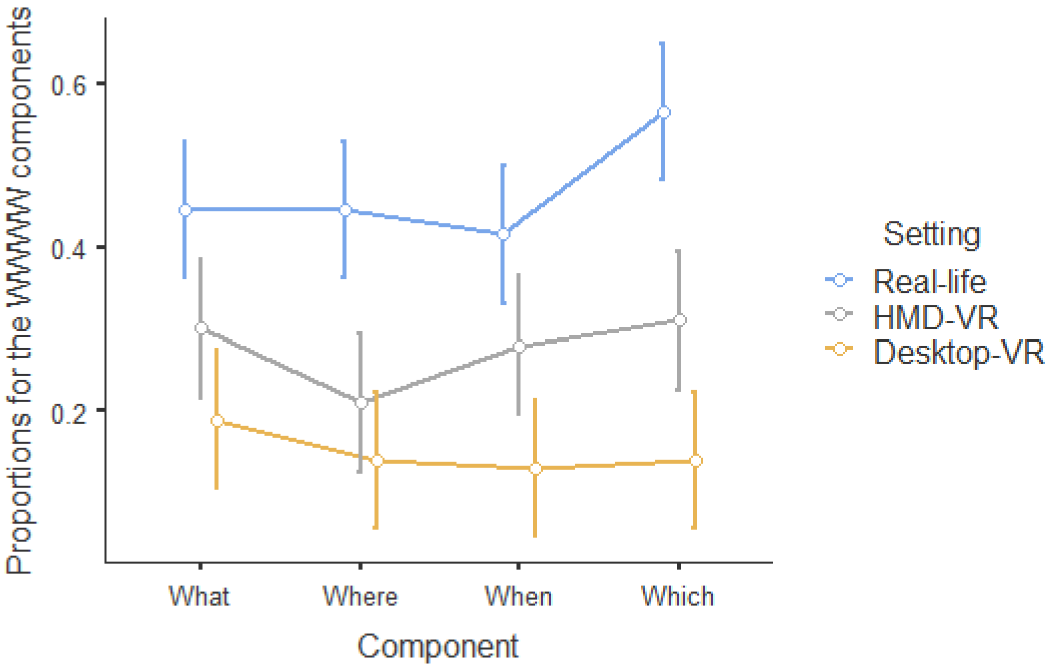
Analysis of each WWWW component separately showed an effect of setting, $F(2,48)=24.11$, $p<.001$ (see Figure 6.3). Overall, proportions for all of the components were higher in the real-life setting than in the HMD-VR, $t(48)=4.07$, $p<.001$, or Desktop-VR settings, $t(48)=6.91$, $p<.001$ and higher in the HMD-VR setting than in the Desktop-VR setting, $t(48)=2.84$, p=.020.

The effect of Component was significant, $F(2,48)=26.59$, $p<.001$. Multiple comparisons revealed higher Which proportions than Where, $t(72)=3.92$, $p=.001$, or When, $t(72)=3.39$, $p=.007$. None of the other comparisons were significant, *ts*<2.5, *ps*>.089.

A significant interaction was found between the Components and the three Settings, $F(6,144)=4.63$, $p<.001$ (see Figure 6.3). As in the previous experiments, further pairwise comparisons were done to explore the differences in the separate component recalls between the three settings (See Table 6.3 for the post-hoc comparisons).

Figure 6.3.

Experiment 4: Mean proportions of recalled separate WWWW components in the real-world, HMD-VR and Desktop-VR settings



Notes. Error bars represent 95% confidence intervals.

Table 6.3.

Experiment 4: Bonferroni corrected comparisons for the component recall between the three settings in Experiment 4.

Component	Setting		Setting	Mean Difference	SD	df	t	*p*
What	Real-life	-	HMD-VR	.15	.25	80.50	2.88	.336
	Real-life	-	Desktop-VR	.26	.25	80.50	5.07	< .001
	HMD-VR	-	Desktop-VR	.11	.25	80.50	2.19	1
Where	Real-life	-	HMD-VR	.24	.25	80.50	4.67	< .001
	Real-life	-	Desktop-VR	.31	.25	80.50	6.06	< .001
	HMD-VR	-	Desktop-VR	.07	.25	80.50	1.39	1
When	Real-life	-	HMD-VR	.14	.25	80.50	2.68	.586
	Real-life	-	Desktop-VR	.29	.25	80.50	5.66	< .001
	HMD-VR	-	Desktop-VR	.15	.25	80.50	2.98	.251
Which	Real-life	-	HMD-VR	.26	.25	80.50	5.07	< .001
	Real-life	-	Desktop-VR	.43	.25	80.50	8.44	< .001
	HMD-VR	-	Desktop-VR	.17	.25	80.50	3.38	.075

6.3.3. Incomplete What-Where-Which-When combinations

For exploration purposes, the WWWW data were broken down by the number of recalled incomplete combinations (similarly to Mazurek et al., 2015), for example, if a participant recalled What-Where-When but not Which or if a participant only recalled the What and nothing else. The data was transformed into proportions (ranging from 0 to 1). The data was also separated into 5 within-subject levels leading to one full combined WWWW proportion and four levels of incomplete combinations. This resulted in the complete WWWW proportion, proportions containing combinations of three components, proportions containing combinations of two components, proportions containing combinations of two components and a proportion of no recalled information (see Table 6.4). For example, a score of 0.5 for a What-Where-Which combination would mean that out of all of the combinations the participant recalled that were not full WWWW, half of them (0.5) were What-Where-Which. The means and standard deviations for all of the incomplete WWWW combinations in each of the three settings can be seen in Table 6.4.

Table 6.4.

Experiment 4: The What-Where-Which-When combination proportions

	Setting		
Combinations	**Real-life**	**HMD-VR**	**Desktop-VR**
What-Where-When-Which	.25 (.18)	.13 (.20)	.02 (.05)
What-Where-When	0	0	.01 (.04)
What-Where-Which	.07 (.11)	.02 (.04)	.03 (.06)
Where-When-Which	0	0	.01 (.03)
What-When-Which	.06 (.06)	.09 (.14)	.04 (.07)
What-Where	0	0	.03 (.09)
What-When	0	.02 (.04)	.02 (.05)
What-Which	.04 (.07)	.02 (.08)	.01 (.04)
Where-Which	.11 (.18)	.04 (.08)	.03 (.06)
Where-When	0	0	0
What	.04 (.07)	.03 (.06)	.04 (.06)
Where	.02 (.05)	.04 (.07)	.03 (.05)
Nothing recalled	.44 (.28)	.65 (.25)	.77 (.17)

The incomplete WWWW pairs were analysed using repeated-measures ANOVAs. Only the three and two component pairs and the individual What and Where proportions were analysed. The complete WWWW combinations were not analysed as they was already explored in section 6.3.1 and proportions of no recall were not analysed as the interest was in the differences in the incomplete combinations. Additionally, as the main interest was in the differences in the incomplete combinations the pairwise comparisons were performed only for the combinations and not settings.

The effect of Combination was significant for combinations with three, $F(3,72)=11.83$, $p<.001$, and two, $F(4,96)=7.76$, $p<.001$, components but not singular What and Where components, $F(1,24)=.89$, $p=.356$. The multiple pairwise comparisons are presented in Table 6.5.

Table 6.5.

Experiment 4: Multiple comparison table for the incomplete combinations

Combination		Combination	Mean Difference	SD	df	t	*p*
What-Where-When	-	What-Where-Which	-.03	.06	72	-2.76	.044
	-	Where-When-Which	.00	.06	72	.15	1
	-	What-When-Which	-.06	.06	72	-4.94	< .001
What-Where-Which	-	Where-When-Which	.03	.06	72	2.90	.029
	-	What-When-Which	-.03	.06	72	-2.18	.196
Where-When-Which	-	What-When-Which	-.06	.06	72	-5.08	< .001
What-Where	-	What-When	-.00	.06	96	-.29	1
	-	What-Which	-.01	.06	96	-1.15	1
	-	Where-Which	-.05	.06	96	-4.32	< .001
	-	Where-When	.01	.06	96	.72	1
What-When	-	What-Which	-.01	.06	96	-.87	1
	-	Where-Which	-.05	.06	96	-4.03	.001
	-	Where-When	.01	.06	96	1.01	1
What-Which	-	Where-Which	-.04	.06	96	-3.17	.020
	-	Where-When	.02	.06	96	1.87	.641
Where-Which	-	Where-When	.06	.06	96	5.04	< .001

6.3.4. The Remember/Know/Guess judgements

A repeated-measures ANOVA was used to analyse the Remember/Know/Guess judgements for the What, Where, Which and When components. The judgement data was transformed into a mean number of each judgement given for one recalled object. The judgements were transformed into overall proportions in a similar fashion to Dewhurst, Conway, & Brandt (2009). For example, adding one participant's Remember, Know and Guess judgement proportions at Baseline would equal to 1. This transformation was undertaken so that the lower number of recalled objects on the second session would not affect the judgement data. The means and standard deviations of the R/K/G judgements are presented in Table 6.6

Table 6.6.

Experiment 4: The means and standard deviations of the Remember/Know/Guess judgements given for one recalled object during the WWWW task

	Judgement		
Setting	**Remember**	**Know**	**Guess**
Real-life	.64 (.24)	.14 (.20)	.18 (.16)
HMD-VR	.48 (.31)	.13 (.19)	.23 (.22)
Desktop-VR	.41 (.36)	.10 (.15)	.29 (.31)

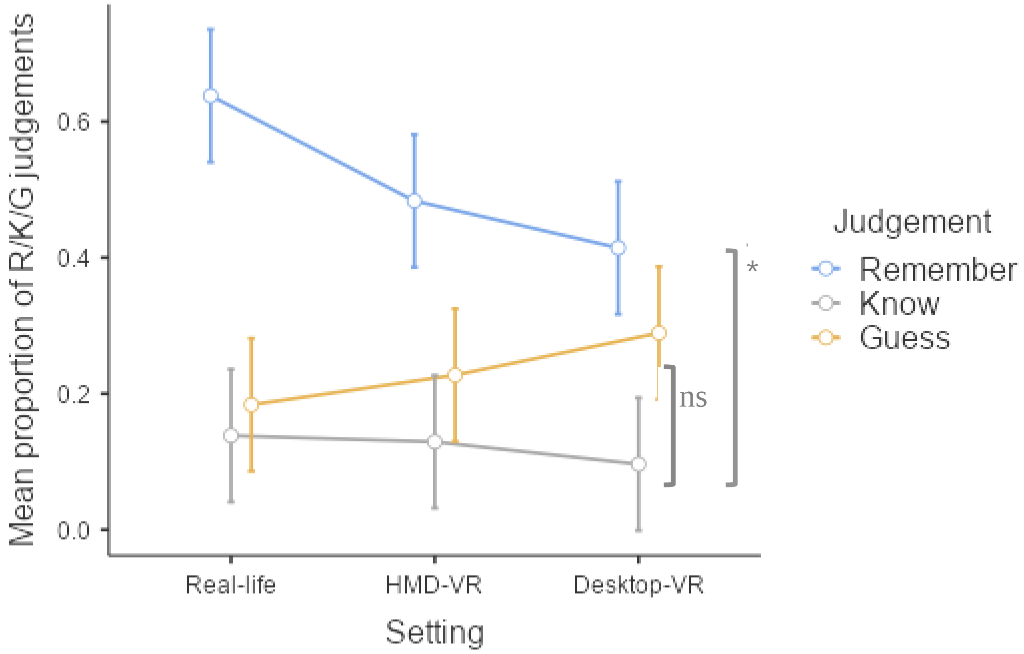
Note. Numbers in brackets indicate standard deviations

The effect of Setting was not significant, $F(2,48)=1.47$, $p=.240$. The effect of Judgement was significant, $F(2,48)=26.98$, $p<.001$. Multiple comparisons revealed that participants gave more Remember judgements than Know, $t(48)=7.13$, $p<.001$, or Guess judgements, $t(48)=5.09$, $p<.001$. The proportions of Know and Guess judgements did not differ, $t(48)=-2.05$, $p=.138$.

A significant Setting x Judgement interaction was detected, $F(4,96)=3.36$, $p=.013$ (see Figure 6.4). The pairwise comparisons can be seen in Table 6.7. For clarity, only the significant comparisons are shown.

Figure 6.4.

Experiment 4: The mean number of judgements given for one recalled object during the WWWW task



Notes. Error bars represent 95% confidence intervals. * indicates a significant difference (p<.05), ns indicates non-significant difference (p>.05)

Table 6.7.

Experiment 4: Pairwise comparison table for the Setting x Judgement interaction between Setting and Judgement in the Remember/Know/Guess data

Setting	Judgement		Setting	Judgement	Mean Difference	SD	df	t	*p*
Real-life	Remember	-	Real-life	Know	.50	.38	125	6.56	< .001
		-	Real-life	Guess	.45	.38	125	5.97	< .001
		-	HMD-VR	Know	.51	.37	129	6.89	< .001
		-	HMD-VR	Guess	.41	.37	129	5.57	< .001
		-	Desktop-VR	Remember	.22	.31	142	3.61	.015
		-	Desktop-VR	Know	.54	.37	129	7.34	< .001
		-	Desktop-VR	Guess	.35	.37	129	4.73	< .001
	Know	-	HMD-VR	Remember	-.35	.37	129	-4.68	< .001
		-	Desktop-VR	Remember	-.28	.37	129	-3.74	.01
	Guess	-	HMD-VR	Remember	-.30	.37	129	-4.06	.003
HMD-VR	Remember	-	HMD-VR	Know	.35	.38	125	4.65	< .001
		-	HMD-VR	Guess	.26	.38	125	3.36	.037
		-	Desktop-VR	Know	.39	.37	129	5.25	< .001
	Know	-	Desktop-VR	Remember	-.29	.37	129	-3.86	.006
Desktop-VR	Remember	-	Desktop-VR	Know	.32	.38	125	4.18	.002

Note. For clarity, only the significant comparisons are shown.

6.3.5. The WWWW vividness scores

After filling in the WWWW information about an object, participants were asked to give an overall rating (on a scale of 1 to 5) on how vividly they recalled that information. The data was transformed to show mean vividness rating for each participant in each of the three settings. The means and standard deviations can be seen in Table 6.8.

Table 6.8.

Experiment 4: Means and standard deviations for vividness scores

Setting	Mean
Real-world	3.34 (.94)
HMD-VR	2.68 (1.01)
Desktop-VR	2.44 (1.02)

Note. Numbers in brackets indicate standard deviations

The effect of Setting was significant, $F(2,32)=5.29$, $p=.01$. Mean vividness ratings did not statistically differ between the real-world and HMD-VR settings, $t(32)=2.30$, $p=.084$. The real-world ratings were significantly higher than Desktop-VR ratings, $t(32)=3.14$, $p=.011$. HMD-VR and Desktop-VR ratings did not differ, $t(32)=.84$, $p=1$.

As the vividness score was added to explore its relationship to the R/K/G judgements a correlation matrix was created (See Table 6.9).

Table 6.9.

Experiment 4: Correlation matrix for between Vividness scores and the number of R/K/G judgements

Measure	1	2	3	4
1. Vividness	—			
2. Remember judgements	.50***	—		
3. Know judgements	.08	-0.54***	—	
4. Guess judgements	-.40***	-0.46***	-0.24***	—

Notes. *p < .05. **p < .01. ***p < .001

6.3.6. Object recognition – d' scores

Recognition data was converted to d' scores (Z(hit rate) − Z(false alarm rate)). The means and standard deviations of the d' scores can be seen in Table 6.10.

Table 6.10.

Experiment 4: The means and standard deviations for the d' scores for each of the three settings

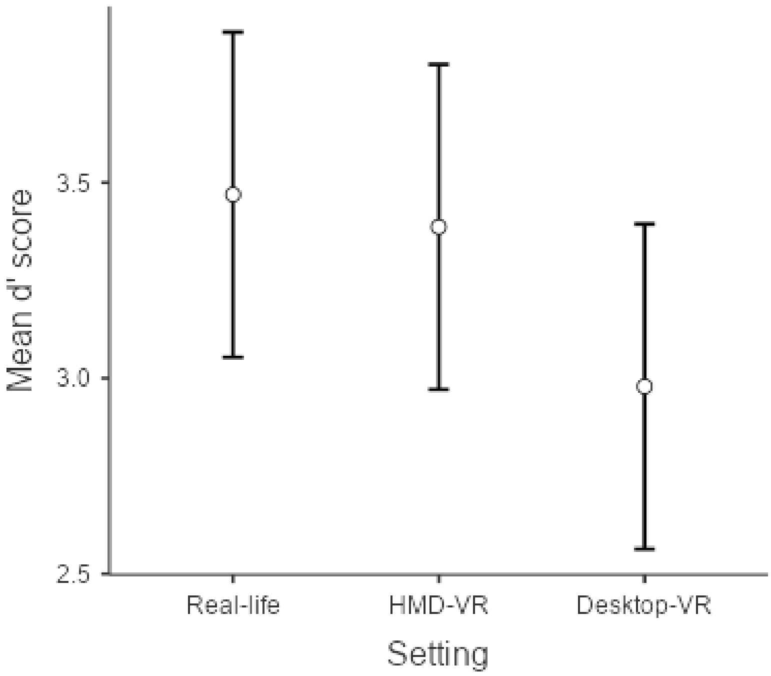
Setting	d'
Real-life	2.19 (1.65)
HMD-VR	1.78 (1.90)
Desktop-VR	1.44 (1.87)

Note. Numbers in brackets indicate standard deviations

The effect of Setting was significant, $F(2,48)=5.22$, $p=.009$ (see Figure 6.5). The d' scores did not differ between the real-life and HMD-VR settings, $t(48)=.51$, $p=.1$. The d' scores were higher in the real-life setting than in Desktop-VR settings, $t(48)=3.02$, $p=.012$. The d’ scores were higher in HMD-VR setting over Desktop-VR setting, $t(48)=2.51$, $p=.047$.

Figure 6.5.

Experiment 4: Mean d’ scores from the object recognition task.



Notes. Error bars represent 95% confidence intervals.

6.3.7. Object recognition – Confidence ratings

During the object recognition task, participants had to indicate how confident they were with their decisions (ranging from 0 to 1). The means and standard deviations for the confidence ratings can be seen in Table 6.11.

Table 6.11.

Experiment 4: Mean object recognition confidence ratings for each of the three settings

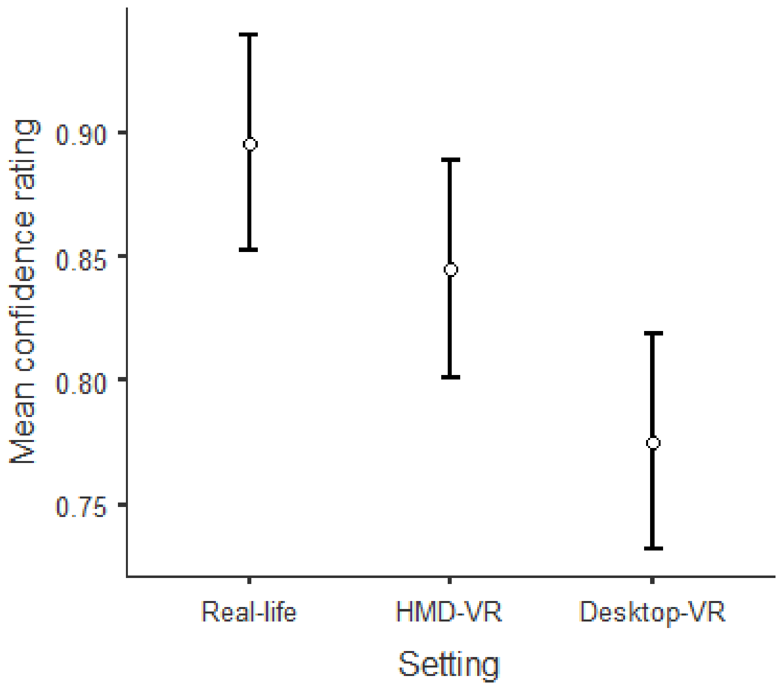
Setting	Confidence rating
Real-life	.90 (.20)
HMD-VR	.85 (.21)
Desktop-VR	.78 (.19)

Note. Numbers in brackets indicate standard deviations

The effect of Setting was significant, $F(2,48)=11.10$, $p<.001$, (see Figure 6.6). The confidence ratings did not differ between the real-life and HMD-VR settings, $t(48)=1.99$, $p=.157$, but were higher than in Desktop-VR setting, $t(48)=4.69$, $p<.001$. The confidence ratings in HMD-VR setting were higher than in Desktop-VR, $t(48)=2.70$, $p=.029$.

Figure 6.6.

Experiment 4: Mean object recognition confidence ratings in each of the three settings.



Notes. Error bars represent 95% confidence intervals.

6.3.8. Object recognition – When/Which

After recognising an object, participants were asked to provide information on the group (related to When component - first or second) and setting (related to Which component – real-life, HMD-VR or Desktop-VR) the object was in. This was done to further explore the differences between the temporal When and contextual Which information. The means and standard deviations for the two measures can be seen in Table 6.12.

Table 6.12.

Experiment 4: Mean correctly recalled proportions for Which and When judgements

Setting	When proportion	Which proportion
Real-life	.69 (.17)	.77 (.20)
HMD-VR	.60 (.24)	.66 (.25)
Desktop-VR	.49 (.19)	.39 (.32)

Note. Numbers in brackets indicate standard deviations

The effect of Setting was significant, *F*(1,24)=38.31, *p*<.001. Overall, the Which and When proportions were higher in the real-life setting than in the HMD-VR, *t*(48)=3.00, *p*=.013, or Desktop-VR settings, *t*(48)=8.62, *p*<.001, and the proportions were higher in the HMD-VR setting than in the Desktop-VR setting, *t*(48)=5.62, *p*<.001. The effect of Component was not significant, *F*(1,24)=.05, *p*=.826. There was a significant Component x Setting interaction, *F*(2,48)=4.57, *p*=.015 (See Figure 6.7). Pairwise comparisons can be seen in Table 6.13.

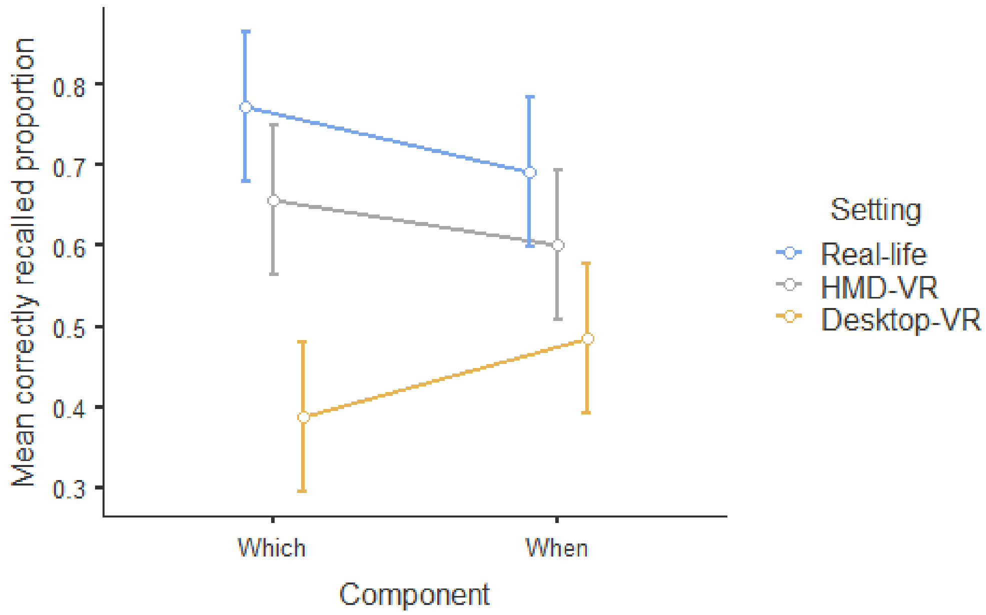
Table 6.13.

Experiment 4: The pairwise comparisons for the Component x Setting interaction

Component	Setting		Component	Setting	Mean Difference	SE	df	t	*p*
Which	Real-life	-	Which	HMD-VR	.12	.24	95.60	2.46	.237
		-	Which	Desktop-VR	.39	.24	95.60	8.20	< .001
		-	When	Real-life	.08	.34	44.20	1.18	1
		-	When	HMD-VR	.17	.35	47.70	2.46	.262
		-	When	Desktop-VR	.29	.35	47.70	4.12	.002
	HMD-VR	-	Which	Desktop-VR	.27	.24	95.60	5.75	< .001
		-	When	Real-life	-.04	.35	47.70	-.50	1
		-	When	HMD-VR	.06	.34	44.20	.81	1
		-	When	Desktop-VR	.17	.35	47.70	2.47	.260
	Screen	-	When	Real-life	-.31	.35	47.70	-4.38	< .001
		-	When	HMD-VR	-.21	.35	47.70	-3.07	.052
		-	When	Desktop-VR	-.10	.34	44.20	-1.43	1
When	Real-life	-	When	HMD-VR	.09	.24	95.60	1.93	.852
		-	When	Desktop-VR	.21	.24	95.60	4.39	< .001
	HMD-VR	-	When	Desktop-VR	.12	.24	95.60	2.46	.234

Figure 6.7.

Experiment 4: Mean correctly recalled proportions for Which and When judgements



Notes. Error bars represent 95% confidence intervals.

To better understand the significant difference between Desktop-VR and HMD-VR Which recalls a table was created showing the participant indicated settings and the real object settings (see Table 6.14).

Table 6.14.

Experiment 4: Participant indicated settings and the real object settings for the object recognition Which recalls

	Actual object setting		
Participant indicated setting	Real-life	HMD-VR	Desktop-VR
Real-life	180	11	14
HMD-VR	7	134	46
Desktop-VR	7	65	85

6.3.9. Presence data

The data from IPQ (Igroup Project Consortium, 2015) and PQ (Witmer & Singer, 1998) questionnaires were combined to create one overall score for that particular questionnaire. For the analysis, only data from the HMD-VR and Desktop-VR settings were used. Data from the real-life setting was not used as it has been shown that such comparisons can be invalid. For instance, it has been shown that when a participant is asked to estimate their ‘sense of being there’ in a real-life setting, they might interpret the question in a way to make it seem sensible. Participants might give a lower than the maximum rating even if they are actually there (Usoh et al., 2000). The means and standard deviation can be seen in Table 6.15.

Table 6.15.

Experiment 4: Mean IPQ and PQ scores for HMD-VR and Desktop-VR settings

	Questionnaire	
Setting	IPQ	PQ
HMD-VR	5.10 (.76)	5.52 (.64)
Desktop-VR	2.96 (.82)	4.16 (1.11)

Note. Numbers in brackets indicate standard deviations

There was a difference in IPQ scores between HMD-VR and Desktop-VR settings, $F(1,24)=87.0$, $p<.001$, with HMD-VR having higher mean scores than Desktop-VR, $t(24)=9.33$, $p<.001$.There was a difference in PQ scores between the two VR settings, $F(1,24)=34.3$, $p<.001$, with HMD-VR having higher mean scores than Desktop-VR, $t(24)=5.86$, $p<.001$.

A correlation matrix was created to explore the relationships between the two presence questionnaires and the number of full WWWW recalls per participant (see Table 6.16). None of the presence questionnaires correlated with the WWWW recalls (rs < .193, n = 50, *ps* > .180). There was a strong positive correlation between the two presence questionnaires (r = .738, n = 50, *p* < .001).

Table 6.16.

Experiment 4: Correlation matrix for the mean scores of the IPQ and PQ presence questionnaires and a number of full WWWW recalls

Measure	1	2	3
1. Combined WWWW proportions	—		
2. IPQ	.19	—	
3. PQ	.13	.74***	—

Note. * *p* < .05, ** *p* < .01, *** *p* < .001

6.4. Discussion

The main aim of the present chapter was to explore how EM differed between real-life, HMD-VR and Desktop-VR settings. The secondary aims were to explore if source information obtained from the Which component can be used to improve understanding and testing of EM and to explore the relationship between the level of presence and the EM performance.

The general prediction was that EM performance in the WWWW and object recognition tests would have been highest in the real-life setting following HMD-VR setting with Desktop-VR having the worst performance. This prediction was based on research showing real-life memory performance being higher than in HMD-VR (Flannery & Walles, 2003; Hoffman et al., 2001; Waller et al., 1998) and better memory performance in HMD-VR than in Desktop-VR (Cárdenas-Delgado et al., 2017; Harman et al., 2017; Krokos et al., 2019; Mania et al., 2003; Repetto et al., 2016). The results partly supported this prediction. The combined WWWW proportion data showed that EM performance was better in the HMD-VR setting compared to the Desktop-VR setting with the real-life setting having higher performance than both of the VR settings. However, the object recognition data showed real-life performance being higher than both of the VR settings with no difference in performance between HMD-VR and Desktop-VR

While it was predicted that the WWWW performance in the HMD-VR setting will be higher than in Desktop-VR, it is important to discuss the present findings in the light of the results from Experiment 3. The results from the previous experiment showed no difference in EM between the two settings, while the present results do show it. As discussed in the introduction, the change in the method was to be able to compare HMD-VR and Desktop-VR to real-life and to overcome behavioural issues (lack of exploration in the Desktop-VR) observed in Experiment 3. The combined WWWW results show that EM performance in HMD-VR is closer to the real-life than Desktop-VR but only if a more active and goal-oriented task is employed as it was done in the present experiment. Indeed, the change in task led to a difference between the two VR settings. However, it is important to point out that Experiment 3 used a between-subjects design to compare the VR environments, whereas the present experiment involved a more statistically powerful within-subjects design. The between-subjects design in Experiment 3 might have not had enough statistical power to find the differences observed in the present experiment.

The task used in the present experiment involved actively interacting with objects and locations; this led to higher WWWW performance in the HMD-VR performance, compared to the Desktop-VR, as the actions and interactions were closer to the real-life counterparts. This could have stemmed from the effect of enactment (Mohr, Engelkamp, & Zimmer, 1989; Zimmer & Engelkamp, 1989). As participants had to use the HMD-VR controllers to pick up and carry the objects as if they were real, this might have led to better memory performance as the action was closer to the real-world counterpart as compared to the Desktop-VR. As discussed in the introduction, research has shown that interactivity positively affects memory (for a review of enactment effect see Mohr, Engelkamp, & Zimmer, 1989; Zimmer & Engelkamp, 1989). Another, closely related reason for the better HMD-VR performance over the Desktop-VR might have been the ability to better inspect the objects. It has been shown that an active inspection (e.g. ability to pick up and inspect objects from various angles) can lead to better memory recall (Harman et al., 1999; James et al., 2002; Trewartha et al., 2015). As discussed in the introduction, this effect is based on the motor information that is combined with the memory of an object which facilitates both encoding and retrieval of memory (Mohr et al., 1989; Russ et al., 2003; Zimmer & Engelkamp, 1989). While participants in the Desktop-VR setting were able to get closer to the objects to better inspect them it was observed that they tended not to whereas in the HMD-VR setting participants tended to inspect the objects. This again shows how the differences in behaviour between the two VR settings that can affect EM. Overall, the combination of the two effects might have enhanced the EM encoding in HMD-VR as participants were able to gain more information about the objects. The present results suggest that more life-like interaction with objects through HMD-VR can lead to a more life-like memory performance. This finding directly relates to the main aim of the thesis to investigate how HMD-VR can be used to test EM in an ecologically valid fashion by showing that HMD-VR leads to more life-like EM performance than the more traditional Desktop-VR.

A number of interesting findings can be seen when looking at the separate WWWW components and also the incomplete combinations. When looking at the separate WWWW components it can be seen that HMD-VR performance did not differ to the real-life performance in the What and When components. However, HMD-VR performance was lower than real-life performance in Where and Which component data. However, HMD-VR performance showed no statistically significant differences to Desktop-VR in any of the components (see Table 6.2 and Figure 6.3). These results indicate that the What or factual information about an episode and the When or information on the serial order of events is recalled comparably well between the real-life and HMD-VR settings. While there is a lack of research supporting the What and When findings, the literature does support the lower Where and Which performance in HMD-VR setting. As discussed in the introduction, research has shown better memory performance in the real-life condition for source memory (Hoffman et al., 2001) and spatial memory (Waller et al., 1998) which corresponds to the Which and Where components.

While the incomplete WWWW combination data was only briefly explored it revealed a number of important findings. First of all, there was a very low proportion of What-Where-When combinations and only in the Desktop-VR setting. There were none in the real-life or HMD-VR settings (see Table 6.4). The low number of the traditionally used in research What-Where-When combinations over the What-Where-When-Which combinations can be interpreted in a way that participants who recalled the What-Where-When information automatically recalled the accompanying setting. However, looking deeper into the combination data a more interesting trend can be observed. The low proportions of Where-When, What-When and What-Where combinations have one thing in common – they are all combinations of components that are part of the traditional What-Where-When EM triad. The low proportions of these pairs mean that if those pairs were recalled, there was a high chance that another component was also recalled. When looking at the combinations of three components the only other combination apart from the traditional What-Where-When that shown low proportions was the Where-When-Which. As before, this means that when this combination was recalled, the remaining What component was also automatically recalled. In general, the incomplete combination data shows the importance of the Where-When pair. All incomplete combinations that included this pair had low proportions, indicating that when participants recalled the temporal and spatial, the remaining item and context information was also recalled.

When looking at other incomplete combinations an interesting finding is the difference between the What-Which and Where-Which proportions. There were more Where-Which combinations than What-Which combinations with more recalls of both pairs in the real-life condition over the two VR conditions. This shows that locations in the real-life setting were better recalled than the locations in HMD-VR or Desktop-VR. It also shows that this better location recall allowed participants to recall the correct setting associated with the location. It is possible to argue that the object hiding in the Real-life setting had more sensory feedback compared to the two VR settings. For example, dropping the token in the box would lead to a sound and the participants had to crouch down to put the hook under the chair. It has been shown that increased sensory feedback can lead to better memory performance (Hoffman et al., 2001) and overall task performance (Pan & Rickard, 2015; Weller & Zachmann, 2012; for a review see Smith, 2019). The increased sensory feedback might have helped with the encoding of the location information in the real-life setting.

The d’ score analysis revealed significant effects of setting. Object recognition performance did not differ between the real-life and HMD-VR settings with the Desktop-VR setting being lower than the other two. While the almost identical performance in the real-life and HMD-VR settings supports the initial prediction that HMD-VR can lead to more life-like memory performance over Desktop-VR, it needs to be looked at with caution. The high p-value (p=.047) and the highly overlapping 95% CIs for the d’ mean values do show that the difference is small and might have happened by chance. The literature with similar VR comparisons is scarce and difficult to use to explain our findings. Studies have shown either no difference (Mania et al., 2003) or slightly better performance in the Desktop VR setting when compared to HMD-VR (Mania & Chalmers, 2001). What adds to the difficulty is that these studies tried to explain their findings purely on the low levels of immersion due to their HMD-VR systems being based on participants sitting and using a computer mouse for controls.

However, an interesting finding can be seen when comparing object recognition results to the WWWW results. When looking at the separate WWWW components data, a similar lack of difference between HMD-VR and Desktop-VR can be seen in the What component. This indicates that both cued (recognition task) and non-cued (WWWW task) recall performance does not differ between the two VR systems. However, the combined WWWW proportions did show higher EM performance in HMD-VR than Desktop-VR indicating that HMD-VR leads to a better binding of WWWW information. A similar observation can be seen when looking at the When/Which proportions obtained in the object recognition task. As it can be seen in Table 6.12, Which and When proportion were higher in HMD-VR setting over Desktop-VR. While not directly, this suggests that information is better binded together in HMD-VR than Desktop-VR. It also shows that the recall or recognition of an object cannot be used as an indicator of EM and only through additional recalled contextual information it is possible to conclude that one has recalled an EM.

When looking at the overall confidence rating differences between the settings there was no difference between HMD-VR and Desktop-VR. Previous studies on confidence ratings and memory recall show mixed results with some showing higher ratings for the Desktop-VR (Mania & Chalmers, 2001) while others are showing no difference (Mania et al., 2003) when comparing the two VR settings. The data shows that participants were more confident with their recognition judgements in the real-life setting and equally confident in both HMD-VR and Desktop-VR. However, taking this data with the results from the d' scores it is visible that the confidence judgements do not follow the same trend. Overall, object recognition confidence ratings does not provide much insight into EM.

An interesting finding can be seen in the When/Which part of the object recognition data. Participants found it difficult to discern the correct setting (Which) associated with the Desktop-VR objects (see Figure 6.7.). This shows that while participants were able to recognise Desktop-VR objects on a similar level as objects from the other settings, the Which information was not well retrieved. However, due to the short period of time between leaning and testing (1h), it is also possible to say that this difference was based on reduced encoding. If this was the case, it would indicate there was something about the Desktop-VR setting that stopped participants from accurately encoding context information. However, an interesting insight can be observed after exploring the participants' provided Which guesses and plotting them against the true settings in which the object actually were presented (see Table 6.14). The data indicate that participants tended to mix Desktop-VR objects with HMD-VR objects. This indicates that while participants knew that the object was virtual (observed in one of the VR settings) they had a difficult time discerning if it was from HMD-VR or Desktop-VR. A possible explanation for this is that the novelty of HMD-VR led to an enhanced encoding of the setting information. The effect of novelty of HMD-VR on memory and general engagement has been shown in the literature (Casu et al., 2015; Kirchhoff et al., 2000; Lee & Wong, 2014). This leads to an interesting question – how would this performance change with the removal of HMD-VR setting, using a between-subjects design or having longer training sessions to familiarise participants with HMD-VR. If removing the effect of novelty from HMD-VR would remove or reduce the difference in Which recall between Desktop-VR and the real-life settings, it would indicate that the difference is mainly due to HMD-VR setting being present. As such, this finding shows the effect of novelty on the encoding of context (which) information.

One of the very initial ideas behind this experiment and the thesis as a whole was that HMD-VR should lead to higher levels of immersion and sense of presence which in turn should result in better memory performance. The results from the IPQ and PQ presence questionnaires do not support this. While the HMD-VR setting was rated higher than the Desktop-VR setting in both of the questionnaires none of the questionnaire scores correlated with the full integrated WWWW recalls. The difference in levels of presence can be associated with the 360 immersion of the HMD-VR setting and the previously mentioned effect of enactment (Schubert, 2002). However, while this might be true, the literature is mixed when exploring differences between the two VR settings. A number of studies have shown no difference in levels of presence (Mania et al., 2003; Mania & Chalmers, 2001) while some are showing higher levels in HMD-VR setting (Buttussi & Chittaro, 2018; Shu et al., 2019). The problem is that there is a lack of research that explores EM using the two VR settings and looks at the levels of presence. In conclusion, the present results suggest that while HMD-VR might lead to higher levels of presence, the level of presence might not be a good indicator for the recall of integrated EMs.

In general, the present experiment provided partial support to the hypothesis that EM performance would be highest in the real-life setting following HMD-VR setting with Desktop-VR having the worst performance. The combined WWWW data showed that EM performance was better in the HMD-VR setting compared to the Desktop-VR setting with the real-life setting having higher performance than both of the VR settings. However, the object recognition data showed real-life performance being better than both of the VR settings with no difference in performance between HMD-VR and Desktop-VR. Additionally, the experiment showed that the higher levels of presence observed in HMD-VR did not correlate with the combined WWWW data, going against the literature regarding positive relationship between presence and EM performance.

Chapter 7 – General discussion

7.1. Summary of aims

The present thesis aimed to explore how HMD-VR might be used to increase the ecological validity of episodic memory (EM) testing. This overarching aim underpinned the two research questions dividing the experiments presented in this thesis into two groups. Firstly, how does sleep-dependant memory consolidation affect event memory (Experiments 1 and 2)? Secondly, how does EM performance in HMD-VR compares to the more traditionally used Desktop-VR (Experiments 3 and 4)? These two major research questions were further intermixed with more specific research questions in each of the individual experiments.

Experiment 1 had two research questions. Firstly, to explore how EM for events might differ to EM for non-events. This question was enclosed in a secondary question of how the sleep dependant memory consolidation and the effect of time of sleep before learning affects event and non-event EM. This was investigated by immediate and delayed (24h) EM tests. These questions were explored using custom made virtual environments (VEs) presented through an HMD-VR system.

Experiment 2 followed with the same research questions as Experiment 1 but extended the EM testing by adding additional sessions after 7 and 30 days. This was done to explore how the sleep dependant memory consolidation and the effect of time of sleep before learning affect EM over a longer period of time. Additionally, one of the secondary aims of the experiment was to explore how different measures of EM relate to each other.

Experiment 3 moved back to the 24h testing to explore the secondary overarching aim of the thesis – how EM performance in HMD-VR differs to Desktop-VR. This research question was investigated by using identical tasks to the ones used in the previous two experiments.

Experiment 4 continued comparing EM performance between HMD-VR and Desktop-VR. The experiment introduced a real-life condition which allowed to better investigate the differences in EM between HMD-VR and Desktop-VR and to answer the question of which VR system better represents real-life EM performance.

While the main aims of the thesis were achieved, the answers to some of the research questions were also clear. The following section (7.2) will provide a brief reminder of the literature that motivated the present thesis. This is followed by a brief summary of the experimental findings (section 7.3). These finding will be discussed in light of related literature. Starting with the use of HMD-VR (section 7.4), the event and non-event memory (section 7.5), memory consolidation (section 7.6) and measuring EM (section 7.7). The chapter will be finished with a conclusion and suggested future direction in similar research (Section 7.8).

7.2. Thesis motivation

EM allows us to receive and store information about events and spatio-temporal relationships between them. In other words, it keeps information regarding *what*, *where* and *when* (Tulving, 1985, 2002). Episodic memory shares features with autobiographical memory, semantic memory, and source memory. In addition to that, the same name is given for two single processes or different names for what might be a single process within the memory field. This mostly refers to the dual-process theory of recognition memory and the links between Remember/Know and Familiarity/Recollection processes (Wixted & Mickes, 2010; Yonelinas, 2002). To make things more difficult, memory tests are sometimes labelled according to the way they assess memory (e.g. free or cued recall; Padilla-Walker & Poole, 2002) or the nature of what is being remembered (e.g. item or source memory; Guo et al., 2006). Due to this, the internal consistency of the EM literature is limited to different tasks being used to assess the same process.

Additionally, research shows that the time interval between encoding and retrieval is crucial due to memory consolidation (Rasch & Born, 2013). However, some experiments tend to test retrieval immediately after encoding (e.g. Plancher et al., 2012) which misses out the memory stabilisation (Dudai, 2012). What complicates things, even more, is the lack of ecological validity in the tasks used to test EM as they tend to evaluate abstract constructs without referencing real-life performance or behaviour (Parsons, 2015; Parsons et al., 2017). One way of achieving this is by using virtual reality. Virtual reality use in EM research has become more prevalent as it provides experiences that are close to daily life while still having high experimental control (Lloyd et al., 2009; Plancher et al., 2010, 2012). By letting participants interact with rich multimodal environments it allows obtaining data that is closer to the real-life compared to pen-and-paper tests, standard computer interfaces or virtual environments presented on computer screens (Mestre & Vercher, 2011).

This brief reminder of the related literature shows how each experiment adds to the understanding of four areas that were explored in this thesis: episodes, memory consolidation, measurements of EM and the use of HMD-VR.

7.3. Summary of findings

Chapter 3 presented the first empirical experiment (Experiment 1), which aimed to explore how EM for event and non-event objects is affected by sleep-dependant memory consolidation. Participants explored VEs containing a number of predesigned event objects in them (e.g. books falling or TV turning on) and performed the free recall, What-When-Where and object recognition tasks. The same memory tasks were performed 24 hours after the initial testing. The experiment also employed the AM/PM testing design to investigate the effect of time of sleep on EM consolidation. As an additional explorative measure, participants wore sleep tracking actigraphy bracelets throughout the 24 hours, which provided sleep data such as time spent in SWS and REM.

As predicted, event objects were better recalled then non-event objects supporting the hypothesis that more life-like events are better recalled than static EMs. This effect was visible in all of the tests (free recall, WWW, and object recognition). When looking at the WWW data, memory for both types of objects did not differ straight after encoding but was higher for the event objects after 24h. This indicated that EM for events might not rely on enhanced encoding but on preferential consolidation. When looking at the effect of time of sleep there were no differences if the period of time between the two testing sessions was filled with wake first and then sleep (AM group) or sleep first and then wake (PM group). This did not support the hypothesis that having sleep closer to the encoding would positively affect the EM performance.

Experiment 2 in Chapter 4 aimed to explore how EM for events and non-events changed over a course of 30 days. This was achieved by repeating the same procedure as in Experiment 1 but adding 7 day and 30 day testing sessions, to elongate the time course of EM retrieval. The experiment also continued to explore the effect of sleep dependant consolidation and the effect of time of sleep (AM/PM) on the EM. Lastly, the experiment explored the validity of different measures of EM.

The experiment replicated the previous findings regarding the event and non-event objects. As previously, the event objects overall were better recalled than the non-event objects. However, after looking at the recall trends, there was a lack of general forgetting between the 7 day and 30 day sessions in all of the tests. One of the explanations for this was the effect of retesting as participants in the 7 day and 30 day sessions were presented with objects that they have been exposed to objects from the previous sessions. Lastly, Experiment 2 explored the relationships between all of the used memory measures. It was found that the combined WWW proportions positively correlated with almost all of the EM measures. This indicates that directly or indirectly all of those measures might relate to EM.

One of the underlying aims of Experiments 1 and 2 and one of the main arguments of the thesis was that HMD-VR use should lead to more life-like memory representations and thus more ecologically valid data as compared to more traditional methods. As such Experiment 3 in Chapter 5 aimed to explore how HMD-VR compared to a more conventional and more widely used Desktop-VR while using identical VEs, tasks and methodology, to determine the validity of HMD-VR.

The general prediction was that EM performance in the free recall, WWW and object recognition tests would be more accurate in the HMD-VR group, compared to the Desktop-VR. However, the memory performance did not differ between the two VR types in almost all of the EM measures. The effects seen in previous experiments such as the more accurate memory for the event than the non-event objects and the lack of differences between the two object types straight after testing but better memory for event objects after 24 hours remained. Overall these results lead to an important conclusion that was counter to the main notion of the thesis, that in terms of EM performance Desktop-VR may be as useful as HMD-VR.

Due to this conclusion, it was important to explore the ecological validity of HMD-VR and Desktop-VR further by comparing EM from the two VR systems to EM obtained in a real-world setting. This was explored by Experiment 4 in Chapter 6. The prediction was that EM performance would be highest in the real-life setting following HMD-VR setting with Desktop-VR having the worst performance. The results only partly supported this prediction. The combined WWWW data showed that EM performance was better in the HMD-VR setting compared to the Desktop-VR setting with the real-life setting having higher performance than both of the VR settings. However, the object recognition data showed real-life performance being better than both of the VR settings with no difference in performance between HMD-VR and Desktop-VR.

7.4. The use of HMD-VR

The main overarching aim of this thesis was to explore how HMD-VR can be used to test EM in an ecologically valid fashion. The following sections discuss the thesis findings in relation to the event and non-event objects, sleep, consolidation and memory measurements. However, each experiment in the present thesis involved participants using HMD-VR to explore virtual environments (VEs) and create EMs in them.

As discussed Experiment 3 shown no differences in memory performance between HMD-VR and Desktop-VR whereas Experiment 4 demonstrated that EM performance in the WWWW task was better in the HMD-VR than in Desktop-VR. The difference in experimental findings was mainly attributed to the change in the task (goal-based instead of free exploration) in Experiment 4 which was employed to offset the passive behaviour observed in Desktop-VR. Indeed, a number of behavioural differences were observed between the two VR settings in both of the experiments. For example, higher levels of inquisitiveness and inspection of objects or more time spent moving around - all associated with HMD-VR. While the results from Experiment 4 might seem to support the main premise of the thesis, caution is needed. For example, the better memory performance was not observed in other measures such as the separate WWWW components or the object recognition scores. Even when looking at the main WWWW scores the difference between the VR settings barely reached the significance level (p=.025). These results, in addition to the findings from Experiment 3, lead to a conclusion that Desktop-VR can be used to explore life-like EM as well as HMD-VR.

However, these results and conclusion lead to an important question – how useful is it to use HMD-VR over Desktop-VR in memory research? As it has been shown in research and in the present thesis, EM can be created in virtual environments presented through Desktop-VR (Sauzéon et al., 2012; Smith, 2019). What is more, as observed in the present thesis (especially in Experiment 3) the trends in memory recall and consolidation are almost identical.

To successfully encode information, attentional resources need to be allocated towards that information (Treisman & Gelade, 1980) (Buckner et al., 2000; Chun & Turk-Browne, 2007; Iidaka et al., 2000). Research has shown that low familiarity with VR system can have an attentional 'burden' increasing the difficulty of the main task which results in a dual-task situation. This can lead to a reduction in memory performance (Makransky et al., 2019; Waller, 2000). It is possible that this affected EM performance in the present experiments and especially led to the lack of differences in Experiments 3 and 4. This demonstrates Desktop-VR advantage, due to it being a well-known system and as such using less attentional resources needed for the task. To overcome this problem studies employing HMD-VR need to include a familiarisation phase at the beginning of the experiments (Camara Lopez et al., 2016). While such phase was used in each of the present experiments it did not last more than 5 minutes and was focused more on the learning of the controls and a general introduction to HMD-VR in a plain VE. It is possible that a longer, more involved familiarisation was needed. Additionally, the training VEs need to be similar quality as the main experimental VEs as participants in every experiment have indicated surprise at the difference in the quality of VEs.

Episodic memories need to have some personal significance for them to be consolidated into the autobiographical memory system (Akhtar et al., 2019; Bauer & Larkina, 2016; Conway, 2001). HMD-VR can provide immersion and real-time interaction with an environment which have been shown to lead to self-experience and body representation. These two qualities have been shown to be central in daily-life experiences (Makowski et al., 2017; Nash et al., 2000) and reinforce EM performance (Bergouignan et al., 2014; Bréchet et al., 2019; Repetto et al., 2016; Tuena et al., 2019). Self-experience and body representation is related to the self-relevance which helps form vivid, life-like memories (Conway, 2005; Marsh & Roediger, 2012) which in turn helps experiences become part of the autobiographical memory system (Cabeza et al., 2004; Svoboda et al., 2006). Self-relevance denotes a feeling of being affected by what is happening in the environment (Schubert et al., 2001) and it has been argued that HMD-VR leads to increased self-relevance than conventional laboratory experiments and even Desktop-VR (Kisker et al., 2019). Overall, the discussed literature shows the clear theoretical advantage of HMD-VR in real-life-like EM research. While the EM data from the present thesis lean more towards the advantage of

HMD-VR use (mainly results from Experiment 4), more research is needed, paying attention to the discussed points such as the behavioural differences and effects of attention. A number of technical insights are also important to discuss if one considers using HMD-VR in EM research.

First and foremost is the crucial need for programming and technological knowledge for the creation of virtual environments. As at the time of writing the present thesis there were no studies that did similar experiments, the only option was to create and code custom environments. However, with the increase in the use of VR in research experimental frameworks (e.g. Brookes et al., 2020) are being created that reduce the need of programming skills. As the aim of the thesis was to explore EM in an ecologically valid fashion while using HMD-VR there was also a need to search for 3D objects that would look life-like. However recently, databases of 3D objects started being published to use in research (Peeters, 2018; Popic et al., 2020; Tromp et al., 2020). These databases would further reduce the time needed for HMD-VR experiment creation. However, it is worth pointing out that in some cases it might not be viable to combine the objects from different databases as their visual quality differs. This relates to the research showing the positive effect on memory from increased levels of visual fidelity and detail (Rauchs et al., 2008; Smith, 2019; Wallet et al., 2011). A related issue is that the participant's ability to inspect objects from various angles can further complicate the creation of virtual environments for HMD-VR. For example, during the piloting of Experiment 1, one of the participants noticed that there was no actual lightbulb in one of the lamps or that the screen of the TV was slightly detached from the TV itself. Such details would be difficult to notice in Desktop-VR. However, it is worth pointing out that this issue shows the life-like exploratory behaviour observed in HMD-VR which links back to the earlier discussed increased chance that the experience will become part of the autobiographical memory system.

Following that, there are some other practical considerations, that are not in favour of HMD-VR. For example, there is a need for an extended space so that participants could explore the virtual environments unhindered by any objects in the real life. After each participant, the headset needs to be cleaned as part of it always touches the participant's skin. During the exploration, the experimenter always needs to be close to the participant to hold the cable connecting the HMD to the computer. This knowledge that the experimenter is always close to the participant can amplify the earlier discussed issue of divided attention and therefore affect EM performance. Lastly, there is an important need to plan the time participants will spend using the headset and how long and often they will have breaks. This is due to the possibility of cybersickness (Smith, 2019; Weech et al., 2019) which again can affect participant's sense of presence and attention.

Research has shown that HMD-VR is useful in many different fields such as treatment of phobias (for a review see Maples-Keller, Bunnell, Kim, & Rothbaum, 2017), Post-traumatic Stress Disorder (Botella et al., 2015; De La Rosa Gómez & López, 2012) and anxiety (Morina et al., 2015; Opriş et al., 2012; Parsons & Rizzo, 2008a). The findings from the present thesis and the discussed current literature show the advantage of HMD-VR use in the EM field. The next section will discuss what insights into EM were provided by the event and non-events object use in Experiments 1, 2 and 3.

7.5. Events and non-events

The main underlying notion behind the present thesis was that to obtain data on EM that is reflective of everyday behaviour it is important to provide stimuli that are reflective of everyday experiences. Experiments 1, 2 and 3 argued that the event objects were a better representation of our everyday episodes than the non-event. The non-event objects were used as an 'alternative' to the usually used stimuli such as pictures of objects or lists of words in EM research. The data from the three experiments have shown that while participants were able to recall and combine the WWW information for both types of objects, event objects (or events) were recalled better than non-events overall. Following the aforementioned notion, regarding the need to use more life-like experiences, these data suggest that the more life-like experiences result in different and more importantly better EM performance than static objects. However, this warrants a deeper examination.

EM in the present thesis was defined using Tulving's definition as a memory store that holds information about events and the spatio-temporal relationships among them (Tulving, 1972, 2002). In all four of the experiments, participants were able to create EMs composed out of the What, When, Where (or in the case of Experiment 4 – What-When-Where-Which) information. While the event objects used in Experiments 1, 2 and 3 represented events as defined in the initial description, participants were also able to recall the WWW information regarding the non-event objects. This shows that participants formed EMs about static objects or in other words – participants recalled seeing those objects in the virtual environments using EM. Such description is almost identical to the ones used in EM research employing pictures of objects If participants were able to create EM for both types of objects why was there such a difference in the recall performance?

A possible explanation for this might be due to the nature of memory traces. In Chapter 1 (section 1.7.1), memory trace theories were identified as the underlying mechanism behind memory consolidation and retrieval (Moscovitch & Nadel, 1998; Nadel & Moscovitch, 1998; Yassa & Reagh, 2013). These theories posit that all episodic information is encoded by the hippocampal neurons which act as an index for the neocortical neurons. The difference between the event and non-event objects or as argued by the thesis, lifelike experiences and static objects, might result from a difference in memory trace connections. This is due to event memory traces having more information which leads to enhanced consolidation and therefore lower forgetting rates. Research has shown that hippocampus plays a role in memory integration, creating links between related memory traces (Horner et al., 2015; Kumaran et al., 2016; Schapiro et al., 2017; Wang, 2019). An ensemble of traces related to an event object will by definition be larger than an ensemble of traces related to a non-event objects due to the additional feature of an event. As such there are more ways to reactivate and in turn retrieve the event object. For example, the memory trace ensemble related to the TV turning itself on, playing static and turning itself off would have memory traces related to the sound of the static, the static image itself and the turning on/off of the TV. When participants were cued as in the WWW or object recognition tasks or self-cued as in the free recall task, they were able to 'initiate' the retrieval from a wider selection of memory traces. Taking the TV example, a participant could have used all of the earlier listed information as a starting point for the episodic recall.

Here it is important to consider if the discussed 'event' is an additional or qualitatively different, feature of encoding. From the previous description of the event memory traces it might seem that the event part of the experience is just an additional piece of information that provides additional ways to access the memory itself. This is not entirely the case. It has been shown that event features depend on one another to and through remembering those features are binded together to create a single coherent event representation (Cooper & Ritchey, 2019; Horner & Burgess, 2013, 2014). Therefore, retrieval of an EM involves coordination a number of areas such as the hippocampus and the prefrontal cortex to reconstruct the particular memory from many different memory representations (Cooper & Ritchey, 2019). As such it is not entirely correct to say that the enhanced event recollection after a period of time, as observed in the present thesis, was based purely on consolidation. It is possible to suggest that events are both encoded and consolidated differently. As mentioned before, to create a coherent representation of an experience all of the needed features need to be binded together. Observing a TV turning itself on and off cannot be reduced to a simple memory trace of an object, an event, its location and when it happened. Such experience involves schemas and previous knowledge which provide interpretation and organisation of ongoing experiences (Dudai, 2012; Moscovitch et al., 2016; Wang & Morris, 2010). To understand the order of experiences, it is necessary to understand those experiences in relation to the previous knowledge of similar situations. For example, the TV turning on, playing static and turning off would require semantic knowledge regarding those experiences. This relates to the mentioned schemas and scripts (Alonso et al., 2020; Bird, 2020; Bransford & Johnson, 1972; Brewer & Treyens, 1981). Therefore, memory for lifelike events is not a simple representation of the incoming information, but a mix of that information and the stored semantic knowledge.

As a result of this additional information, it is possible to argue that the nature of encoding is also different for event as compared to non-event EMs. This is due to the fact that to encode and combine the event information with the current knowledge requires engagement of additional processes and as a result additional brain regions (Grilli et al., 2019; Uncapher et al., 2006). While indirectly, the free recall data from Experiments 1, 2 and 3 support this hypothesis. When participants had to recall their experiences using internal cues, events were better recalled than non-events even immediately after encoding. While it is not possible to rule out that the memory trace size had an effect on these results, the enhanced encoding and binding of the event information provides a better explanation.

A similar hypothesis of enhanced encoding can also be used to explain the difference in perceptual detail recall. In Experiments 1, 2 and 3 participants recalled more perceptual details about the event objects than non-event objects. This is visible in both the free recall task and the Detail part of the WWW task. Research has shown that EM is sensory-perceptual in nature (Conway, 2001). The combination of the higher number of recalled perceptual details and the overall better memory for event objects indicates that the more life-like experiences are accompanied with more information than memories for static objects. This again shows the complexity of EM for life-like events which results in both enhanced encoding and retrieval.

However, the presented hypothesis regarding the enhanced encoding does not explain how the encoding is initiated. To successfully encode multiple features, one needs to allocate attentional resources towards the object or event (Treisman & Gelade, 1980). This allocation of attention and the resulting enhanced configural processing is what leads to the enhanced encoding. Research has shown that attention modulates memory (Buckner et al., 2000; Chun & Turk-Browne, 2007; Iidaka et al., 2000). What is more is that it has been shown that information that shows a 'contrast' to previous knowledge leads to increased attention (Fyhn et al., 2002; Lisman & Grace, 2005; Sousa et al., 2015). Going back to the findings in the present thesis, this can be linked to the events being better recalled than the non-events. The static object stimuli (non-events) did not create a 'contrast' with participants' previous knowledge and as such perhaps did not lead to increased attention. However, event objects did lead to increased attention due to the events happening with them being unexpected and thus requiring more attentional resources.

Taken together, the event and non-event data from the present thesis provided an important insight into EM. The experiments showed that EMs for life-like experiences are both better encoded and retrieved over EMs for static objects, which are traditionally used in EM research. The events may, therefore, either be real, valid EMs, or a particularly poignant memory trace, leading to greater accessibility over time.

7.6. Consolidation

So far the results from the thesis have been discussed in terms of overall encoding and retrieval. As mentioned, one of the aims of the thesis was to explore the effect of memory consolidation on the retrieval of EM over time. This exploration revealed a number of important findings.

One of the more interesting findings was the lack of effect of AM/PM testing or time of sleep on EM performance. As it has been discussed, the time the sleep takes place after learning also affects memory formation which has been shown by studies that have compared the effect of sleep just after learning to sleep at a later time (Benson & Feinberg, 1977; Payne et al., 2008; Scullin, 2014; Talamini et al., 2008). In general, the closer learning is to sleep the better memory retention becomes (Benson & Feinberg, 1977; Payne et al., 2008; Talamini et al., 2008). As such, Experiments 1 and 2 employed an AM/PM testing paradigm (e.g. Aly &

Moscovitch, 2010; Hasher et al., 2002; Scullin, 2014) to explore how this paradigm would affect the consolidation of events and non-events. The data from both experiments indicated that reduced time between encoding and consolidation and therefore reduced time for interference (Dudai, 2012; Ellenbogen et al., 2006; Frankland & Bontempi, 2005) did not improve memory performance. The combined findings added to the growing body of literature showing lack of support for such effect (Sheth et al., 2009; Studte et al., 2015; van der Helm et al., 2011; Wilhelm et al., 2011; also see Cordi & Rasch, 2021 for a review). More importantly, a recent review of studies and meta-analyses (Cordi & Rasch, 2021) regarding the effect of sleep on declarative memory showed that this effect might be a lot smaller and less reliable as suggested by the literature. While data from Experiments 1 and 2 showed a lack of support for the early time of sleep effect, it did provide important insights into EM consolidation.

The difference in EM retrieval straight after encoding and after a period of 24h observed in Experiments 1, 2 and 3 shows the importance of including a period of consolidation in EM research. The consistent finding was the lack of statistical difference between EM for events and non-events straight after the encoding but higher recall performance for the events after 24 hours. Testing EM straight after encoding likely would not have revealed this effect. This is due to the fact, and as seen in the present thesis, that it can be difficult to detect the discussed enhanced encoding straight after experiencing the stimuli.

It is important to discuss this finding in the light of memory trace theories (Moscovitch & Nadel, 1998; Nadel & Moscovitch, 1997, 1998; Sutherland et al., 2019; Yassa & Reagh, 2013). As is has been argued, the additional processes and information related to the events, lead to their enhanced retrieval. However, the results from the present experiments show that the mentioned processes and information are mainly beneficial after consolidation, during which the EMs for events are deemed important for the future and as a result strengthened. Such change in retrieval can be linked to the earlier discussed effects of attention which is then followed by goal processing and memory becoming part of the autobiographical store (Conway, 2001, 2009). The effect of goal-centered processing can be described as a function of human memory to preserve information regarding the progress of personal goals (Conway & Pleydell-Pearce, 2000; also see Sousa et al., 2015). For example, completion of a goal ‘taking a break’ can be a part of a wider goal of ‘writing a chapter’. As such, recalling that one has had a break and therefore fulfilled the sub-goal might prompt that the larger goal of 'writing the chapter' has not been achieved. However, this implies that all EMs are based on specific goal attainment hierarchies. What goal led participants to recall the TV playing static in one of the virtual rooms? The results from the present thesis regarding the differences between the events and non-events and the discussed effect of enhanced encoding suggest that EM consolidation might be rather based on the anticipation of possible future scenarios. This also relates to the Constructive Episodic Simulation Hypothesis (Clayton & Wilkins, 2018; Schacter et al., 2008, 2012; Schacter & Addis, 2007) which explains that the function of EM is to use past experiences and use them to learn and plan for the future. The events were recalled better after consolidation, and throughout a 30 day period as shown in Experiment 4, perhaps as they were deemed 'important' for the future. For example, knowing that if participants were to explore similar virtual environments there might be a chance of similar events happening again. Using this hypothesis it is also possible to explain the lack of differences between the events and non-events straight after encoding (e.g. combined WWW performance). As the virtual environments and everything that happened in them were completely novel, there was not enough time for the salience of the events to affect their EMs

A feature of memories considered to be important for the future is that they are often actively retrieved many times which reflect real-life behaviour. It has been shown that rehearsal of information improves later recall (Baddeley et al., 2019; Karpicke & Roediger, 2008; Roediger & Butler, 2011; Soderstrom et al., 2016). One of the explanations for the forgetting trends observed in Experiment 2 was the effect of retesting. However, in the light of the just discussed notion of EM being based on the anticipation of the future, it can also be argued that findings from Experiment 2 reflect real-world memory rehearsal and therefore life-like behaviour. It is important to bring up a study by Baddeley et al. (2019) discussed in Chapter 4. The study showed that when participants were tested after one day, one week and one month (the Four Doors Test) they showed significantly less or no forgetting at all, compared to when they were tested only immediately and after a month (the Crimes Test). What is important to the present thesis is that forgetting was still observed in the experiment that used more ecologically valid stimuli (Four Doors Test) unlike in the experiment which used semantic knowledge-based stories. This shows that some trends regarding EM consolidation become visible only when using more lifelike stimuli as their complexity better reflects everyday memory consolidation.

Overall, the experiments in the present thesis showed that events tend to be preferentially consolidated over non-events. Taking into consideration the findings discussed in the previous section, the results from the present thesis indicate that EM for events is enhanced at all three stages of memory: encoding, consolidation and retrieval. Continuing with the discussion regarding trends and the effect of consolidation, it is important to discuss the different data that was provided by the various measures of EM used in the present thesis.

7.7. Measuring episodic memory

One of the thesis aims was to explore the different measures of EM and the relationships between them. As it has been discussed in Chapter 1 (section 1.4.7) a common problem in EM testing is the lack of links to the definition and/or the main components of EM, such as the WWW triad or the autonoetic consciousness. What is more, research on the tests themselves shows that not all of them relate to one another (Cheke & Clayton, 2013, 2015).

Table 4.10 illustrated that the combined WWW measure positively correlated with almost all other measures (excluding the Know and Guess judgement proportions). Most importantly, the free recall and object recognition tasks showed the highest positive correlations, which add validity to the decision to choose those three tests as the main measurements of EM. However, it is important to point out the differences in results and trends the tests revealed.

For example, the earlier discussed higher EM accuracy after at least 24h period but not straight after encoding was only observed in the WWW and d' object recognition scores and not in the free recall task. In the free recall task, event objects were always better remembered than the non-event objects event straight after testing. This is interesting as the free recall task is more reflective of an everyday remembering (Morris & Frey, 1997) and the internal cueing during the task should lead to items being remembered more episodically (Tulving, 1972). While one argument is that the free recall measure of the number of recalled objects might reflect more semantic information than episodic, the same was not observed in the separate What component data. In Experiments 1, 2 and 3, the recall of the What component did not differ between the two object types when tested straight after encoding. The only difference between these two measures was that the objects in the free recall task were cued internally while in the What task – externally (participants were given a name of an object). This is an important finding as it shows that measuring EM with just one type of test might not reveal important effect and/or interactions.

It is important to go back to the earlier mentioned (section 7.4) notion that EM is closely related to semantic memory (Alonso et al., 2020; Bird, 2020; Bransford & Johnson, 1972; Brewer & Treyens, 1981). Research has shown that episodic and semantic memory systems are interdependent, share many attributes and are more along a continuum (Craik, 2000, 2020; Greenberg & Verfaellie, 2010; Renoult & Rugg, 2020; Saive et al., 2015; Strikwerda- Brown et al., 2019). This relates to the very early discussion regarding the episodic/semantic memory distinction in Chapter 1 (section 1.1) and that it is not possible to test just one memory system (Jacoby, 1991; McCabe, Roediger, et al., 2011). However, as discussed in Chapter 1, the main distinction between the two memory systems and the main attribute of EM is autonoetic consciousness (Tulving, 2002, 2004; also see Klein, 2013) which is measured by the Remember/Know paradigm (Tulving, 1985; Wais et al., 2008; Wixted, 2009). Remembering refers to the mental recollection of personally experienced events and is associated with EM, whereas knowing refers to the retrieval of decontextualized earlier learnt information and is associated with semantic memory (Tulving, 1972; Wais et al., 2008; Wixted, 2009). Interestingly the insights into EM and the effect events and non-events from the R/K/G data ware mixed. Nevertheless, while Experiment 2 showed that events lead to more Remember judgements, this was not observed in Experiments 1 or 3. One of the reasons for that could have been that in Experiment 2, the R/K/G judgement data was taken for all of the WWW components and not just one overall judgement per object (e.g. Dewhurst et al., 2009; Mickes et al., 2013; Saive et al., 2015). However, this should not have had a detectable effect as if the individual object was recalled using EM (remembered) the accompanying information such as when and where the object was should also be recalled episodically and result in remember judgements. Additionally, it is important to note that there was no Remember-Know shift observed in any of the experiments (Dewhurst et al., 2009; Herbert & Burt, 2003). This was especially predicted in Experiment 2 due to the length of the experiment.

The more interesting insights are revealed when looking at the relationships between the R/K/G judgements and the other measures. For example, the remember judgements yielded the lowest (however still statistically significant) correlation out of any other measures to the combined WWW proportions. This is interesting as it is argued that the remember judgements indicate autonoetic consciousness – an integral part of the episodic recall. Here it is possible to argue that a different measure might be more useful to test autonoetic and therefore episodic recall – detail data.

As discussed by Conway (2001), episodic information is conceived as being largely sensory-perceptual in nature. Indeed, research shows that perceptual richness is one of the main features of EM that contributes to how vividly events are re-experienced (Brewer, 1986, 1995; Conway, 2009; Rubin et al., 2003). What is more, perceptual details are context-specific and therefore are not easily generalised and semanticised (Winocur et al., 2010; Winocur & Moscovitch, 2011; Yonelinas et al., 2019) resulting in what has been called – high-resolution content (Yonelinas, 2013). Due to this, memory for perceptual details can be an indicator of recollection (St-Laurent et al., 2016).

The WWW detail data from the present thesis does lend some support to this hypothesis. For example, in all three experiments, there was a drop in details recalled over the initial 24 hour period showing the loss of perceptual richness of the recalled objects. More interestingly, there was a significant downwards trend in the 30-day period indicating that participants kept losing perceptual information regarding the objects they had experienced. Due to the perceptual detail research discussed earlier (Brewer, 1986, 1995; Conway, 2009; Rubin et al., 2003), the present detail data and its trends can be likened to the R-K shift (Conway et al., 1997; Herbert & Burt, 2001). The loss of perceptual details has been shown to be associated with gist extraction and memory semantisation (Furman et al., 2007, 2012; Sekeres et al., 2016) However, this way of exploring recollection and autonoetic consciousness has an advantage over the traditional R/K/G judgements as the data is easily quantifiable and does not suffer from the lack of objective criteria (Dunn, 2004; Wixted, 2007).

Going back to the discussed effect of perceptual richness, it has been shown that hippocampal engagement correlates with ratings of memory vividness (Gilboa, 2004; Rabin et al., 2010; Sheldon & Levine, 2013; St-Laurent et al., 2016). In Experiment 4, participants had to provide overall vividness ratings regarding their recalled information. A correlation table between the vividness ratings and the R/K/G judgements (see Table 6.8) revealed a positive correlation with remember judgements, no correlation with Know judgements and negative correlation with Guess judgements. If vividness is associated with the perceptual nature of EM and it correlates with remember judgements, which indicate episodic recall of information, it might be used instead of the R/K/G judgements. The use of vividness might overcome R/K/G issues such as its subjectivity and problems with understanding instructions (see Migo et al., 2012 for a review). Overall, the data add further support for the use of perceptual details as a measure of episodic recall.

7.8. Conclusions and future directions

The present thesis aimed to investigate how HMD-VR can be used to test episodic memory in an ecologically valid fashion. The central premise was that HMD-VR should provide experiences closer to real-life than the more conventional Desktop-VR. The results from the present thesis combined with current literature have shown that HMD-VR can be and should be used as a tool for exploring daily-life-like EM.

In future research, it is important to explore how habituation to HMD-VR or lack of it affects episodic memory. As HMD-VR is still a relatively new system and not a lot of people have experienced it, it is important to explore if and how memory performance is affected by extended training sessions in HMD-VR. In the present thesis, participants spent no more than five minutes in the training rooms which might not have been enough to offset the novelty of the HMD-VR. When participants are equally comfortable using both HMD-VR and Desktop-VR one should explore the differences between different object types or the effects of memory consolidation. Only then it would be possible to explore the true differences in encoding, consolidation and retrieval of episodic memory between the two VR systems.

One of the secondary aims of the thesis was to explore the differences in memory retrieval for events and non-events with the non-events representing static stimuli which are often used in episodic memory research. The thesis has shown that events were better retrieved then non-events again indicating the issue of lack of ecological validity in memory research and the need to use more life-like stimuli. Future research should continue focusing on the use of life-like stimuli and move towards the recreation of everyday activities. Combining events and goal-based tasks (both used in the present thesis) with HMD-VR would allow exploring episodic memory while maintaining close to life-like behaviour which the literature lacks.

Lastly, the majority of the thesis was interested in the sleep dependant memory consolidation and the effect of time of sleep before learning affects event and non-event episodic memory. The results showed that the increased retrieval of episodic memory for events is mainly based on enhanced consolidation as the retrieval performance showed no difference compared to non-events when tested immediately after encoding. However, after discussing the results in relation to memory trace and consolidation theories a conclusion was reached that both encoding and consolidation of episodic memory is enhanced for events. The findings show the importance of sleep dependant memory consolidation and its use in episodic memory research. As such, future research should include more than one testing session with a period of sleep between them. This way a fuller picture of episodic memory performance would be available. Combining this with the life-like experiences and HMD-VR would lead to episodic memory performance that truly represent the real-life experiences.

References

Addis, D. R., Moscovitch, M., & McAndrews, M. P. (2007). Consequences of hippocampal damage across the autobiographical memory network in left temporal lobe epilepsy. *Brain*, *130*(9), 2327–2342.

Addis, D. R., & Schacter, D. L. (2008). Constructive episodic simulation: Temporal distance and detail of past and future events modulate hippocampal engagement. *Hippocampus*.

Aggleton, J. P., & Brown, M. W. (2006). Interleaving brain systems for episodic and recognition memory. *Trends in Cognitive Sciences*, *10*(10), 455–463.

Agranoff, B. W., Davis, R. E., & Brink, J. J. (1966). Chemical studies on memory fixation in goldfish. *Brain Research*.

Akhtar, S., Justice, L. V., Morrison, C. M., Conway, M. A., & Howe, M. L. (2019). What Are Autobiographical Memories? A Reply to Bauer, Baker-Ward, Krøjgaard, Peterson, and Wang (2019). *Psychological Science*, 1–3.

Allen, T. A., & Fortin, N. J. (2013). The evolution of episodic memory. *Proceedings of the National Academy of Sciences*, *110*(Supplement 2), 10379–10386.

Alonso, A., van der Meij, J., Tse, D., & Genzel, L. (2020). Naïve to expert: Considering the role of previous knowledge in memory. *Brain and Neuroscience Advances*, *4*.

Aly, M., & Moscovitch, M. (2010). The effects of sleep on episodic memory in older and younger adults. *Memory*, *18*(3), 327–334.

Amin, A., Gromal, D., Tong, X., Shaw, C., Gromala, D., Tong, X., Shaw, C., Gromal, D., Tong, X., & Shaw, C. (2016). Immersion in cardboard VR compared to a traditional head-mounted display. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *9740*, 269–276.

Arthur, E. J., Hancock, P. A., & Chrysler, S. T. (1997). The perception of spatial layout in real and virtual worlds. *Ergonomics*, *40*(1), 69–77.

Ashcraft, M. H. (1994). *Human memory and cognition*. Harper Collins.

Ashton, J. E., Cairney, S. A., & Gaskell, M. G. (2018). No effect of targeted memory reactivation during slow-wave sleep on emotional recognition memory. *Journal of Sleep Research*, *27*(1), 129–137.

Atienza, M., & Cantero, J. L. (2008). Modulatory effects of emotion and sleep on recollection and familiarity. *Journal of Sleep Research*, *17*(3), 285–294.

Axmacher, N., Draguhn, A., Elger, C. E., & Fell, J. (2009). Memory processes during sleep: Beyond the standard consolidation theory. *Cellular and Molecular Life Sciences*, *66*(14), 2285–2297.

Babb, S. J., & Crystal, J. D. (2005). Discrimination of what, when, and where: Implications for episodic-like memory in rats. *Learning and Motivation*, *36*(2), 177–189.

Backhaus, J., & Junghanns, K. (2006). Daytime naps improve procedural motor memory. *Sleep Medicine*, *7*(6), 508–512.

Bäckman, L., Small, B. J., & Fratiglioni, L. (2001). Stability of the preclinical episodic memory deficit in Alzheimer's disease. *Brain*, *124*(1), 96–102.

Baddeley, A. D., Atkinson, A., Kemp, S., & Allen, R. J. (2019). The problem of detecting long-term forgetting: Evidence from the Crimes Test and the Four Doors Test. *Cortex*, *110*, 69–79.

Baddeley, A. D., Hatter, J. E., Scott, D., & Snashall, A. (1970). Memory and Time of Day. *Quarterly Journal of Experimental Psychology*, *22*(4), 605–609.

Bahrick, H. P. (1984). Semantic memory content in permastore: Fifty years of memory for Spanish learned in school. *Journal of Experimental Psychology: General*, *113*(1), 1–29.

Barba, G. D. (1993). Confabulation: Knowledge and recollective experience. *Cognitive Neuropsychology*, *10*(1), 1–20.

Barba, G. D., Mantovan, M. C., Ferruzza, E., & Denes, G. (1997). Remembering and Knowing the Past: A Case Study of Isolated Retrograde Amnesia. *Cortex*, *33*(1), 143–154.

Barclay, C. R. (1986). Schematization of autobiographical memory. In *Autobiographical Memory* (pp. 82–99). Cambridge University Press.

Barclay, C. R. (1988). Truth and accuracy in autobiographical memory. In *Practical aspects of memory: Current research and issues, Vol. 1: Memory in everyday life*.

Barham, M. P., Enticott, P. G., Conduit, R., & Lum, J. A. G. (2016). Transcranial electrical stimulation during sleep enhances declarative (but not procedural) memory consolidation: Evidence from a meta-analysis. *Neuroscience & Biobehavioral Reviews*, *63*, 65–77.

Barr, W. B., Goldberg, E., Wasserstein, J., & Novelly, R. A. (1990). Retrograde amnesia following unilateral temporal lobectomy. *Neuropsychologia*, *28*(3), 243–255.

Barrett, T. R., & Ekstrand, B. R. (1972). Effect of sleep on memory: III. Controlling for time-of-day effects. *Journal of Experimental Psychology*, *96*(2), 321–327.

Barry, D. N., & Maguire, E. A. (2019). Remote Memory and the Hippocampus: A Constructive Critique. *Trends in Cognitive Sciences*, *23*(2), 128–142.

Bartsch, T., Döhring, J., Rohr, A., Jansen, O., & Deuschl, G. (2011). CA1 neurons in the human hippocampus are critical for autobiographical memory, mental time travel, and autonoetic consciousness. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(42), 17562–17567.

Bauer, P. J., & Larkina, M. (2016). Predicting remembering and forgetting of autobiographical memories in children and adults: a 4-year prospective study. *Memory*, *24*(10), 1345–1368.

Ben-Yakov, A., & Dudai, Y. (2011). Constructing Realistic Engrams: Poststimulus Activity of Hippocampus and Dorsal Striatum Predicts Subsequent Episodic Memory. *Journal of Neuroscience*, *31*(24), 9032–9042.

Benedict, R. H. B., & Zgaljardic, D. J. (1998). Practice Effects During Repeated Administrations of Memory Tests With and Without Alternate Forms. *Journal of Clinical and Experimental Neuropsychology*, *20*(3), 339–352.

Benjamin, S. (2007). The Human Frontal Lobes: Functions and Disorders, 2nd ed. *The Journal of Clinical Psychiatry*, *68*(11), 1817.

Benson, K., & Feinberg, I. (1977). The beneficaial effect of sleep in an extended Jenkins and Dallenbach paradigm. *Psychophysiology*, *14*(4), 375–384.

Bergouignan, L., Nyberg, L., & Ehrsson, H. H. (2014). Out-of-body-induced hippocampal amnesia. *Proceedings of the National Academy of Sciences*, *111*(12), 4421–4426.

Bird, C. M. (2020). How do we remember events? *Current Opinion in Behavioral Sciences*, *32*, 120–125.

Bird, C. M., & Burgess, N. (2008). The hippocampus and memory: insights from spatial processing. *Nature Reviews Neuroscience*, *9*(3), 182–194.

Bird, L. R., Roberts, W. A., Abroms, B. D., Kit, K. A., & Crupi, C. (2003). Spatial memory for food hidden by rats (Rattus norvegicus) on the radial maze: Studies of memory for where, what, and when. *Journal of Comparative Psychology*, *117*(2), 176–187.

Blessed, G., Tomlinson, B. E., & Roth, M. (1968). The association between quantitative measures of dementia and of senile change in the cerebral grey matter of elderly subjects. *The British Journal of Psychiatry : The Journal of Mental Science*.

Born, J., Rasch, B., & Gais, S. (2006). Sleep to remember. *The Neuroscientist : A Review Journal Bringing Neurobiology, Neurology and Psychiatry*, *12*(5), 410–424.

Born, J., & Wilhelm, I. (2012). System consolidation of memory during sleep. *Psychological Research*, *76*(2), 192–203.

Borson, S., Scanlan, J., Brush, M., Vitaliano, P., & Dokmak, A. (2000). The mini-cog: A cognitive "vital signs" measure for dementia screening in multi-lingual elderly. *International Journal of Geriatric Psychiatry*.

Botella, C., Serrano, B., Baños, R., & García-Palacios, A. (2015). Virtual reality exposure-based therapy for the treatment of post-traumatic stress disorder: a review of its efficacy, the adequacy of the treatment protocol, and its acceptability. *Neuropsychiatric Disease and Treatment*, *11*, 2533.

Bowman, D. A., Sowndararajan, A., Ragan, E. D., & Kopper, R. (2009). Higher Levels of Immersion Improve Procedure Memorization Performance. *Proceedings of the 15th Joint Virtual Reality Eurographics Conference on Virtual Environments*.

Boyd, C. (1997). Does immersion make a virtual environment more usable? *CHI '97 Extended Abstracts on Human Factors in Computing Systems Looking to the Future - CHI '97*, 325.

Brainerd, C. J., & Reyna, V. F. (1993). Memory independence and memory interference in cognitive development. *Psychological Review*, *100*(1), 42–67.

Bransford, J. D., & Johnson, M. K. (1972). Contextual prerequisites for understanding: Some investigations of comprehension and recall. *Journal of Verbal Learning and Verbal Behavior*, *11*(6), 717–726.

Bréchet, L., Mange, R., Herbelin, B., Theillaud, Q., Gauthier, B., Serino, A., & Blanke, O. (2019). First-person view of one's body in immersive virtual reality: Influence on episodic memory. *PLOS ONE*, *14*(3), e0197763.

Brewer, W. F. (1986). What is autobiographical memory? In D. C. Rubin (Ed.), *Autobiographical Memory* (pp. 25–49). Cambridge University Press.

Brewer, W. F. (1995). What is recollective memory? In D. C. Rubin (Ed.), *Remembering our past: Studies in autobiographical memory* (pp. 19–66). Cambridge University Press.

Brewer, W. F., & Treyens, J. C. (1981). Role of schemata in memory for places. *Cognitive Psychology*, *13*(2), 207–230.

Brookes, J., Warburton, M., Alghadier, M., Mon-Williams, M., & Mushtaq, F. (2020). Studying human behavior with virtual reality: The Unity Experiment Framework. *Behavior Research Methods*, *52*(2), 455–463.

Brooks, B. M., & Rose, F. D. (2003). The use of virtual reality in memory rehabilitation: Current findings and future directions. *NeuroRehabilitation*.

Buckner, R. L., & Carroll, D. C. (2007). Self-projection and the brain. *Trends in Cognitive Sciences*, *11*(2), 49–57.

Buckner, R. L., Kelley, W. M., & Petersen, S. E. (1999). Frontal cortex contributes to human memory formation. In *Nature Neuroscience*.

Buckner, R. L., Logan, J., Donaldson, D. I., & Wheeler, M. E. (2000). Cognitive neuroscience of episodic memory encoding. *Acta Psychologica*, *105*(2–3), 127–139.

Burgess, N., Maguire, E. A., Spiers, H. J., & O'Keefe, J. (2001). A Temporoparietal and Prefrontal Network for Retrieving the Spatial Context of Lifelike Events. *NeuroImage*, *14*(2), 439–453.

Burgess, P. W., Alderman, N., Forbes, C., Costello, A., M-A. Coates, L., Dawson, D. R., Anderson, N. D., Gilbert, S. J., Dumontheil, I., & Channon, S. (2006). The case for the development and use of "ecologically valid" measures of executive function in experimental and clinical neuropsychology. *Journal of the International Neuropsychological Society*, *12*(2), 194–209.

Burke, D. M., MacKay, D. G., Worthley, J. S., & Wade, E. (1991). On the tip of the tongue:

What causes word finding failures in young and older adults? *Journal of Memory and Language*, *30*(5), 542–579.

Busey, T. A., Tunnicliff, J., Loftus, G. R., & Loftus, E. F. (2000). Accounts of the confidence-accuracy relation in recognition memory. *Psychonomic Bulletin & Review*, *7*(1), 26–48.

Buttussi, F., & Chittaro, L. (2018). Effects of Different Types of Virtual Reality Display on Presence and Learning in a Safety Training Scenario. *IEEE Transactions on Visualization and Computer Graphics*, *24*(2), 1063–1076.

Bywaters, M., Andrade, J., & Turpin, G. (2004). Intrusive and non‐ intrusive memories in a non‐ clinical sample: The effects of mood and affect on imagery vividness. *Memory*, *12*(4), 467–478.

Cabeza, R., Prince, S. E., Daselaar, S. M., Greenberg, D. L., Budde, M., Dolcos, F., LaBar, K. S., & Rubin, D. C. (2004). Brain Activity during Episodic Retrieval of Autobiographical and Laboratory Events: An fMRI Study using a Novel Photo Paradigm. *Journal of Cognitive Neuroscience*, *16*(9), 1583–1594.

Cabeza, R., Stanley, M. L., & Moscovitch, M. (2018). Process-Specific Alliances (PSAs) in Cognitive Neuroscience. *Trends in Cognitive Sciences*, *22*(11), 996–1010.

Cairney, S. A., Durrant, S. J., Hulleman, J., & Lewis, P. A. (2014). Targeted Memory Reactivation During Slow Wave Sleep Facilitates Emotional Memory Consolidation. *Sleep*, *37*(4), 701–707.

Cairney, S. A., Durrant, S. J., Musgrove, H., & Lewis, P. A. (2011). Sleep and environmental context: interactive effects for memory. *Experimental Brain Research*, *214*(1), 83–92.

Cairney, S. A., Durrant, S. J., Power, R., & Lewis, P. A. (2015). Complementary Roles of Slow-Wave Sleep and Rapid Eye Movement Sleep in Emotional Memory Consolidation. *Cerebral Cortex*, *25*(6), 1565–1575.

Camara Lopez, M., Deliens, G., & Cleeremans, A. (2016). Ecological assessment of divided attention: What about the current tools and the relevancy of virtual reality. *Revue Neurologique*, *172*(4–5), 270–280.

Carcaillon, L., Amieva, H., Auriacombe, S., Helmer, C., & Dartigues, J. F. (2009). A Subtest of the MMSE as a valid test of episodic memory? comparison with the free and cued reminding test. *Dementia and Geriatric Cognitive Disorders*, *27*(5), 429–438.

Cárdenas-Delgado, S., Méndez-López, M., Juan, M.-C., Pérez-Hernández, E., Lluch, J., & Vivó, R. (2017). Using a Virtual Maze Task to Assess Spatial Short-term Memory in Adults. *Proceedings of the 12th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications*, *Visigrapp*, 46–57.

Carr, M. F., Jadhav, S. P., & Frank, L. M. (2011). Hippocampal replay in the awake state: a potential substrate for memory consolidation and retrieval. *Nature Neuroscience*, *14*(2), 147–153.

Casey, S. J., Solomons, L. C., Steier, J., Kabra, N., Burnside, A., Pengo, M. F., Moxham, J., Goldstein, L. H., & Kopelman, M. D. (2016). Slow wave and REM sleep deprivation effects on explicit and implicit memory during sleep. *Neuropsychology, 30*(8), 931–945.

Cassel, J.-C., Cassel, D., & Manning, L. (2012). From Augustine of Hippo's Memory Systems to Our Modern Taxonomy in Cognitive Psychology and Neuroscience of Memory: A 16-Century Nap of Intuition before Light of Evidence. *Behavioral Sciences*, *3*(1), 21–41.

Casu, A., Spano, L. D., Sorrentino, F., & Scateni, R. (2015). RiFtArt: Bringing masterpieces in the classroom through immersive virtual reality. *Italian Chapter Conference 2015 - Smart Tools and Apps in Computer Graphics, STAG 2015*.

Cellini, N., Torre, J., Stegagno, L., & Sarlo, M. (2016). Sleep before and after learning promotes the consolidation of both neutral and emotional information regardless of REM presence. *Neurobiology of Learning and Memory*.

Cermak, L. S. (1972). *Human memory: Research and theory*. Ronald Press.

Cermak, L. S., & O'Connor, M. (1983). The anterograde and retrograde retrieval ability of a patient with amnesia due to encephalitis. *Neuropsychologia*, *21*(3), 213–234.

Chadwick, M. J., Hassabis, D., Weiskopf, N., & Maguire, E. A. (2010). Decoding Individual Episodic Memory Traces in the Human Hippocampus. *Current Biology*.

Chaytor, N., & Schmitter-Edgecombe, M. (2003). The Ecological Validity of Neuropsychological Tests: A Review of the Literature on Everyday Cognitive Skills. *Neuropsychology Review*, *13*(4), 181–197.

Cheke, L. G. (2016). What–where–when memory and encoding strategies in healthy aging. *Learning & Memory*, *23*(3), 121–126.

Cheke, L. G., & Clayton, N. S. (2013). Do different tests of episodic memory produce consistent results in human adults? *Learning & Memory*, *20*(9), 491–498.

Cheke, L. G., & Clayton, N. S. (2015). The six blind men and the elephant: Are episodic memory tasks tests of different things or different tests of the same thing? *Journal of Experimental Child Psychology*, *137*, 164–171.

Chen, H.-Y. Y., Gilmore, A. W., Nelson, S. M., & McDermott, K. B. (2017). Are There Multiple Kinds of Episodic Memory? An fMRI Investigation Comparing Autobiographical and Recognition Memory Tasks. *The Journal of Neuroscience*, *37*(10), 2764–2775.

Chrastil, E. R., & Warren, W. H. (2013). Active and passive spatial learning in human navigation: Acquisition of survey knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*(5), 1520–1537.

Chun, M. M., & Turk-Browne, N. B. (2007). Interactions between attention and memory. *Current Opinion in Neurobiology*, *17*(2), 177–184.

Civan, M. M., Peterson-Yantorno, K., Coca-Prados, M., & Yantorno, R. E. (1992). Regulatory volume decrease by cultured non-pigmented ciliary epithelial cells. *Experimental Eye Research*, *54*(2), 181–191.

Clayton, N. S., Bussey, T. J., & Dickinson, A. (2003). Can animals recall the past and plan for the future? *Nature Reviews Neuroscience*, *4*(8), 685–691.

Clayton, N. S., & Dickinson, A. (1998). Episodic-like memory during cache recovery by scrub jays. *Nature*.

Clayton, N. S., & Dickinson, A. (1999). Memory for the content of caches by scrub jays (Aphelocoma coerulescens). *Journal of Experimental Psychology: Animal Behavior Processes*.

Clayton, N. S., Griffiths, D. P., Emery, N. J., & Dickinson, A. (2001). Elements of episodic–like memory in animals. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *356*(1413), 1483–1491.

Clayton, N. S., & Wilkins, C. (2018). Seven myths of memory. *Behavioural Processes*, *152*, 3–9.

Clayton, N. S., Yu, K. S., & Dickinson, A. (2001). Scrub jays (Aphelocoma coerulescens) form integrated memories of the multiple features of caching episodes. *Journal of Experimental Psychology: Animal Behavior Processes*.

Clemens, Z., Fabó, D., & Halász, P. (2005). Overnight verbal memory retention correlates with the number of sleep spindles. *Neuroscience*, *132*(2), 529–535.

Conway, M. (2003). Commentary Cognitive-affective mechanisms and processes in autobiographical memory. *Memory*, *11*(2), 217–224.

Conway, M. A. (2001). Sensory–perceptual episodic memory and its context: autobiographical memory. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *356*(1413), 1375–1384.

Conway, M. A. (2005). Memory and the self. *Journal of Memory and Language*, *53*(4), 594–628.

Conway, M. A. (2009). Episodic memories. *Neuropsychologia*, *47*(11), 2305–2313.

Conway, M. A., Gardiner, J. M., Perfect, T. J., Anderson, S. J., & Cohen, G. M. (1997). Changes in Memory Awareness During Learning: The Acquisition of Knowledge by Psychology Undergraduates. *Journal of Experimental Psychology: General Copyright*, *126*(4), 393–413.

Conway, M. A., & Pleydell-Pearce, C. W. (2000). The construction of autobiographical memories in the self-memory system. *Psychological Review*, *107*(2), 261–288.

Cook, J. D., Eftekari, S. C., Dallmann, E., Sippy, M., & Plante, D. T. (2019). Ability of the Fitbit Alta HR to quantify and classify sleep in patients with suspected central disorders of hypersomnolence: A comparison against polysomnography. *Journal of Sleep Research*, *28*(4), e12789.

Cook, J. D., Prairie, M. L., & Plante, D. T. (2018). Ability of the Multisensory Jawbone UP3 to Quantify and Classify Sleep in Patients With Suspected Central Disorders of Hypersomnolence: A Comparison Against Polysomnography and Actigraphy. *Journal of Clinical Sleep Medicine*, *14*(05), 841–848.

Coolican, H. (1992). *Reseach Methods and Statistics in Psychology*. Hodder.

Cooper, R. A., & Ritchey, M. (2019). Cortico-hippocampal network connections support the multidimensional quality of episodic memory. *ELife*, *8*.

Cordi, M. J., & Rasch, B. (2021). How robust are sleep-mediated memory benefits? *Current Opinion in Neurobiology*, *67*, 1–7.

Corriveau-Lecavalier, N., Ouellet, É., Boller, B., Belleville, S., Corriveau Lecavalier, N., Ouellet, É., Boller, B., & Belleville, S. (2018). Use of immersive virtual reality to assess episodic memory: A validation study in older adults. *Neuropsychological Rehabilitation*, *30*(3), 462–480.

Craik, F. I. M. (2000). Age-related changes in human. In D. C. Park & N. Schwarz (Eds.), *Cognitive aging: A primer* (pp. 75–92). Psychology Press Ltd.

Craik, F. I. M. (2020). Remembering: An Activity of Mind and Brain. *Annual Review of Psychology*, *71*(1), 1–24.

Crovitz, H. F., & Schiffman, H. (1974). Frequency of episodic memories as a function of their age. *Bulletin of the Psychonomic Society*, *4*(5), 517–518.

Dalla Barba, G., & La Corte, V. (2013). The hippocampus, a time machine that makes errors. *Trends in Cognitive Sciences*, *17*(3), 102–104.

Danckert, J., Ferber, S., Pun, C., Broderick, C., Striemer, C., Rock, S., & Stewart, D. (2007). Neglected time: Impaired temporal perception of multisecond intervals in unilateral neglect. *Journal of Cognitive Neuroscience*.

Darken, R. P., Bernatovich, D., Lawson, J. P., & Peterson, B. (1999). Quantitative measures of presence in virtual environments: The roles of attention and spatial comprehension. *Cyberpsychology and Behavior*.

Daurat, A., Terrier, P., Foret, J., & Tiberge, M. (2007). Slow wave sleep and recollection in recognition memory. *Consciousness and Cognition*, *16*(2), 445–455.

Davachi, L. (2006). Item, context and relational episodic encoding in humans. In *Current Opinion in Neurobiology*.

Davachi, L., & DuBrow, S. (2015). How the hippocampus preserves order: the role of prediction and context. *Trends in Cognitive Sciences*, *19*(2), 92–99.

Davachi, L., Mitchell, J. P., & Wagner, A. D. (2003). Multiple routes to memory: Distinct medial temporal lobe processes build item and source memories. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(4), 2157–2162.

Davison, S. M. C., Deeprose, C., & Terbeck, S. (2018). A comparison of immersive virtual reality with traditional neuropsychological measures in the assessment of executive functions. *Acta Neuropsychiatrica*.

De La Rosa Gómez, A., & López, G. C. (2012). Posttraumatic stress disorder: Efficacy of a treatment program using virtual reality for victims of criminal violence in Mexican population. *Anuario de Psicologia*, *42*(3), 377–391.

De Zambotti, M., Claudatos, S., Inkelis, S., Colrain, I. M., & Baker, F. C. (2015). Evaluation of a consumer fitness-tracking device to assess sleep in adults. *Chronobiology International*, *32*(7), 1024–1028.

Dehn, L. B., Kater, L., Piefke, M., Botsch, M., Driessen, M., & Beblo, T. (2018). Training in a comprehensive everyday-like virtual reality environment compared to computerized cognitive training for patients with depression. *Computers in Human Behavior*, *79*, 40–52.

Delis, D., Kramer, J., Kaplan, E., & Ober, B. (2000). California Verbal Learning Test. In *TX: Psychological Corporation*. Psychological Corporation.

Deuker, L., Olligs, J., Fell, J., Kranz, T. A., Mormann, F., Montag, C., Reuter, M., Elger, C. E., & Axmacher, N. (2013). Memory Consolidation by Replay of Stimulus-Specific Neural Activity. *Journal of Neuroscience*, *33*(49), 19373–19383.

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.

Dewhurst, S. A., Conway, M. A., & Brandt, K. R. (2009). Tracking the R-to-K shift: Changes in memory awareness across repeated tests. *Applied Cognitive Psychology*, *23*(6), 849–858.

Dewhurst, S. A., Holmes, S. J., Brandt, K. R., & Dean, G. M. (2006). Measuring the speed of the conscious components of recognition memory: Remembering is faster than knowing. *Consciousness and Cognition*, *15*(1), 147–162.

Diana, R. A., Yonelinas, A. P., & Ranganath, C. (2008). The effects of unitization on familiarity-based source memory: Testing a behavioral prediction derived from neuroimaging data. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(4), 730–740.

Diba, K., & Buzsáki, G. (2007). Forward and reverse hippocampal place-cell sequences during ripples. *Nature Neuroscience*, *10*(10), 1241–1242.

Diekelmann, S., & Born, J. (2010). The memory function of sleep. *Nat Rev Neurosci*, *11*(2), 114–126.

Diekelmann, S., Wilhelm, I., & Born, J. (2009). The whats and whens of sleep-dependent memory consolidation. *Sleep Medicine Reviews*, *13*(5), 309–321.

Dobbins, I. G., Foley, H., Schacter, D. L., & Wagner, A. D. (2002). Executive Control during Episodic Retrieval. *Neuron*, *35*(5), 989–996.

Dolcos, F., Labar, K. S., & Cabeza, R. (2005). Remembering one year later: Role of the amygdala and the medial temporal lobe memory system in retrieving emotional memories. *Proceedings of the National Academy of Sciences of the United States of America*.

Donaldson, W. (1996). The role of decision processes in remembering and knowing. *Memory & Cognition*, *24*(4), 523–533.

Donix, M., Brons, C., Jurjanz, L., Poettrich, K., Winiecki, P., & Holthoff, V. A. (2010). Overgenerality of autobiographical memory in people with amnestic mild cognitive impairment and early Alzheimer's disease. *Archives of Clinical Neuropsychology*.

Dougal, S., & Rotello, C. M. (2007). "Remembering" emotional words is based on response bias, not recollection. *Psychonomic Bulletin and Review*.

Dragoi, G., & Tonegawa, S. (2011). Preplay of future place cell sequences by hippocampal cellular assemblies. *Nature*, *469*(7330), 397–401.

Drosopoulos, S. (2005). Sleep enhances explicit recollection in recognition memory. *Learning & Memory*, *12*(1), 44–51.

Drosopoulos, S., Windau, E., Wagner, U., & Born, J. (2007). Sleep Enforces the Temporal Order in Memory. *PLoS ONE*, *2*(4), e376.

Drummey, A. B., & Newcombe, N. S. (2002). Developmental changes in source memory. *Developmental Science*.

Dudai, Y. (2004). The Neurobiology of Consolidations, Or, How Stable is the Engram? *Annual Review of Psychology*, *55*(1), 51–86.

Dudai, Y. (2012). The Restless Engram: Consolidations Never End. *Annual Review of Neuroscience*, *35*(1), 227–247.

Dudai, Y., Karni, A., & Born, J. (2015). The Consolidation and Transformation of Memory. *Neuron*, *88*(1), 20–32.

Dunn, J. C. (2004). Remember-Know: A Matter of Confidence. *Psychological Review*, *111*(2), 524–542.

Dunn, J. C. (2008). The dimensionality of the remember-know task: A state-trace analysis. *Psychological Review*, *115*(2), 426–446.

Durrant, S. J., Cairney, S. A., & Lewis, P. A. (2013). Overnight Consolidation Aids the Transfer of Statistical Knowledge from the Medial Temporal Lobe to the Striatum. *Cerebral Cortex*, *23*(10), 2467–2478.

Durrant, S. J., Taylor, C., Cairney, S. A., & Lewis, P. A. (2011). Sleep-dependent consolidation of statistical learning. *Neuropsychologia*.

Easton, A., & Eacott, M. J. (2008). A new working definition of episodic memory: replacing “when” with “which.” In *Handbook of Behavioral Neuroscience* (Vol. 18, Issue 08, pp. 185–196).

Eichenbaum, H. (2004). Hippocampus. *Neuron*, *44*(1), 109–120.

Eichenbaum, H. (2014). Time cells in the hippocampus: a new dimension for mapping memories. *Nature Reviews Neuroscience*, *15*(11), 732–744.

Eichenbaum, H., & Cohen, N. J. (2004). From Conditioning to Conscious Recollection. In *From Conditioning to Conscious Recollection: Memory systems of the brain*. Oxford University Press.

Eichenbaum, H., Yonelinas, A. P., & Ranganath, C. (2007). The Medial Temporal Lobe and Recognition Memory. *Annual Review of Neuroscience*, *30*(1), 123–152.

Ellenbogen, J. M., Payne, J. D., & Stickgold, R. (2006). The role of sleep in declarative memory consolidation: passive, permissive, active or none? *Current Opinion in Neurobiology*, *16*(6), 716–722.

Eysenck, M. W. (1990). *The Blackwell Dictionary of Cognitive Psychology*. Blackwell.

Eysenck, M. W., & Keane, M. T. (2000). *Cognitive Psychology: A student's handbook* (4th ed.). Psychology Press.

Fabes, R. A., Martin, C. L., Hanish, L. D., & Updegraff, K. A. (2000). Criteria for evaluating the significance of developmental research in the twenty-first century: Force and counterforce. *Child Development*.

Fernández, G., & Morris, R. G. M. (2018). Memory, Novelty and Prior Knowledge. *Trends in Neurosciences*, *41*(10), 654–659.

Fillenbaum, G. G., van Belle, G., Morris, J. C., Mohs, R. C., Mirra, S. S., Davis, P. C., Tariot, P. N., Silverman, J. M., Clark, C. M., Welsh-Bohmer, K. A., & Heyman, A. (2008). Consortium to Establish a Registry for Alzheimer's Disease (CERAD): The first twenty years. *Alzheimer's and Dementia*, *4*(2), 96–109.

Flannery, K. A., & Walles, R. (2003). How Does Schema Theory Apply to Real versus Virtual Memories? *CyberPsychology & Behavior*, *6*(2), 151–159.

Folkard, S., & Monk, T. H. (1980). Circadian rhythms in human memory. *British Journal of Psychology*, *71*(2), 295–307.

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, *12*(3), 189–198.

Foster, D. J., & Wilson, M. A. (2006). Reverse replay of behavioural sequences in hippocampal place cells during the awake state. *Nature*, *440*(7084), 680–683.

Frances Wang, R., & Simons, D. J. (1999). Active and passive scene recognition across views. *Cognition*, *70*(2), 191–210.

Frances Wang, R., & Spelke, E. S. (2002). Human spatial representation: insights from animals. *Trends in Cognitive Sciences*, *6*(9), 376–382.

Frankland, P. W., & Bontempi, B. (2005). The organization of recent and remote memories. *Nature Reviews Neuroscience*, *6*(2), 119–130.

Franzen, M. D., & Wilhelm, K. L. (1996). Conceptual foundations of ecological validity in neuropsychological assessment. In *Ecological validity of neuropsychological testing.*

Friedman, W. J. (1993). Memory for the time of past events. *Psychological Bulletin*, *113*(1), 44–66.

Fuchs, P., Moreau, G., & Guitton, P. (2011). *Virtual reality: Concepts and technologies*. CRC Press.

Fujii, T., Suzuki, M., Okuda, J., Ohtake, H., Tanji, K., Yamaguchi, K., Itoh, M., & Yamadori, A. (2004). Neural correlates of context memory with real-world events. *NeuroImage*, *21*(4), 1596–1603.

Fuller, P. M., Gooley, J. J., & Saper, C. B. (2006). Neurobiology of the sleep-wake cycle: Sleep architecture, circadian regulation, and regulatory feedback. In *Journal of Biological Rhythms*.

Furht, B. (Ed.). (2008). *Encyclopedia of multimedia* (2nd ed.). Springer Science & Business Media.

Furman, O., Dorfman, N., Hasson, U., Davachi, L., & Dudai, Y. (2007). They saw a movie: long-term memory for an extended audiovisual narrative. *Learning & Memory (Cold Spring Harbor, N.Y.)*.

Furman, O., Mendelsohn, A., & Dudai, Y. (2012). The episodic engram transformed: Time reduces retrieval-related brain activity but correlates it with memory accuracy. *Learning and Memory*, *19*(12), 575–587.

Fyhn, M., Molden, S., Hollup, S., Moser, M.-B., & Moser, E. I. (2002). Hippocampal Neurons Responding to First-Time Dislocation of a Target Object. *Neuron*, *35*(3), 555–566.

Gais, S., Lucas, B., & Born, J. (2006). Sleep after learning aids memory recall. *Learning & Memory*, *13*(3), 259–262.

Gais, S., Mölle, M., Helms, K., & Born, J. (2002). Learning-dependent increases in sleep spindle density. *Journal of Neuroscience*.

Gais, S., Plihal, W., Wagner, U., & Born, J. (2000). Early sleep triggers memory for early visual discrimination skills. *Nature Neuroscience*, *3*(12), 1335–1339.

Gallo, D. A. (2010). False memories and fantastic beliefs: 15 years of the DRM illusion. In *Memory and Cognition*.

Gamberini, L. (2000). Virtual reality as a new research tool for the study of human memory. *Cyberpsychology and Behavior*, *3*(3), 337–342.

Gardiner, J. M. (1988). Functional aspects of recollective experience. *Memory & Cognition*, *16*(4), 309–313.

Gardiner, J. M., Ramponi, C., & Richardson-Klavehn, A. (2002). Recognition memory and decision processes: A meta-analysis of remember, know, and guess responses. *Memory*, *10*(2), 83–98.

Genzel, L., Kroes, M. C. W., Dresler, M., & Battaglia, F. P. (2014). Light sleep versus slow wave sleep in memory consolidation: A question of global versus local processes? In *Trends in Neurosciences* (Vol. 37, Issue 1, pp. 10–19).

Gibbs, J. C. (1979). The meaning of ecologically oriented inquiry in contemporary psychology. *American Psychologist*, *34*(2), 127–140.

Gilboa, A. (2004). Remembering Our Past: Functional Neuroanatomy of Recollection of Recent and Very Remote Personal Events. *Cerebral Cortex*, *14*(11), 1214–1225.

Gioia, G. A., & Isquith, P. K. (2004). Ecological Assessment of Executive Function in Traumatic Brain Injury. *Developmental Neuropsychology*, *25*(1–2), 135–158.

Giuditta, A. (2014). Sleep memory processing: the sequential hypothesis. *Frontiers in Systems Neuroscience, 8*.

Giuditta, A., Ambrosini, M. V., Montagnese, P., Mandile, P., Cotugno, M., Zucconi, G. G., & Vescia, S. (1995). The sequential hypothesis of the function of sleep. *Behavioural Brain Research*, *69*(1–2), 157–166.

Golding, J. F. (1998). Motion sickness susceptibility questionnaire revised and its relationship to other forms of sickness. *Brain Research Bulletin*, *47*(5), 507–516.

Greenberg, D. L., & Verfaellie, M. (2010). Interdependence of episodic and semantic memory: Evidence from neuropsychology. *Journal of the International Neuropsychological Society*, *16*(5), 748–753.

Grewe, P., Lahr, D., Kohsik, A., Dyck, E., Markowitsch, H. J., Bien, C. G., Botsch, M., &

Piefke, M. (2014). Real-life memory and spatial navigation in patients with focal epilepsy: Ecological validity of a virtual reality supermarket task. *Epilepsy and Behavior*, *31*, 57–66.

Griessenberger, H., Hoedlmoser, K., Heib, D. P. J., Lechinger, J., Klimesch, W., & Schabus, M. (2012). Consolidation of temporal order in episodic memories. *Biological Psychology*, *91*(1), 150–155.

Grilli, M. D., Coste, S., Landry, J. E., & Mangen, K. (2019). Evidence that an episodic mode of thinking facilitates encoding of perceptually rich memories for naturalistic events relative to a gist-based mode of thinking. *Memory*, *27*(10), 1468–1474.

Groch, S., Wilhelm, I., Diekelmann, S., & Born, J. (2013). The role of REM sleep in the processing of emotional memories: Evidence from behavior and event-related potentials. *Neurobiology of Learning and Memory*, *99*, 1–9.

Grundgeiger, T., Bayen, U. J., & Horn, S. S. (2014). Effects of sleep deprivation on prospective memory. *Memory*, *22*(6), 679–686.

Guan, Z., Peng, X., & Fang, J. (2004). Sleep deprivation impairs spatial memory and decreases extracellular signal-regulated kinase phosphorylation in the hippocampus. *Brain Research*, *1018*(1), 38–47.

Guderian, S., Dzieciol, A. M., Gadian, D. G., Jentschke, S., Doeller, C. F., Burgess, N., Mishkin, M., & Vargha-Khadem, F. (2015). Hippocampal Volume Reduction in Humans Predicts Impaired Allocentric Spatial Memory in Virtual-Reality Navigation. *Journal of Neuroscience*, *35*(42), 14123–14131.

Guillery-Girard, B., Martins, S., Deshayes, S., Hertz-Pannier, L., Chiron, C., Jambaqué, I., Landeau, B., Clochon, P., Chételat, G., & Eustache, F. (2013). Developmental Trajectories of Associative Memory from Childhood to Adulthood: A Behavioral and Neuroimaging Study. *Frontiers in Behavioral Neuroscience*, *7*.

Guo, C., Duan, L., Li, W., & Paller, K. A. (2006). Distinguishing source memory and item memory: Brain potentials at encoding and retrieval. *Brain Research*, *1118*(1), 142–154.

Gutiérrez, F., Pierce, J., Vergara, V. M., Coulter, R., Saland, L., Caudell, T. P., Goldsmith, T. E., & Alverson, D. C. (2007). The effect of degree of immersion upon learning performance in virtual reality simulations for medical education. *Studies in Health*

Technology and Informatics, *125*, 155–160.

Haatveit, B. C., Sundet, K., Hugdahl, K., Ueland, T., Melle, I., & Andreassen, O. A. (2010). The validity of d prime as a working memory index: Results from the "Bergen n -back task." *Journal of Clinical and Experimental Neuropsychology*, *32*(8), 871–880.

Habib, M., & Sirigu, A. (1987). Pure Topographical Disorientation: A Definition and Anatomical Basis. *Cortex*, *23*(1), 73–85.

Hamann, S. B., Ely, T. D., Grafton, S. T., & Kilts, C. D. (1999). Amygdala activity related to enhanced memory for pleasant and aversive stimuli. *Nature Neuroscience*, *2*(3), 289–293.

Hamilton, D. A., & Sutherland, R. J. (1999). Blocking in human place learning: Evidence from virtual navigation. *Psychobiology*, *27*(4), 453–461.

Hampton, R. R., Hampstead, B. M., & Murray, E. A. (2005). Rhesus monkeys (Macaca mulatta) demonstrate robust memory for what and where, but not when, in an open-field test of memory. *Learning and Motivation*, *36*(2), 245–259.

Hampton, R. R., & Schwartz, B. L. (2004). Episodic memory in nonhumans: what, and where, is when? *Current Opinion in Neurobiology*, *14*(2), 192–197.

Harand, C., Bertran, F., la Joie, R., Landeau, B., Mézenge, F., Desgranges, B., Peigneux, P., Eustache, F., & Rauchs, G. (2012). The hippocampus remains activated over the long term for the retrieval of truly episodic memories. *PLoS ONE*.

Hardt, O., Nader, K., & Nadel, L. (2013). Decay happens: the role of active forgetting in memory. *Trends in Cognitive Sciences*, *17*(3), 111–120.

Harman, J., Brown, R., & Johnson, D. (2017). Improved Memory Elicitation in Virtual Reality: New Experimental Results and Insights. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (pp. 128–146).

Harman, K. L., Humphrey, G. K., & Goodale, M. A. (1999). Active manual control of object views facilitates visual recognition. *Current Biology*, *9*(22), 1315–1318.

Hasher, L., Chung, C., May, C. P., & Foong, N. (2002). Age, time of testing, and proactive interference. *Canadian Journal of Experimental Psychology/Revue Canadienne de*

Psychologie Expérimentale, *56*(3), 200–207.

Hashtroudi, S., Johnson, M. K., & Chrosniak, L. D. (1990). Aging and qualitative characteristics of memories for perceived and imagined complex events. *Psychology and Aging*, *5*(1), 119–126.

Hassabis, D., & Maguire, E. A. (2007). Deconstructing episodic memory with construction. *Trends in Cognitive Sciences*, *11*(7), 299–306.

Hassabis, D., & Maguire, E. A. (2009). The construction system of the brain. *Philosophical Transactions of the Royal Society B: Biological Sciences*.

Hauer, B. J. A., Wessel, I., & Merckelbach, H. (2006). Intrusions, avoidance and overgeneral memory in a non-clinical sample. *Clinical Psychology & Psychotherapy*, *13*(4), 264–268.

Hauner, K. K., Howard, J. D., Zelano, C., & Gottfried, J. A. (2013). Stimulus-specific enhancement of fear extinction during slow-wave sleep. *Nature Neuroscience*, *16*(11), 1553–1555.

Hayes, S. M., Ryan, L., Schnyer, D. M., & Nadel, L. (2004). An fMRI Study of Episodic Memory: Retrieval of Object, Spatial, and Temporal Information. *Behavioral Neuroscience*, *118*(5), 885–896.

Hayne, H., & Imuta, K. (2011). Episodic memory in 3- and 4-year-old children. *Developmental Psychobiology*, *53*(3), 317–322.

Heine, R. (1914). Uber der Wiedererkennen und rückwirkende Hemmung. *Zeitschrift Für Psychologie*, *68*, 161–236.

Helbing, J., Draschkow, D., & Võ, M. L.-H. L. H. (2020). Search superiority: Goal-directed attentional allocation creates more reliable incidental identity and location memory than explicit encoding in naturalistic virtual environments. *Cognition*, *196*(August 2019), 104147.

Held, R. (1992). Telepresence. In *The Journal of the Acoustical Society of America* (Vol. 92, Issue 4, p. 2458).

Herbert, D. M. ., & Burt, J. S. (2003). The effects of different review opportunities on schematisation of knowledge. *Learning and Instruction*, *13*(1), 73–92.

Herbert, D. M. B., & Burt, J. S. (2001). Memory awareness and schematization: learning in the university context. *Applied Cognitive Psychology*, *15*(6), 617–637.

Herbert, D. M. B., & Burt, J. S. (2004). What do students remember? Episodic memory and the development of schematization. *Applied Cognitive Psychology*, *18*(1), 77–88.

Herlitz, A., Nilsson, L.-G. G., & Bäckman, L. (1997). Gender differences in episodic memory. *Memory and Cognition*, *25*(6), 801–811.

Herron, J. ., & Wilding, E. . (2004). An electrophysiological dissociation of retrieval mode and retrieval orientation. *NeuroImage*, *22*(4), 1554–1562.

Hirshman, E., & Henzler, A. (1998). The Role of Decision Processes in Conscious Recollection. *Psychological Science*, *9*(1), 61–65.

Hoc, J.-M. (2001). Towards ecological validity of research in cognitive ergonomics. *Theoretical Issues in Ergonomics Science*, *2*(3), 278–288.

Hoffman, H. G., Garcia-Palacios, A., Thomas, A. K., & Schmidt, A. (2001). Virtual Reality Monitoring: Phenomenal Characteristics of Real, Virtual, and False Memories. *CyberPsychology & Behavior*, *4*(5), 565–572.

Hokkanen, L., Launes, J., Vataja, R., Valanne, L., & Iivanainen, M. (1995). Isolated retrograde amnesia for autobiographical material associated with acute left temporal lobe encephalitis. *Psychological Medicine*, *25*(1), 203–208.

Holland, S. M., & Smulders, T. V. (2011). Do humans use episodic memory to solve a What-Where-When memory task? *Animal Cognition*, *14*(1), 95–102.

Hornberger, M., & Piguet, O. (2012). Episodic memory in frontotemporal dementia: A critical review. *Brain*, *135*(3), 678–692.

Horner, A. J., Bisby, J. A., Bush, D., Lin, W.-J. J., & Burgess, N. (2015). Evidence for holistic episodic recollection via hippocampal pattern completion. *Nature Communications*, *6*(May), 7462.

Horner, A. J., Bisby, J. A., Zotow, E., Bush, D., & Burgess, N. (2016). Grid-like processing of imagined navigation. *Current Biology*, *26*(6), 842–847.

Horner, A. J., & Burgess, N. (2013). The associative structure of memory for multi-element events. *Journal of Experimental Psychology: General*, *142*(4), 1370–1383.

Horner, A. J., & Burgess, N. (2014). Pattern Completion in Multielement Event Engrams. *Current Biology*, *24*(9), 988–992.

Horton, A. M., Monahan, M. C., & Fennell, E. B. (2001). Children's Memory Scale. *Archives of Clinical Neuropsychology*, *16*(2), 193–198.

Howard, M. W., & Eichenbaum, H. (2013). The hippocampus, time, and memory across scales. *Journal of Experimental Psychology: General*, *142*(4), 1211–1230.

Howard, M. W., & Kahana, M. J. (2002). When Does Semantic Similarity Help Episodic Retrieval? *Journal of Memory and Language*, *46*(1), 85–98.

HTC Corporation. (2015). HTC Vive. *Htcvr.Com*, 1–6.

Huber, R., Felice Ghilardi, M., Massimini, M., & Tononi, G. (2004). Local sleep and learning. *Nature*, *430*(6995), 78–81.

Humphreys, M. S., Smith, S., Pachana, N. A., Tehan, G., & Byrne, G. J. (2010). Measuring episodic memory: A novel approach with an indefinite number of alternative forms. *Applied Cognitive Psychology*, *24*(8), 1080–1094.

Hunt, R. R., & Mcdaniel, M. A. (1993). The Enigma of Organization and Distinctiveness. *Journal of Memory and Language*, *32*(4), 421–445.

Husserl, E. (1964). *The Phenomenology of Internal Time-Consciousness*. Indiana University Press.

Hutchison, I. C., & Rathore, S. (2015). The role of REM sleep theta activity in emotional memory. *Frontiers in Psychology*, *6*.

Hyman, I. E., & Loftus, E. F. (1998). Errors in autobiographical memory. *Clinical Psychology Review*, *18*(8), 933–947.

Iber, C., Ancoli-Israel, S., Chesson, A., & Quan, S. F. (2007). The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specification. In *Journal of Clinical Sleep Medicine*.

Idzikowski, C. (1984). Sleep and Memory. *British Journal of Psychology*, 439–449.

Igroup Project Consortium. (2015). *Igroup Presence Questionnaire*. Igroup.Org.

Iidaka, T., Anderson, N. D., Kapur, S., Cabez, R., & Craik, F. I. M. (2000). The Effect of

Divided Attention on Encoding and Retrieval in Episodic Memory Revealed by Positron Emission Tomography. *Journal of Cognitive Neuroscience*, *12*(2), 267–280.

Inostroza, M., & Born, J. (2013). Sleep for Preserving and Transforming Episodic Memory. *Annual Review of Neuroscience*, *36*(1), 79–102.

Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language*.

James, K. H., Humphrey, G. K., Vilis, T., Corrie, B., Baddour, R., & Goodale, M. A. (2002). “Active” and “passive” learning of three-dimensional object structure within an immersive virtual reality environment. *Behavior Research Methods, Instruments, & Computers*, *34*(3), 383–390.

Jansari, A. S., Davis, K., McGibbon, T., Firminger, S., & Kapur, N. (2010). When “long-term memory” no longer means “forever”: Analysis of accelerated long-term forgetting in a patient with temporal lobe epilepsy. *Neuropsychologia*, *48*(6), 1707–1715.

JASP Team. (2019). *JASP*.

Jebara, N., Orriols, E., Zaoui, M., Berthoz, A., & Piolino, P. (2014). Effects of Enactment in Episodic Memory: A Pilot Virtual Reality Study with Young and Elderly Adults. *Frontiers in Aging Neuroscience*, *6*(DEC), 338.

Jenkins, J. G., & Dallenbach, K. M. (1924). Obliviscence during Sleep and Waking. *The American Journal of Psychology*, *35*(4), 605–612.

Jennings, J. M., & Jacoby, L. L. (1993). Automatic versus intentional uses of memory: Aging, attention, and control. *Psychology and Aging*, *8*(2), 283–293.

Jensen, L., & Konradsen, F. (2018). A review of the use of virtual reality head-mounted displays in education and training. *Education and Information Technologies*, *23*(4), 1515–1529.

Ji, D., & Wilson, M. a. (2007). Coordinated memory replay in the visual cortex and hippocampus during sleep. *Nature Neuroscience*, *10*(1), 100–107.

Johnson, M. K., & Raye, C. L. (1981). Reality monitoring. *Psychological Review*, *88*(1), 67–85.

Kandel, E. R. (2001). The Molecular Biology of Memory Storage: A Dialogue Between

Genes and Synapses. *Science*, *294*(5544), 1030–1038.

Kapur, N., Millar, J., Colbourn, C., Abbott, P., Kennedy, P., & Docherty, T. (1997). Very Long-Term Amnesia in Association with Temporal Lobe Epilepsy: Evidence for Multiple-Stage Consolidation Processes. *Brain and Cognition*, *35*(1), 58–70.

Karpicke, J. D., & Roediger, H. L. (2008). The Critical Importance of Retrieval for Learning. *Science*, *319*(5865), 966–968.

Kennedy, R. S., Lane, N. E., Berbaum, K. S., & Lilienthal, M. G. (1993). Simulator Sickness Questionnaire: An Enhanced Method for Quantifying Simulator Sickness. *The International Journal of Aviation Psychology*, *3*(3), 203–220.

Kensinger, E. A. (2009). Remembering the details: Effects of emotion. In *Emotion Review* (Vol. 1, Issue 2, pp. 99–113).

Kensinger, E. A., & Ford, J. H. (2020). Retrieval of Emotional Events from Memory. *Annual Review of Psychology*, *71*(1), 251–272.

Kessels, R. P. C., Hobbel, D., & Postma, A. (2007). Aging, context memory and binding: a comparison of “what, where and when” in young and older adults. *International Journal of Neuroscience*, *117*(6), 795–810.

Kilb, A., & Naveh-Benjamin, M. (2007). Paying attention to binding: Further studies assessing the role of reduced attentional resources in the associative deficit of older adults. *Memory & Cognition*, *35*(5), 1162–1174.

Kim, H., Park, J. Y., & Kim, K. K. (2018). Spatial Learning and Memory Using a Radial Arm Maze with a Head-Mounted Display. *Psychiatry Investigation*, *15*(10), 935–944.

King, J. A., Burgess, N., Hartley, T., Vargha-Khadem, F., & O’Keefe, J. (2002). Human hippocampus and viewpoint dependence in spatial memory. In *Hippocampus* (Vol. 12, Issue 6, pp. 811–820).

Kinsbourne, M., & Wood, F. (1975). *Short-term memory processes and the amnesic syndrome* (D. Deutsch & J. A. Deutsch (Eds.)). Academic Press.

Kinugawa, K., Schumm, S., Pollina, M., Depre, M., Jungbluth, C., Doulazmi, M., Sebban, C., Zlomuzica, A., Pietrowsky, R., Pause, B., Mariani, J., & Dere, E. (2013). Aging-related episodic memory decline: are emotions the key? *Frontiers in Behavioral Neuroscience*.

Kirchhoff, B. A., Wagner, A. D., Maril, A., & Stern, C. E. (2000). Prefrontal–Temporal Circuitry for Episodic Encoding and Subsequent Memory. *The Journal of Neuroscience*, *20*(16), 6173–6180.

Kisker, J., Gruber, T., & Schöne, B. (2019). Experiences in virtual reality entail different processes of retrieval as opposed to conventional laboratory settings: A study on human memory. *Current Psychology*, 1–8.

Klein, S. B. (2013). Making the case that episodic recollection is attributable to operations occurring at retrieval rather than to content stored in a dedicated subsystem of long-term memory. *Frontiers in Behavioral Neuroscience*, *7*.

Knowlton, B. J., & Squire, L. R. (1995). Remembering and knowing: Two different expressions of declarative memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(3), 699–710.

Kober, S. E., & Neuper, C. (2012). Using auditory event-related EEG potentials to assess presence in virtual reality. *International Journal of Human-Computer Studies*, *70*(9), 577–587.

Konishi, K., Mckenzie, S., Etchamendy, N., Roy, S., & Bohbot, V. D. (2017). Hippocampus-dependent spatial learning is associated with higher global cognition among healthy older adults. *Neuropsychologia*, *106*, 310–321.

Kopelman, M. D., Wilson, B. A., & Baddeley, A. D. (1989). The autobiographical memory interview: a new assessment of autobiographical and personal semantic memory in amnesic patients. *Journal of Clinical and Experimental Neuropsychology : Official Journal of the International Neuropsychological Society*, *11*(5), 724–744.

Koulack, D. (1997). Recognition memory, circadian rhythms, and sleep. *Percept Mot Skills*, *85*(1), 99–104.

Krokos, E., Plaisant, C., & Varshney, A. (2019). Virtual memory palaces: immersion aids recall. *Virtual Reality*, *23*(1), 1–15.

Kumaran, D., Hassabis, D., & McClelland, J. L. (2016). What Learning Systems do Intelligent Agents Need? Complementary Learning Systems Theory Updated. *Trends in Cognitive Sciences*, *20*(7), 512–534.

Kwan, D., Carson, N., Addis, D. R., & Rosenbaum, R. S. (2010). Deficits in past remembering extend to future imagining in a case of developmental amnesia. *Neuropsychologia*, *48*(11), 3179–3186.

LaBar, K. S., & Cabeza, R. (2006). Cognitive neuroscience of emotional memory. *Nature Reviews Neuroscience*, *7*(1), 54–64.

Lechner, H. A., Squire, L. R., & Byrne, J. H. (1999). 100 Years of Consolidation—Remembering Müller and Pilzecker. *Learning & Memory*, *6*(2), 77–87.

Lee, A. K., & Wilson, M. A. (2002). Memory of Sequential Experience in the Hippocampus during Slow Wave Sleep. *Neuron*, *36*(6), 1183–1194.

Lee, E. A.-L., & Wong, K. W. (2014). Learning with desktop virtual reality: Low spatial ability learners are more positively affected. *Computers and Education*, *79*, 49–58.

Lehmann, M., Schreiner, T., Seifritz, E., & Rasch, B. (2016). Emotional arousal modulates oscillatory correlates of targeted memory reactivation during NREM, but not REM sleep. *Scientific Reports*, *6*(1), 39229.

Lepage, M., Ghaffar, O., Nyberg, L., & Tulving, E. (2000). Prefrontal cortex and episodic memory retrieval mode. *Proceedings of the National Academy of Sciences*, *97*(1), 506–511.

Leplow, B., Holl, D., Zeng, L., & Mehdorn, M. (1998). Spatial orientation and spatial memory within a `locomotor' maze for humans. *Spatial Cognition*, 429–446.

Lessiter, J., Freeman, J., Keogh, E., & Davidoff, J. (2001). A Cross-Media Presence Questionnaire: The ITC-Sense of Presence Inventory. *Presence: Teleoperators and Virtual Environments*, *10*(3), 282–297.

Levine, B. (1998). Episodic memory and the self in a case of isolated retrograde amnesia. *Brain*, *121*(10), 1951–1973.

Levine, B., Svoboda, E., Hay, J. F., Winocur, G., & Moscovitch, M. (2002). Aging and autobiographical memory: Dissociating episodic from semantic retrieval. *Psychology and Aging*, *17*(4), 677–689.

Lewis, P. A., & Durrant, S. J. (2011). Overlapping memory replay during sleep builds cognitive schemata. In *Trends in Cognitive Sciences* (Vol. 15, Issue 8, pp. 343–351).

Li, C.-L., Aivar, M. P., Kit, D. M., Tong, M. H., & Hayhoe, M. M. (2016). Memory and visual search in naturalistic 2D and 3D environments. *Journal of Vision*, *16*(8), 9.

Lin, J. J.-W., Duh, H. B. L., Parker, D. E., Abi-Rached, H., & Furness, T. A. (2002). Effects of field of view on presence, enjoyment, memory, and simulator sickness in a virtual environment. *Proceedings IEEE Virtual Reality 2002*, 164–171.

Lisman, J. E., & Grace, A. A. (2005). The Hippocampal-VTA Loop: Controlling the Entry of Information into Long-Term Memory. *Neuron*, *46*(5), 703–713.

Lloyd, J., Persaud, N. V, & Powell, T. E. (2009). Equivalence of real-world and virtual-reality route learning: a pilot study. *Cyberpsychology & Behavior : The Impact of the Internet, Multimedia and Virtual Reality on Behavior and Society*, *12*(4), 423–427.

Long, C. J. (1996). Neuropsychological tests: A look at our past and the impact that ecological issues may have on our future. In C. J. Long & R. J. Sbordone (Eds.), *Ecological validity of neuropsychological testing* (pp. 1–14). Gr Press/St Lucie Press, Inc.

Louie, K., & Wilson, M. A. (2001). Temporally structured replay of awake hippocampal ensemble activity during rapid eye movement sleep. *Neuron*, *29*(1), 145–156.

Lundstrom, B. N., Ingvar, M., & Petersson, K. M. (2005). The role of precuneus and left inferior frontal cortex during source memory episodic retrieval. *NeuroImage*, *27*(4), 824–834.

Machery, E. (2012). Dissociations in Neuropsychology and Cognitive Neuroscience. *Philosophy of Science*, *79*(4), 490–518.

Mackey, W. E., & Curtis, C. E. (2017). Distinct contributions by frontal and parietal cortices support working memory. *Scientific Reports*, *7*(1), 6188.

Mackiewicz, K. L., Sarinopoulos, I., Cleven, K. L., & Nitschke, J. B. (2006). The effect of anticipation and the specificity of sex differences for amygdala and hippocampus function in emotional memory. *Proceedings of the National Academy of Sciences*, *103*(38), 14200–14205.

Macmillan, N. a N. A., & Creelman, C. D. (2005). Detection Theory: A User's Guide. In *Detection theory: A user's guide (2nd ed.).*

MacPherson, S. E., Bozzali, M., Cipolotti, L., Dolan, R. J., Rees, J. H., & Shallice, T. (2008).

Effect of frontal lobe lesions on the recollection and familiarity components of recognition memory. *Neuropsychologia*, *46*(13), 3124–3132.

MacPherson, S. E., Turner, M. S., Bozzali, M., Cipolotti, L., & Shallice, T. (2016). The Doors and People Test: The effect of frontal lobe lesions on recall and recognition memory performance. *Neuropsychology*, *30*(3), 332–337.

Maguire, E. A., Burgess, N., Donnett, J. G., Frackowiak, R. S. J., Frith, C. D., & O'Keefe, J. (1998). Knowing where and getting there: A human navigation network. *Science*, *280*(5365), 921–924.

Maguire, E. A., Frith, C. D., Burgess, N., Donnett, J. G., & O'Keefe, J. (1998). Knowing Where Things Are: Parahippocampal Involvement in Encoding Object Locations in Virtual Large-Scale Space. *Journal of Cognitive Neuroscience*, *10*(1), 61–76.

Maguire, E. A., & Mullally, S. L. (2013). The hippocampus: A manifesto for change. *Journal of Experimental Psychology: General*, *142*(4), 1180–1189.

Maguire, E. A., & Mummery, C. J. (1999). Differential modulation of a common memory retrieval network revealed by positron emission tomography. *Hippocampus*, *9*(1), 54–61.

Maillot, P., Dommes, A., Dang, N.-T., & Vienne, F. (2017). Training the elderly in pedestrian safety: Transfer effect between two virtual reality simulation devices. *Accident Analysis & Prevention*, *99*, 161–170.

Makowski, D., Sperduti, M., Nicolas, S., & Piolino, P. (2017). "Being there" and remembering it: Presence improves memory encoding. *Consciousness and Cognition*, *53*, 194–202.

Makransky, G., Terkildsen, T. S., & Mayer, R. E. (2019). Adding immersive virtual reality to a science lab simulation causes more presence but less learning. *Learning and Instruction*, *60*, 225–236.

Mandler, G. (1980). Recognizing: The judgment of previous occurrence. *Psychological Review*, *87*(3), 252–271.

Mandler, G. (2008). Familiarity Breeds Attempts: A Critical Review of Dual-Process Theories of Recognition. *Perspectives on Psychological Science*, *3*(5), 390–399.

Mania, K., & Chalmers, A. (2001). The Effects of Levels of Immersion on Memory and Presence in Virtual Environments: A Reality Centered Approach. *CyberPsychology & Behavior*, *4*(2), 247–264.

Mania, K., Troscianko, T., Hawkes, R., & Chalmers, A. (2003). Fidelity metrics for virtual environment simulations based on spatial memory awareness states. *Presence: Teleoperators and Virtual Environments*, *12*(3), 296–310.

Maples-Keller, J. L., Bunnell, B. E., Kim, S.-J., & Rothbaum, B. O. (2017). The Use of Virtual Reality Technology in the Treatment of Anxiety and Other Psychiatric Disorders. *Harvard Review of Psychiatry*, *25*(3), 103–113.

Markowitsch, H. J., & Staniloiu, A. (2011). Memory, autonoetic consciousness, and the self. *Consciousness and Cognition*, *20*(1), 16–39.

Marks, D. F. (1973). Visual Imagery Differences in the Recall of Pictures. *British Journal of Psychology*, *64*(1), 17–24.

Marsh, E., & Roediger, H. (2012). Episodic and Autobiographical Memory. In *Handbook of Psychology* (Second Edi, pp. 472–494).

Marshall, L., & Born, J. (2007). The contribution of sleep to hippocampus-dependent memory consolidation. In *Trends in Cognitive Sciences* (Vol. 11, Issue 10, pp. 442–450).

Marshall, L., Helgadóttir, H., Mölle, M., & Born, J. (2006). Boosting slow oscillations during sleep potentiates memory. *Nature*, *444*(7119), 610–613.

Marshall, L., Kirov, R., Brade, J., Mölle, M., & Born, J. (2011). Transcranial Electrical Currents to Probe EEG Brain Rhythms and Memory Consolidation during Sleep in Humans. *PLoS ONE*, *6*(2), e16905.

Martin-Ordas, G., & Atance, C. M. (2019). Adults’ Performance in an Episodic-Like Memory Task: The Role of Experience. *Frontiers in Psychology*, *9*(JAN), 2688.

Martin-Ordas, G., Atance, C. M., & Caza, J. (2017). Did the popsicle melt? Preschoolers’ performance in an episodic-like memory task. *Memory*, *25*(9), 1260–1271.

Martinelli, P., Anssens, A., Sperduti, M., & Piolino, P. (2013). The influence of normal aging and Alzheimer’s disease in autobiographical memory highly related to the self.

Neuropsychology, *27*(1), 69–78.

Martini, M., Zamarian, L., Sachse, P., Martini, C., & Delazer, M. (2019). Wakeful resting and memory retention: a study with healthy older and younger adults. *Cognitive Processing*, *20*(1), 125–131.

May, C. P., Hasher, L., & Stoltzfus, E. R. (1993). Optimal Time of Day and the Magnitude of Age Differences in Memory. *Psychological Science*, *4*(5), 326–330.

Mayes, A. R., Holdstock, J. S., Isaac, C. L., Hunkin, N. M., & Roberts, N. (2002). Relative sparing of item recognition memory in a patient with adult-onset damage limited to the hippocampus. *Hippocampus*, *12*(3), 325–340.

Mazurek, A., Bhoopathy, R. M., Read, J. C. A., Gallagher, P., & Smulders, T. V. (2015). Effects of age on a real-world What-Where-When memory task. *Frontiers in Aging Neuroscience*, *7*(APR), 0–17.

McCabe, D. P., Geraci, L., Boman, J. K., Sensenig, A. E., & Rhodes, M. G. (2011). On the validity of remember–know judgments: Evidence from think aloud protocols. *Consciousness and Cognition*, *20*(4), 1625–1633.

McCabe, D. P., Langer, K. G., Borod, J. C., & Bender, H. A. (2011). Practice Effects. In J. S. Kreutzer, J. DeLuca, & B. Caplan (Eds.), *Encyclopedia of Clinical Neuropsychology* (pp. 1988–1989). Springer New York.

McCabe, D. P., Roediger, H. L., & Karpicke, J. D. (2011). Automatic processing influences free recall: converging evidence from the process dissociation procedure and remember-know judgments. *Memory & Cognition*, *39*(3), 389–402.

McCarley, R. W. (2007). Neurobiology of REM and NREM sleep. *Sleep Medicine*, *8*(4), 302–330.

McCarthy, R. A., Evans, J. J., & Hodges, J. R. (1996). Topographic amnesia: Spatial memory disorder, perceptual dysfunction, or category specific semantic memory impairment? *Journal of Neurology Neurosurgery and Psychiatry*, *60*(3), 318–325.

McClelland, J. L., McNaughton, B. L., & O’Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: *Psychological Review*, *102*(3), 419–457.

McElree, B., Dolan, P. O., & Jacoby, L. L. (1999). Isolating the contributions of familiarity and source information to item recognition: A time course analysis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*(3), 563–582.

McGeoch, J. A. (1932). Forgetting and the law of disuse. *Psychological Review*, *39*(4), 352–370.

McNair, D. M., Kahn, R. J., Crook, S. F. T., & Bartus, R. (1983). The Cognitive Difficulties Scale. *Assessment in Geriatric Psychopharmacology*, *137*, 143.

Meeter, M., & Murre, J. M. J. (2004). Consolidation of Long-Term Memory: Evidence and Alternatives. *Psychological Bulletin*, *130*(6), 843–857.

Mestre, D., & Vercher, J.-L. (2011). Immersion and Presence. *Virtual Reality: Concepts and Technologies*, 81–96.

Meyers, J. E., & Meyers, K. R. (1995). Rey complex figure test under four different administration procedures. *The Clinical Neuropsychologist*, *9*(1), 63–67.

Mickes, L., Seale-Carlisle, T. M., & Wixted, J. T. (2013). Rethinking familiarity: Remember/Know judgments in free recall. *Journal of Memory and Language*, *68*(4), 333–349.

Mickley Steinmetz, K. R., Schmidt, K., Zucker, H. R., & Kensinger, E. A. (2012). The effect of emotional arousal and retention delay on subsequent-memory effects. *Cognitive Neuroscience*, *3*(3–4), 150–159.

Migo, E. M., Mayes, A. R., & Montaldi, D. (2012). Measuring recollection and familiarity : Improving the remember / know procedure. *Consciousness and Cognition*, *21*(3), 1435–1455.

Miller, J. F., Neufang, M., Solway, A., Brandt, A., Trippel, M., Mader, I., Hefft, S., Merkow, M., Polyn, S. M., Jacobs, J., Kahana, M. J., & Schulze-Bonhage, A. (2013). Neural Activity in Human Hippocampal Formation Reveals the Spatial Context of Retrieved Memories. *Science*, *342*(6162), 1111–1114.

Mohr, G., Engelkamp, J., & Zimmer, H. D. (1989). Recall and recognition of self-performed acts. *Psychological Research*, *51*(4), 181–187.

Montaldi, D., & Mayes, A. R. (2011). Familiarity, recollection and medial temporal lobe function: an unresolved issue. *Trends in Cognitive Sciences*, *15*(8), 339–340.

Moreton, B. J., & Ward, G. (2010). Time scale similarity and long-term memory for autobiographical events. *Psychonomic Bulletin and Review*, *17*(4), 510–515.

Morgenthaler, J., Wiesner, C. D., Hinze, K., Abels, L. C., Prehn-Kristensen, A., & Göder, R. (2014). Selective REM-Sleep Deprivation Does Not Diminish Emotional Memory Consolidation in Young Healthy Subjects. *PLoS ONE*, *9*(2), e89849.

Morina, N., Ijntema, H., Meyerbröker, K., & Emmelkamp, P. M. G. (2015). Can virtual reality exposure therapy gains be generalized to real-life? A meta-analysis of studies applying behavioral assessments. *Behaviour Research and Therapy*, *74*, 18–24.

Morris, R. G. M., & Frey, U. (1997). Hippocampal synaptic plasticity: role in spatial learning or the automatic recording of attended experience? *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *352*(1360), 1489–1503.

Moscovitch, M. (1992). Memory and Working-with-Memory: A Component Process Model Based on Modules and Central Systems. *Journal of Cognitive Neuroscience*, *4*(3), 257–267.

Moscovitch, M. (1995). Recovered consciousness: A hypothesis concerning modularity and episodic memory. *Journal of Clinical and Experimental Neuropsychology*, *17*(2), 276–290.

Moscovitch, M., Cabeza, R., Winocur, G., & Nadel, L. (2016). Episodic Memory and Beyond: The Hippocampus and Neocortex in Transformation. *Annual Review of Psychology*, *67*(1), 105–134.

Moscovitch, M., & Nadel, L. (1998). Consolidation and the hippocampal complex revisited: in defense of the multiple-trace model. *Current Opinion in Neurobiology*, *8*(2), 297–300.

Moscovitch, M., & Winocur, G. (1992). The neuropsychology of memory and aging. *The Handbook of Aging and Cognition*.

Mou, W., McNamara, T. P., Valiquette, C. M., & Rump, B. (2004). Allocentric and Egocentric Updating of Spatial Memories. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*(1), 142–157.

Murcia-López, M., & Steed, A. (2016). The effect of environmental features, self-avatar, and immersion on object location memory in virtual environments. *Frontiers in ICT*, *3*(NOV), 24.

Murphy, K. J., Troyer, A. K., Levine, B., & Moscovitch, M. (2008). Episodic, but not semantic, autobiographical memory is reduced in amnestic mild cognitive impairment. *Neuropsychologia*, *46*(13), 3116–3123.

Nadel, L. (2008). The Hippocampus and Context Revisited. In *Hippocampal Place Fields* (pp. 3–15). Oxford University Press.

Nadel, L., & Moscovitch, M. (1997). Memory consolidation, retrograde amnesia and the hippocampal complex. In *Current Opinion in Neurobiology* (Vol. 7, Issue 2, pp. 217–227).

Nadel, L., & Moscovitch, M. (1998). Hippocampal contributions to cortical plasticity. *Neuropharmacology*, *37*(4–5), 431–439.

Nadel, L., & Peterson, M. A. (2013). The hippocampus: Part of an interactive posterior representational system spanning perceptual and memorial systems. *Journal of Experimental Psychology: General*, *142*(4), 1242–1254.

Nadel, L., Samsonovich, A., Ryan, L., & Moscovitch, M. (2000). Multiple trace theory of human memory: Computational, neuroimaging, and neuropsychological results. *Hippocampus*, *10*(4), 352–368.

Nash, E. B., Edwards, G. W., Thompson, J. A., & Barfield, W. (2000). A Review of Presence and Performance in Virtual Environments. *International Journal of Human-Computer Interaction*, *12*(1), 1–41.

Nasreddine, Z. S., Phillips, N. A., Badirian, V., Charbonneau, S., Whitehead, V., Collin, I., Cummings, J. L., & Chertkow, H. (2005). The Montreal Cognitive Assessment, MoCA: A Brief Screening Tool For Mild Cognitive Impairment. *Journal of the American Geriatrics Society*, *53*(4), 695–699.

Naveh-Benjamin, M. (2000). Adult age differences in memory performance: Tests of an associative deficit hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(5), 1170–1187.

Neath, I. (1993a). Distinctiveness and serial position effects in recognition. *Memory &*

Cognition, *21*(5), 689–698.

Neath, I. (1993b). Contextual and Distinctive Processes and the Serial Position Function. *Journal of Memory and Language*, *32*(6), 820–840.

Neisser, U. (1985). The role of theory in the ecological study of memory: Comment on Bruce. *Journal of Experimental Psychology: General*, *114*(2), 272–276.

Newman, E. B. (1939). Forgetting of meaningful material during sleep and waking. *Am J Psychol*, *52*, 65–71.

Nielsen, T., O'Reilly, C., Carr, M., Dumel, G., Godin, I., Solomonova, E., Lara-Carrasco, J., Blanchette-Carrière, C., & Paquette, T. (2015). Overnight improvements in two REM sleep-sensitive tasks are associated with both REM and NREM sleep changes, sleep spindle features, and awakenings for dream recall. *Neurobiology of Learning and Memory*, *122*, 88–97.

Nilsson, L.-G. (2000). Remembering actions and words. In E. Tulving & F. I. M. Craik (Eds.), *The Oxford handbook of memory.* (pp. 137–148). Oxford University Press.

Nishida, M., Pearsall, J., Buckner, R. L., & Walker, M. P. (2009). REM Sleep, Prefrontal Theta, and the Consolidation of Human Emotional Memory. *Cerebral Cortex*, *19*(5), 1158–1166.

Nyberg, L., McIntosh, A. R., Cabeza, R., Habib, R., Houle, S., & Tulving, E. (1996). General and specific brain regions involved in encoding and retrieval of events: what, where, and when. *Proceedings of the National Academy of Sciences*, *93*(20), 11280–11285.

Nyberg, L., Tulving, E., Habib, R., Nilsson, L.-G., Kapur, S., Houle, S., Cabeza, R., & McIntosh, A. R. (1995). Functional brain maps of retrieval mode and recovery of episodic information. *NeuroReport*, *7*(1), 249–252.

O'Keefe, J., & Nadel, L. (1978). *The Hippocampus as a Cognitive Map*. Oxford: Clarendon Press.

O'Neill, J., Pleydell-Bouverie, B., Dupret, D., & Csicsvari, J. (2010). Play it again: reactivation of waking experience and memory. In *Trends in Neurosciences* (Vol. 33, Issue 5, pp. 220–229).

Old, S. R., & Naveh-Benjamin, M. (2008). Differential effects of age on item and associative measures of memory: A meta-analysis. *Psychology and Aging*, *23*(1), 104–118.

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*.

Opriş, D., Pintea, S., García-Palacios, A., Botella, C., Szamosközi, Ş., & David, D. (2012). Virtual reality exposure therapy in anxiety disorders: A quantitative meta-analysis. *Depression and Anxiety*, *29*(2), 85–93.

Ormer, E. B. (1933). Sleep and retention. *Psychological Bulletin*, *30*, 415–439.

Ouellet, É., Boller, B., Corriveau-Lecavalier, N., Cloutier, S., & Belleville, S. (2018). The Virtual Shop: A new immersive virtual reality environment and scenario for the assessment of everyday memory. *Journal of Neuroscience Methods*, *303*, 126–135.

Oyanedel, C. N., Sawangjit, A., Born, J., & Inostroza, M. (2019). Sleep-dependent consolidation patterns reveal insights into episodic memory structure. *Neurobiology of Learning and Memory*, *160*, 67–72.

Padilla-Walker, L. M., & Poole, D. A. (2002). Memory for previous recall: A comparison of free and cued recall. *Applied Cognitive Psychology*, *16*(5), 515–524.

Pan, S. C., & Rickard, T. C. (2015). Sleep and motor learning: Is there room for consolidation? *Psychological Bulletin*, *141*(4), 812–834.

Parsons, T. D. (2015). Virtual Reality for Enhanced Ecological Validity and Experimental Control in the Clinical, Affective and Social Neurosciences. *Frontiers in Human Neuroscience*, *9*.

Parsons, T. D., Carlew, A. R., Magtoto, J., & Stonecipher, K. (2017). The potential of function-led virtual environments for ecologically valid measures of executive function in experimental and clinical neuropsychology. *Neuropsychological Rehabilitation*, *27*(5), 777–807.

Parsons, T. D., & McMahan, T. (2017). An initial validation of the Virtual Environment Grocery Store. *Journal of Neuroscience Methods*, *291*, 13–19.

Parsons, T. D., & Rizzo, A. A. (2008a). Affective outcomes of virtual reality exposure therapy for anxiety and specific phobias: A meta-analysis. *Journal of Behavior Therapy* *and Experimental Psychiatry*, *39*(3), 250–261.

Parsons, T. D., & Rizzo, A. A. (2008b). Initial validation of a virtual environment for assessment of memory functioning: virtual reality cognitive performance assessment test. *Cyberpsychology & Behavior : The Impact of the Internet, Multimedia and Virtual Reality on Behavior and Society*, *11*(1), 17–25.

Patrick, G. T. W., & Gilbert, J. A. (1896). Studies from the psychological laboratory of the University of Iowa: On the effects of loss of sleep. *Psychological Review*, *3*(5), 469–483.

Pause, B. M., Jungbluth, C., Adolph, D., Pietrowsky, R., & Dere, E. (2010). Induction and measurement of episodic memories in healthy adults. *Journal of Neuroscience Methods*, *189*(1), 88–96.

Pause, B. M., Zlomuzica, A., Kinugawa, K., Mariani, J., Pietrowsky, R., & Dere, E. (2013). Perspectives on Episodic-Like and Episodic Memory. *Frontiers in Behavioral Neuroscience*, *7*(April), 33.

Pavlides, C., & Winson, J. (1989). Influences of hippocampal place cell firing in the awake state on the activity of these cells during subsequent sleep episodes. *The Journal of Neuroscience*, *9*(8), 2907–2918.

Payne, J. D., Stickgold, R., Swanberg, K., & Kensinger, E. A. (2008). Sleep preferentially enhances memory for emotional components of scenes. *Psychological Science*, *19*(8), 781–788.

Payne, J. D., Tucker, M. A., Ellenbogen, J. M., Wamsley, E. J., Walker, M. P., Schacter, D. L., & Stickgold, R. (2012). Memory for semantically related and unrelated declarative information: The benefit of sleep, the cost of wake. *PLoS ONE*, *7*(3), e33079.

Peeters, D. (2018). A standardized set of 3-D objects for virtual reality research and applications. *Behavior Research Methods*, *50*(3), 1047–1054.

Peigneux, P., Laureys, S., Delbeuck, X., & Maquet, P. (2001). Sleeping brain, learning brain. The role of sleep for memory systems. *Neuroreport*, *12*(18), A111–A124.

Peigneux, P., Laureys, S., Fuchs, S., Collette, F., Perrin, F., Reggers, J., Phillips, C., Degueldre, C., Del Fiore, G., Aerts, J., Luxen, A., & Maquet, P. (2004). Are Spatial Memories Strengthened in the Human Hippocampus during Slow Wave Sleep? *Neuron*, *44*(3), 535–545.

Phelps, E. A., & LeDoux, J. E. (2005). Contributions of the Amygdala to Emotion Processing: From Animal Models to Human Behavior. *Neuron*, *48*(2), 175–187.

Picard, L., Abram, M., Orriols, E., & Piolino, P. (2017). Virtual reality as an ecologically valid tool for assessing multifaceted episodic memory in children and adolescents. *International Journal of Behavioral Development*, *41*(2), 211–219.

Piolino, P., Desgranges, B., & Eustache, F. (2009). Episodic autobiographical memories over the course of time: Cognitive, neuropsychological and neuroimaging findings. *Neuropsychologia*, *47*(11), 2314–2329.

Pitel, A. L., Beaunieux, H., Witkowski, T., Vabret, F., Guillery-Girard, B., Quinette, P., Desgranges, B., & Eustache, F. (2007). Genuine Episodic Memory Deficits and Executive Dysfunctions in Alcoholic Subjects Early in Abstinence. *Alcoholism: Clinical and Experimental Research*, *31*(7), 1169–1178.

Placencia, M., Sander, J., Shorvon, S., Ellison, R., & Cascante, S. (1992). Validation of a Screening Questionnaire for the Detection of Epileptic Seizures in Epidemiologic Studies. *Brain*, *115*, 783–794.

Plancher, G., Barra, J., Orriols, E., & Piolino, P. (2013). The influence of action on episodic memory: A virtual reality study. *Quarterly Journal of Experimental Psychology*, *66*(5), 895–909.

Plancher, G., Gyselinck, V., Nicolas, S., & Piolino, P. (2010). Age effect on components of episodic memory and feature binding: A virtual reality study. *Neuropsychology*, *24*(3), 379–390.

Plancher, G., Nicolas, S., & Piolino, P. (2008). Virtual reality as a tool for assessing episodic memory. *Proceedings of the 2008 ACM Symposium on Virtual Reality Software and Technology - VRST '08, Celda*, 179.

Plancher, G., Tirard, A., Gyselinck, V., Nicolas, S., & Piolino, P. (2012). Using virtual reality to characterize episodic memory profiles in amnestic mild cognitive impairment and Alzheimer's disease: Influence of active and passive encoding. *Neuropsychologia*, *50*(5), 592–602.

Plihal, W., & Born, J. (1997). Effects of Early and Late Nocturnal Sleep on Declarative and

Procedural Memory. *Journal of Cognitive Neuroscience*, *9*(4), 534–547.

Polcar, J., & Horejsi, P. (2015). Knowledge acquisition and cyber sickness: A comparison of VR devices in virtual tours. *MM Science Journal*, *2015*(02), 613–616.

Popic, D., Pacozzi, S. G., & Martarelli, C. S. (2020). Database of virtual objects to be used in psychological research. *PLOS ONE*, *15*(9), e0238041.

Poppenk, J., Evensmoen, H. R., Moscovitch, M., & Nadel, L. (2013). Long-axis specialization of the human hippocampus. In *Trends in Cognitive Sciences*.

Postma, A., Van Asselen, M., Keuper, O., Wester, A. J., & Kessels, R. P. C. (2006). Spatial and temporal order memory in Korsakoff patients. *Journal of the International Neuropsychological Society*, *12*(03).

R Core Team 2019. (2019). R: A language and environment for statistical computing. In *R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/*.

Rabin, J. S., Gilboa, A., Stuss, D. T., Mar, R. A., & Rosenbaum, R. S. (2010). Common and Unique Neural Correlates of Autobiographical Memory and Theory of Mind. *Journal of Cognitive Neuroscience*, *22*(6), 1095–1111.

Ragan, E. D. (2010). The Effects of Higher Levels of Immersion on Procedure Memorization Performance and Implications for Educational Virtual Environments. *Presence: Teleoperators and Virtual Environments*, *19*(6), 527–543.

Rajaram, S. (1993). Remembering and knowing: two means of access to the personal past. *Memory & Cognition*, *21*(1), 89–102.

Ranganath, C., Johnson, M. K., & D'Esposito, M. (2000). Left Anterior Prefrontal Activation Increases with Demands to Recall Specific Perceptual Information. *The Journal of Neuroscience*, *20*(22), RC108–RC108.

Rasch, B., & Born, J. (2007). Maintaining memories by reactivation. In *Current Opinion in Neurobiology* (Vol. 17, Issue 6, pp. 698–703).

Rasch, B., & Born, J. (2013). About Sleep's Role in Memory. *Physiological Reviews*, *93*(2), 681–766.

Ratcliff, R., & Murdock, B. B. (1976). Retrieval processes in recognition memory. *Psychological Review*, *83*(3), 190–214.

Ratcliff, R., Van Zandt, T., & McKoon, G. (1995). Process dissociation, single-process theories, and recognition memory. *Journal of Experimental Psychology: General*, *124*(4), 352–374.

Rauchs, G., Bertran, F., Guillery-Girard, B., Desgranges, B., Kerrouche, N., Denise, P., Foret, J., & Eustache, F. (2004). Consolidation of strictly episodic memories mainly requires rapid eye movement sleep. *Sleep*, *27*(3), 395–401.

Rauchs, G., Feyers, D., Landeau, B., Bastin, C., Luxen, A., Maquet, P., & Collette, F. (2011). Sleep Contributes to the Strengthening of Some Memories Over Others, Depending on Hippocampal Activity at Learning. *Journal of Neuroscience*, *31*(7), 2563–2568.

Rauchs, G., Orban, P., Balteau, E., Schmidt, C., Degueldre, C., Luxen, A., Maquet, P., & Peigneux, P. (2008). Partially segregated neural networks for spatial and contextual memory in virtual navigation. *Hippocampus*, *18*(5), 503–518.

Ready, R. E., Stierman, L., & Paulsen, J. S. (2001). Ecological Validity of Neuropsychological and Personality Measures of Executive Functions. *The Clinical Neuropsychologist*, *15*(3), 314–323.

Reagh, Z. M., & Yassa, M. A. (2014). Repetition strengthens target recognition but impairs similar lure discrimination: evidence for trace competition. *Learning & Memory*, *21*(7), 342–346.

Reggente, N., Essoe, J. K.-Y. K.-Y., Aghajan, Z. M., Tavakoli, A. V., McGuire, J. F., Suthana, N. A., & Rissman, J. (2018). Enhancing the Ecological Validity of fMRI Memory Research Using Virtual Reality. *Frontiers in Neuroscience*, *12*(JUN), 408.

Reggev, N., Sharoni, R., & Maril, A. (2018). Distinctiveness Benefits Novelty (and Not Familiarity), but Only Up to a Limit: The Prior Knowledge Perspective. *Cognitive Science*, *42*(1), 103–128.

Reid, L. M., & MacLullich, A. M. J. (2006). Subjective Memory Complaints and Cognitive Impairment in Older People. *Dementia and Geriatric Cognitive Disorders*, *22*(5–6), 471–485.

Renoult, L., & Rugg, M. D. (2020). An historical perspective on Endel Tulving's episodic-semantic distinction. *Neuropsychologia*, *139*(July 2019), 107366.

Repetto, C., Serino, S., Macedonia, M., & Riva, G. (2016). Virtual Reality as an Embodied

Tool to Enhance Episodic Memory in Elderly. *Frontiers in Psychology*, *7*(NOV), 1839.

Rich, J. B., & Troyer, A. K. (2008). Clinical neuropsychology. In C. A. Meyers & J. R. Perry (Eds.), *Cognition and Cancer* (pp. 6–18). Cambridge University Press.

Rimmele, U., Davachi, L., Petrov, R., Dougal, S., & Phelps, E. A. (2011). Emotion enhances the subjective feeling of remembering, despite lower accuracy for contextual details. *Emotion*, *11*(3), 553–562.

Ritchey, M., Dolcos, F., & Cabeza, R. (2008). Role of Amygdala Connectivity in the Persistence of Emotional Memories Over Time: An Event-Related fMRI Investigation. *Cerebral Cortex*, *18*(11), 2494–2504.

Roberts, W. A., Feeney, M. C., MacPherson, K., Petter, M., McMillan, N., & Musolino, E. (2008). Episodic-Like Memory in Rats: Is It Based on When or How Long Ago? *Science*, *320*(5872), 113–115.

Robin, J., Hirshhorn, M., Rosenbaum, R. S., Winocur, G., Moscovitch, M., & Grady, C. L. (2015). Functional connectivity of hippocampal and prefrontal networks during episodic and spatial memory based on real-world environments. *Hippocampus*, *25*(1), 81–93.

Robin, J., & Moscovitch, M. (2014). The effects of spatial contextual familiarity on remembered scenes, episodic memories, and imagined future events. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(2), 459–475.

Roediger, H. L., & Butler, A. C. (2011). The critical role of retrieval practice in long-term retention. *Trends in Cognitive Sciences*, *15*(1), 20–27.

Roediger, H. L., & McDermott, K. B. (2013). Two types of event memory. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(52), 20856–20857.

Roozendaal, B., McEwen, B. S., & Chattarji, S. (2009). Stress, memory and the amygdala. *Nature Reviews Neuroscience*, *10*(6), 423–433.

Rosenbaum, R. S., Gilboa, A., Levine, B., Winocur, G., & Moscovitch, M. (2009). Amnesia as an impairment of detail generation and binding: Evidence from personal, fictional, and semantic narratives in K.C. *Neuropsychologia*, *47*(11), 2181–2187.

Rosenbaum, R. S., Köhler, S., Schacter, D. L., Moscovitch, M., Westmacott, R., Black, S. E., Gao, F., & Tulving, E. (2005). The case of K.C.: contributions of a memory-impaired person to memory theory. *Neuropsychologia*, *43*(7), 989–1021.

Rosenbaum, R. S., Stuss, D. T., Levine, B., & Tulving, E. (2007). Theory of Mind Is Independent of Episodic Memory. *Science*, *318*(5854), 1257–1257.

Rotello, C. M., & Macmillan, N. A. (2006). Remember–know models as decision strategies in two experimental paradigms. *Journal of Memory and Language*, *55*(4), 479–494.

Roumis, D. K., & Frank, L. M. (2015). Hippocampal sharp-wave ripples in waking and sleeping states. In *Current Opinion in Neurobiology* (Vol. 35, pp. 6–12).

Rubin, D. C., Schrauf, R. W., & Greenberg, D. L. (2003). Belief and recollection of autobiographical memories. *Memory & Cognition*, *31*(6), 887–901.

Rubin, D. C., & Umanath, S. (2015). Event memory: A Theory of Memory for Laboratory, Autobiographical, and Fictional Events. *Psychological Review*, *122*(1), 1–23.

Ruddle, R. A., Payne, S. J., & Jones, D. M. (1997). Navigating buildings in “desk-top” virtual environments: Experimental investigations using extended navigational experience. *Journal of Experimental Psychology: Applied*, *3*(2), 143–159.

Ruddle, R. A., Payne, S. J., & Jones, D. M. (1999). Navigating Large-Scale Virtual Environments: What Differences Occur Between Helmet-Mounted and Desk-Top Displays? *Presence: Teleoperators and Virtual Environments*, *8*(2), 157–168.

Ruddle, R. A., Volkova, E., & Bülthoff, H. H. (2011). Walking improves your cognitive map in environments that are large-scale and large in extent. *ACM Transactions on Computer-Human Interaction*, *18*(2), 1–20.

Rugg, M. D., Johnson, J. D., & Uncapher, M. R. (2015). Encoding and Retrieval in Episodic Memory. In *The Wiley Handbook on the Cognitive Neuroscience of Memory* (pp. 84–107). John Wiley & Sons, Ltd.

Rugg, M. D., & Vilberg, K. L. (2013). Brain networks underlying episodic memory retrieval. *Current Opinion in Neurobiology*, *23*(2), 255–260.

Rugg, M. D., & Yonelinas, A. P. (2003). Human recognition memory: A cognitive neuroscience perspective. *Trends in Cognitive Sciences*, *7*(7), 313–319.

Russ, M. O., Mack, W., Grama, C.-R., Lanfermann, H., & Knopf, M. (2003). Enactment effect in memory: evidence concerning the function of the supramarginal gyrus. *Experimental Brain Research*, *149*(4), 497–504.

Sadeh, T., Ozubko, J. D., Winocur, G., & Moscovitch, M. (2014). How we forget may depend on how we remember. *Trends in Cognitive Sciences*, *18*(1), 26–36.

Saito, T., & Sadoshima, J. (2016). Evaluation of a Consumer Fitness- Tracking Device to Asses Sleep in Adults. *Chronobiol. Int.*, *116*(8), 1477–1490.

Saive, A. L., Royet, J. P., Garcia, S., Thévenet, M., & Plailly, J. (2015). "What-where-which" episodic retrieval requires conscious recollection and is promoted by semantic knowledge. *PLoS ONE*, *10*(12).

Sandrini, M., Cappa, S. F., Rossi, S., Rossini, P. M., & Miniussi, C. (2003). The Role of Prefrontal Cortex in Verbal Episodic Memory: rTMS Evidence. *Journal of Cognitive Neuroscience*, *15*(6), 855–861.

Saper, C. B. (2013). The Neurobiology of Sleep. *CONTINUUM: Lifelong Learning in Neurology*, *19*(1), 19–31.

Sauzéon, H., Arvind Pala, P., Larrue, F., Wallet, G., Déjos, M., Zheng, X., Guitton, P., & N'Kaoua, B. (2012). The Use of Virtual Reality for Episodic Memory Assessment. *Experimental Psychology*, *59*(2), 99–108.

Sbordone, R. J. (2008). Ecological validity of neuropsychological testing: Critical issues. *The Neuropsychology Handbook (3rd Ed.). BT - The Neuropsychology Handbook (3rd Ed.).*

Sbordone, R. J., & Guilmette, T. J. (1999). Ecological validity: Prediction of everyday and vocational functioning from neuropsychological test data. In *Forensic neuropsychology: Fundamentals and practice*.

Schacter, D. L., & Addis, D. R. (2007). The cognitive neuroscience of constructive memory: remembering the past and imagining the future. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *362*(1481), 773–786.

Schacter, D. L., Addis, D. R., & Buckner, R. L. (2008). Episodic Simulation of Future Events. *Annals of the New York Academy of Sciences*, *1124*(1), 39–60.

Schacter, D. L., Addis, D. R., Hassabis, D., Martin, V. C., Spreng, R. N., & Szpunar, K. K. (2012). The Future of Memory: Remembering, Imagining, and the Brain. *Neuron*, *76*(4), 677–694.

Schacter, D. L., Harbluk, J. L., & McLachlan, D. R. (1984). Retrieval without recollection: An experimental analysis of source amnesia. *Journal of Verbal Learning and Verbal Behavior*, *23*(5), 593–611.

Schacter, D. L., & Madore, K. P. (2016). Remembering the past and imagining the future: Identifying and enhancing the contribution of episodic memory. *Memory Studies*, *9*(3), 245–255.

Schapiro, A. C., Turk-Browne, N. B., Botvinick, M. M., & Norman, K. A. (2017). Complementary learning systems within the hippocampus: a neural network modelling approach to reconciling episodic memory with statistical learning. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *372*(1711), 20160049.

Schmidt, M. H., Swang, T. W., Hamilton, I. M., & Best, J. A. (2017). State-dependent metabolic partitioning and energy conservation: A theoretical framework for understanding the function of sleep. *PLOS ONE*, *12*(10), e0185746.

Schmidt, S. (2004). Autobiographical memories for the September 11th attacks: Reconstructive errors and emotional impairment of memory. *Memory & Cognition*, *32*(3), 443–454.

Schmidt, S. R. (1991). Can we have a distinctive theory of memory? *Memory & Cognition*, *19*(6), 523–542.

Schomaker, J., Roos, R., & Meeter, M. (2014). Expecting the unexpected: The effects of deviance on novelty processing. *Behavioral Neuroscience*, *128*(2), 146–160.

Schöne, B., Wessels, M., & Gruber, T. (2019). Experiences in Virtual Reality: a Window to Autobiographical Memory. *Current Psychology*, *38*(3), 715–719.

Schubert, T. (2002). Real and Illusory Interactions Enhance Presence. *Presence: Teleoperators and Virtual Environments*, *11*(4), 425–434.

Schubert, T., Friedmann, F., & Regenbrecht, H. (2001). The Experience of Presence: Factor Analytic Insights. *Presence: Teleoperators and Virtual Environments*, *10*(3), 266–281.

Schwabe, L., Nader, K., & Pruessner, J. C. (2014). Reconsolidation of human memory: Brain mechanisms and clinical relevance. In *Biological Psychiatry* (Vol. 76, Issue 4, pp. 274–

280). Elsevier.

Scullin, M. K. (2014). Sleep, Memory, and Aging: The Link Between Slow-Wave Sleep and Episodic Memory Changes from Younger to Older Adults. *Psychololgy and Aging*, *28*(1), 105–114.

Seibert, J., & Shafer, D. M. (2018). Control mapping in virtual reality: effects on spatial presence and controller naturalness. *Virtual Reality*, *22*(1), 79–88.

Sekeres, M. J., Bonasia, K., St-Laurent, M., Pishdadian, S., Winocur, G., Grady, C., & Moscovitch, M. (2016). Recovering and preventing loss of detailed memory: differential rates of forgetting for detail types in episodic memory. *Learning & Memory*, *23*(2), 72–82.

Selzer, M. N., Gazcon, N. F., & Larrea, M. L. (2019). Effects of virtual presence and learning outcome using low-end virtual reality systems. *Displays*, *59*, 9–15.

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

Sharot, T., & Yonelinas, A. P. (2008). Differential time-dependent effects of emotion on recollective experience and memory for contextual information. *Cognition*, *106*(1), 538–547.

Sheldon, S., & Levine, B. (2013). Same as it ever was: Vividness modulates the similarities and differences between the neural networks that support retrieving remote and recent autobiographical memories. *NeuroImage*, *83*, 880–891.

Sheth, B. R., Nguyen, N. D., & Janvelyan, D. (2009). Does Sleep Really Influence Face Recognition Memory? *PLoS ONE*, *4*(5), e5496.

Shimamura, A. P., & Squire, L. R. (1987). A neuropsychological study of fact memory and source amnesia. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *13*(3), 464–473.

Shimamura, A. P., & Wickens, T. D. (2009). Superadditive memory strength for item and source recognition: The role of hierarchical relational binding in the medial temporal lobe. *Psychological Review*, *116*(1), 1–19.

Shin, M. S., Park, S. Y., Park, S. R., Seol, S. H., & Kwon, J. S. (2006). Clinical and empirical applications of the Rey-Osterrieth Complex Figure Test. *Nature Protocols*, *1*(2), 892–899.

Shu, Y., Huang, Y. Z., Chang, S. H., & Chen, M. Y. (2019). Do virtual reality head-mounted displays make a difference? A comparison of presence and self-efficacy between head-mounted displays and desktop computer-facilitated virtual environments. *Virtual Reality*, *23*(4), 437–446.

Siegel, J. M. (2001). The REM sleep-memory consolidation hypothesis. *Science (New York, N.Y.)*, *294*(5544), 1058–1063.

Silver, C. H. (2000). Ecological Validity of Neuropsychological Assessment in Childhood Traumatic Brain Injury. *Journal of Head Trauma Rehabilitation*, *15*(4), 973–988.

Simons, J. S. (2002). Recollection-based memory in frontotemporal dementia: implications for theories of long-term memory. *Brain*, *125*(11), 2523–2536.

Simons, J. S., & Spiers, H. J. (2003). Prefrontal and medial temporal lobe interactions in long-term memory. *Nature Reviews Neuroscience*, *4*(8), 637–648.

Simpson, P. M., Surmon, D. J., Wesnes, K. A., & Wilcock, G. K. (1991). The cognitive drug research computerized assessment system for demented patients: A validation study. *International Journal of Geriatric Psychiatry*, *6*(2), 95–102.

Skaggs, W. E., & McNaughton, B. L. (1996). Replay of neuronal firing sequences in rat hippocampus during sleep following spatial experience. *Science (New York, N.Y.)*, *271*(5257), 1870–1873.

Skinner, E. I., & Fernandes, M. A. (2007). Neural correlates of recollection and familiarity: A review of neuroimaging and patient data. *Neuropsychologia*, *45*(10), 2163–2179.

Slater, M. (2009). Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1535), 3549–3557.

Slotnick, S. D., & Dodson, C. S. (2005). Support for a continuous (single-process) model of recognition memory and source memory. *Memory & Cognition*, *33*(1), 151–170.

Smith, S. A. (2019). Virtual reality in episodic memory research: A review. *Psychonomic Bulletin & Review*, *26*(4), 1213–1237.

Smulders, T. V., Black-Dominique, A., Choudhury, T. S., Constantinescu, S. E., Foka, K., Walker, T. J., Dick, K., Bradwel, S., McAllister-Williams, R. H., & Gallagher, P. (2017a). A Real-world What-Where-When Memory Test. *Journal of Visualized Experiments*, *123*, 1–8.

Smulders, T. V., Black-Dominique, A., Choudhury, T. S., Constantinescu, S. E., Foka, K., Walker, T. J., Dick, K., Bradwel, S., McAllister-Williams, R. H., & Gallagher, P. (2017b). A Real-world What-Where-When Memory Test. *Journal of Visualized Experiments*, *123*, 1–8.

Soderstrom, N. C., Kerr, T. K., & Bjork, R. A. (2016). The Critical Importance of Retrieval—and Spacing—for Learning. *Psychological Science*, *27*(2), 223–230.

Solomon, P. R., Hirschoff, A., Kelly, B., Relin, M., Brush, M., DeVeaux, R. D., & Pendlebury, W. W. (1998). A 7 Minute Neurocognitive Screening Battery Highly Sensitive to Alzheimer's Disease. *Archives of Neurology*, *55*(3), 349.

Sousa, N., Cammarota, M., Cheng, S., & Numan, R. (2015). A Prefrontal-Hippocampal Comparator for Goal-Directed Behavior: The Intentional Self and Episodic Memory. *Frontiers in Behavioral Neuroscience*, *9*, 323.

Sousa Santos, B., Dias, P., Pimentel, A., Baggerman, J.-W. W., Ferreira, C., Silva, S., & Madeira, J. (2009). Head-mounted display versus desktop for 3D navigation in virtual reality: a user study. *Multimedia Tools and Applications*, *41*(1), 161–181.

Sparks, F. T., Lehmann, H., Hernandez, K., & Sutherland, R. J. (2011). Suppression of Neurotoxic Lesion-Induced Seizure Activity: Evidence for a Permanent Role for the Hippocampus in Contextual Memory. *PLoS ONE*, *6*(11), e27426.

Spiers, H. J., Burgess, N., Hartley, T., Vargha-Khadem, F., & O'Keefe, J. (2001). Bilateral hippocampal pathology impairs topographical and episodic memory but not visual pattern matching. *Hippocampus*, *11*(6), 715–725.

Spiers, H. J., Burgess, N., Maguire, E. A., Baxendale, S. A., Hartley, T., Thompson, P. J., & O'Keefe, J. (2001). Unilateral temporal lobectomy patients show lateralized topographical and episodic memory deficits in a virtual town. *Brain : A Journal of Neurology*, *124*(Pt 12), 2476–2489.

Spinnler, H., & Tognoni, G. (1987). Italian standardization and classification of

Neuropsychological tests. The Italian Group on the Neuropsychological Study of Aging. *Italian Journal of Neurological Sciences*, *Suppl 8*, 1–120.

Spreng, R. N., Mar, R. A., & Kim, A. S. N. (2009). The Common Neural Basis of Autobiographical Memory, Prospection, Navigation, Theory of Mind, and the Default Mode: A Quantitative Meta-analysis. *Journal of Cognitive Neuroscience*, *21*(3), 489–510.

Squire, L. R. (2004). Memory systems of the brain: A brief history and current perspective. *Neurobiology of Learning and Memory*, *82*(3), 171–177.

Squire, L. R., & Alvarez, P. (1995). Retrograde amnesia and memory consolidation: a neurobiological perspective. *Current Opinion in Neurobiology*, *5*(2), 169–177.

Squire, L. R., & Wixted, J. T. (2011). The Cognitive Neuroscience of Human Memory Since H.M. *Annual Review of Neuroscience*, *34*(1), 259–288.

Srivastava, P., Rimzhim, A., Vijay, P., Singh, S., & Chandra, S. (2019). Desktop VR Is Better Than Non-ambulatory HMD VR for Spatial Learning. *Frontiers in Robotics and AI*, *6*.

St-Laurent, M., Moscovitch, M., Jadd, R., & McAndrews, M. P. (2014). The perceptual richness of complex memory episodes is compromised by medial temporal lobe damage. *Hippocampus*, *24*(5), 560–576.

St-Laurent, M., Moscovitch, M., & McAndrews, M. P. (2016). The retrieval of perceptual memory details depends on right hippocampal integrity and activation. *Cortex*, *84*, 15–33.

Stamenova, V., Gao, F., Black, S. E., Schwartz, M. L., Kovacevic, N., Alexander, M. P., & Levine, B. (2017). The effect of focal cortical frontal and posterior lesions on recollection and familiarity in recognition memory. *Cortex*, *91*, 316–326.

Stevens, J., Kincaid, P., & Sottilare, R. (2015). Visual modality research in virtual and mixed reality simulation. *The Journal of Defense Modeling and Simulation: Applications, Methodology, Technology*, *12*(4), 519–537.

Stickgold, R. (2005). Sleep-dependent memory consolidation. *Nature*, *437*(7063), 1272–1278.

Stickgold, R., James, L., & Hobson, J. A. (2000). Visual discrimination learning requires sleep after training. *Nature Neuroscience*, *3*(12), 1237–1238.

Stickgold, R., & Walker, M. P. (2005). Memory consolidation and reconsolidation: What is the role of sleep? *Trends in Neurosciences*, *28*(8), 408–415.

Stickgold, R., & Walker, M. P. (2007). Sleep-dependent memory consolidation and reconsolidation. *Sleep Medicine*, *8*(4), 331–343.

Stickgold, R., Whidbee, D., Schirmer, B., Patel, V., & Hobson, J. A. (2000). Visual Discrimination Task Improvement: A Multi-Step Process Occurring During Sleep. *Journal of Cognitive Neuroscience*, *12*(2), 246–254.

Strange, B. A., Witter, M. P., Lein, E. S., & Moser, E. I. (2014). Functional organization of the hippocampal longitudinal axis. *Nature Reviews Neuroscience*, *15*(10), 655–669.

Strikwerda- Brown, C., Mothakunnel, A., Hodges, J. R., Piguet, O., & Irish, M. (2019). External details revisited – A new taxonomy for coding 'non- episodic' content during autobiographical memory retrieval. *Journal of Neuropsychology*, *13*(3), 371–397.

Studte, S., Bridger, E., & Mecklinger, A. (2015). Nap sleep preserves associative but not item memory performance. *Neurobiology of Learning and Memory*, *120*, 84–93.

Suddendorf, T., & Corballis, M. C. (2007). The evolution of foresight: What is mental time travel, and is it unique to humans? *The Behavioral and Brain Sciences*, *30*(3), 299–313; discussion 313-51.

Sutherland, R. J., Lee, J. Q., McDonald, R. J., & Lehmann, H. (2019). Has multiple trace theory been refuted? *Hippocampus*, hipo.23162.

Svoboda, E., McKinnon, M. C., & Levine, B. (2006). The functional neuroanatomy of autobiographical memory: A meta-analysis. In *Neuropsychologia*.

Swets, J. A., Tanner, W. P., & Birdsall, T. G. (1961). Decision processes in perception. *Psychological Review*, *68*(5), 301–340.

Szpunar, K. K. (2011). On subjective time. *Cortex*, *47*(3), 409–411.

Szpunar, K. K., & Tulving, E. (2011). Varieties of future experience. In M. Bar (Ed.), *Predictions and the Brain: Using Our Past to Generate a Future* (pp. 3–12). Oxford University Press.

Takashima, A., Petersson, K. M., Rutters, F., Tendolkar, I., Jensen, O., Zwarts, M. J., McNaughton, B. L., & Fernández, G. (2006). Declarative memory consolidation in humans: a prospective functional magnetic resonance imaging study. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(3), 756–761.

Talamini, L. M., & Gorree, E. (2012). Aging memories: Differential decay of episodic memory components. *Learning and Memory*, 239–246.

Talamini, L. M., Nieuwenhuis, I. L. C., Takashima, A., & Jensen, O. (2008). Sleep directly following learning benefits consolidation of spatial associative memory. *Learning & Memory*, *15*(4), 233–237.

Talarico, J. M., & Rubin, D. C. (2003). Confidence, Not Consistency, Characterizes Flashbulb Memories. *Psychological Science*, *14*(5), 455–461.

Tanaka, Y., Miyazawa, Y., Hashimoto, R., Nakano, I., & Obayashi, T. (1999). Postencephalitic focal retrograde amnesia after bilateral anterior temporal lobe damage. *Neurology*, *53*(2), 344–344.

Teyler, T. J., & Rudy, J. W. (2007). The hippocampal indexing theory and episodic memory: Updating the index. *Hippocampus*, *17*(12), 1158–1169.

Thapa, R., Sparks, F. T., Hanif, W., Gulbrandsen, T., & Sutherland, R. J. (2014). Recent memory for socially transmitted food preferences in rats does not depend on the hippocampus. *Neurobiology of Learning and Memory*, *114*, 113–116.

Thompson, C. P. (1982). Memory for unique personal events: The roommate study. *Memory & Cognition*, *10*(4), 324–332.

Thorndyke, P. W. (1977). Cognitive structures in comprehension and memory of narrative discourse. *Cognitive Psychology*, *9*(1), 77–110.

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, *12*(1), 97–136.

Trevarthen, C. (1977). Review: Cognition and Reality: Principles and Implications of Cognitive Psychology. *Perception*, *6*(5), 605–610.

Trewartha, K. M., Case, S., & Flanagan, J. R. (2015). Integrating actions into object location memory: A benefit for active versus passive reaching movements. *Behavioural Brain*

Research, *279*, 234–239.

Tromp, J., Klotzsche, F., Krohn, S., Akbal, M., Pohl, L., Quinque, E. M., Belger, J., Villringer, A., & Gaebler, M. (2020). *OpenVirtualObjects (OVO): An open set of standardized and validated 3D household objects for virtual reality-based research, assessment, and therapy.*

Tuena, C., Serino, S., Dutriaux, L., Riva, G., & Piolino, P. (2019). Virtual Enactment Effect on Memory in Young and Aged Populations: A Systematic Review. *Journal of Clinical Medicine*, *8*(5), 620.

Tulving, E. (1967). The effects of presentation and recall of material in free-recall learning. *Journal of Verbal Learning and Verbal Behavior*, *6*(2), 175–184.

Tulving, E. (1972). Episodic and semantic memory. In *Organization of memory* (Vol. 1, pp. 381–403).

Tulving, E. (1983). Elements of Episodic Memory. *Canadian Psychology*, *26*(3), 351.

Tulving, E. (1985). Memory and consciousness. *Canadian Psychology/Psychologie Canadienne*, *26*(1), 1–12.

Tulving, E. (1993). What Is Episodic Memory? *Current Directions in Psychological Science*, *2*(3), 67–70.

Tulving, E. (2001). Episodic memory and common sense: how far apart? *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *356*(1413), 1505–1515.

Tulving, E. (2002). EPISODIC MEMORY : From Mind to Brain. *Annual Review of Psychology*, *53*, 1–25.

Tulving, E. (2004). Origin of autonoesis in episodic memory. In *The nature of remembering: Essays in honor of Robert G. Crowder.* (pp. 17–34). American Psychological Association.

Tulving, E., Kapur, S., Craik, F. I. M., Moscovitch, M., & Houle, S. (1994). Hemispheric encoding/retrieval asymmetry in episodic memory: Positron emission tomography findings. *Proceedings of the National Academy of Sciences of the United States of America*.

Tulving, E., & Markowitsch, H. J. (1998). Episodic and declarative memory: role of the hippocampus. *Hippocampus*, *8*(3), 198–204.

Tulving, E., Stuss, D. T., & Tulving, E. (1995). Frontal Lobe Damage Produces Episodic Memory Impairment. *Journal of the International Neuropsychological Society*.

Tunney, R. J. (2010). Do changes in the subjective experience of recognition over time suggest independent processes? *British Journal of Mathematical and Statistical Psychology*.

Tunney, R. J., & Bezzina, G. (2007). Effects of retention intervals on receiver operating characteristics in artificial grammar learning. *Acta Psychologica*, *125*(1), 37–50.

Uncapher, M. R., Otten, L. J., & Rugg, M. D. (2006). Episodic Encoding Is More than the Sum of Its Parts: An fMRI Investigation of Multifeatural Contextual Encoding. *Neuron*, *52*(3), 547–556.

Unity. (2017). *Unity - Game Engine*. Unity3d.Com.

Usoh, M., Arthur, K., Whitton, M. C., Bastos, R., Steed, A., Slater, M., & Brooks, F. P. (1999). Walking > walking-in-place > flying, in virtual environments. *Proceedings of the 26th Annual Conference on Computer Graphics and Interactive Techniques - SIGGRAPH '99*, 359–364.

Usoh, M., Catena, E., Arman, S., & Slater, M. (2000). Using Presence Questionnaires in Reality. *Presence: Teleoperators and Virtual Environments*, *9*(5), 497–503.

van der Ham, I. J. M., Faber, A. M. E., Venselaar, M., van Kreveld, M. J., & Löffler, M. (2015). Ecological validity of virtual environments to assess human navigation ability. *Frontiers in Psychology*, *6*, 637.

van der Helm, E., Gujar, N., Nishida, M., Walker, M. P., Helm, E. Van Der, Gujar, N., Nishida, M., & Walker, M. P. (2011). Sleep-Dependent Facilitation of Episodic Memory Details. *PLoS ONE*, *6*(11).

Van Kesteren, M. T. R. R., Ruiter, D. J., Fernández, G., & Henson, R. N. (2012). How schema and novelty augment memory formation. *Trends in Neurosciences*, *35*(4), 211–219.

Viard, A., Desgranges, B., Eustache, F., & Piolino, P. (2012). Factors affecting medial temporal lobe engagement for past and future episodic events: An ALE meta-analysis of neuroimaging studies. *Brain and Cognition*, *80*(1), 111–125.

Wagner, A. D., Schacter, D. L., Rotte, M., Koutstaal, W., Maril, A., Dale, A. M., Rosen, B. R., & Buckner, R. L. (1998). Building memories: Remembering and forgetting of verbal experiences as predicted by brain activity. *Science*.

Wais, P. E., Mickes, L., & Wixted, J. T. (2008). Remember/Know Judgments Probe Degrees of Recollection. *Journal of Cognitive Neuroscience*, *20*(3), 400–405.

Walker, M. P. (2009). The role of sleep in cognition and emotion. *Annals of the New York Academy of Sciences*, *197*, 168–197.

Waller, D. (2000). Individual differences in spatial learning from computer-simulated environments. *Journal of Experimental Psychology: Applied*, *6*(4), 307–321.

Waller, D., Hunt, E., & Knapp, D. (1998). The transfer of spatial knowledge in virtual environment training. *Presence: Teleoperators and Virtual Environments*, *7*(2), 129–143.

Wallet, G., Sauzéon, H., Pala, P. A., Larrue, F., Zheng, X., & N'Kaoua, B. (2011). Virtual/real transfer of spatial knowledge: benefit from visual fidelity provided in a virtual environment and impact of active navigation. *Cyberpsychology, Behavior and Social Networking*, *14*(7), 417–423.

Wang, J.-H. (2019). Searching basic units in memory traces: associative memory cells. *F1000Research*, *8*, 457.

Wang, J.-Y., Weber, F. D., Zinke, K., Inostroza, M., & Born, J. (2018). More Effective Consolidation of Episodic Long-Term Memory in Children Than Adults-Unrelated to Sleep. *Child Development*, *89*(5), 1720–1734.

Wang, S.-H., & Morris, R. G. M. (2010). Hippocampal-Neocortical Interactions in Memory Formation, Consolidation, and Reconsolidation. *Annual Review of Psychology*, *61*(1), 49–79.

Wechsler, D. (2008). Wechsler adult intelligence scale–Fourth Edition (WAIS–IV). *San Antonio, TX:*, NCS Pearson..

Weech, S., Kenny, S., & Barnett-Cowan, M. (2019). Presence and Cybersickness in Virtual

Reality Are Negatively Related: A Review. *Frontiers in Psychology*, *10*.

Weidemann, C. T., & Kahana, M. J. (2016). Assessing recognition memory using confidence ratings and response times. *Royal Society Open Science*, *3*(4), 150670.

Weintraub, S., Peavy, G. M., O'Connor, M., Johnson, N. A., Acar, D., Sweeney, J., & Janssen, I. (2000). Three Words - Three Shapes: A clinical test of memory. *Journal of Clinical and Experimental Neuropsychology*.

Weller, R., & Zachmann, G. (2012). User performance in complex bi-manual haptic manipulation with 3 DOFs vs. 6 DOFs. *Haptics Symposium 2012, HAPTICS 2012 - Proceedings*.

Welsh, K. A., Butters, N., Hughes, J. P., Mohs, R. C., & Heyman, A. (1992). Detection and Staging of Dementia in Alzheimer's Disease: Use of the Neuropsychological Measures Developed for the Consortium to Establish a Registry for Alzheimer's Disease. *Archives of Neurology*.

Wesley Regian, J., & Yadrick, R. M. (1994). Assessment of configurational knowledge of naturally- and artificially-acquired large-scale space. *Journal of Environmental Psychology*, *14*(3), 211–223.

Wheeler, M. A., Stuss, D. T., & Tulving, E. (1997). Toward a theory of episodic memory: the frontal lobes and autonoetic consciousness. *Psychological Bulletin*, *121*(3), 331–354.

Whitcombe, E. ., & Robinson, E. . (2000). Children's decisions about what to believe and their ability to report the source of their belief. *Cognitive Development*, *15*(3), 329–346.

Wichert, S., Wolf, O. T., & Schwabe, L. (2013). Updating of episodic memories depends on the strength of new learning after memory reactivation. *Behavioral Neuroscience, 127*(3), 331–338.

Widmann, C. N., Beinhoff, U., & Riepe, M. W. (2012). Everyday memory deficits in very mild Alzheimer's disease. *Neurobiology of Aging*, *33*(2), 297–303.

Wiesner, C. D., Pulst, J., Krause, F., Elsner, M., Baving, L., Pedersen, A., Prehn-Kristensen, A., & Göder, R. (2015). The effect of selective REM-sleep deprivation on the consolidation and affective evaluation of emotional memories. *Neurobiology of Learning and Memory*, *122*, 131–141.

Wilhelm, I., Diekelmann, S., Molzow, I., Ayoub, A., Molle, M., Born, J., Mölle, M., & Born, J. (2011). Sleep Selectively Enhances Memory Expected to Be of Future Relevance. *Journal of Neuroscience*, *31*(5), 1563–1569.

Wilson, B. A. (2013). The Assessment, Evaluation and Rehabilitation of Everyday Memory Problems. In *The Assessment, Evaluation and Rehabilitation of Everyday Memory Problems: Selected Papers of Barbara A. Wilson*. Psychology Press.

Wilson, B. A., Cockburn, J., & Baddeley, A. D. (1985). The Rivermead Behavioural Memory Test. In *Thames Valley Company*.

Wilson, C. J., & Soranzo, A. (2015). The Use of Virtual Reality in Psychology: A Case Study in Visual Perception. *Computational and Mathematical Methods in Medicine*, *2015*, 1–7.

Wilson, R. S., Bacon, L. D., Kaszniak, A. W., & Fox, J. H. (1982). The episodic–semantic memory distinction and paired associate learning. *Journal of Consulting and Clinical Psychology*, *50*(1), 154–155.

Winocur, G., & Moscovitch, M. (2011). Memory transformation and systems consolidation. In *Journal of the International Neuropsychological Society* (Vol. 17, Issue 5, pp. 766–780).

Winocur, G., Moscovitch, M., & Bontempi, B. (2010). Memory formation and long-term retention in humans and animals: Convergence towards a transformation account of hippocampal-neocortical interactions. *Neuropsychologia*, *48*(8), 2339–2356.

Winters, B. D., Tucci, M. C., & DaCosta-Furtado, M. (2009). Older and stronger object memories are selectively destabilized by reactivation in the presence of new information. *Learning & Memory (Cold Spring Harbor, N.Y.)*, *16*(9), 545–553.

Witmer, B. G., Bailey, J. H., Knerr, B. W., & Parsons, K. C. (1996). Virtual spaces and real world places: transfer of route knowledge. *International Journal of Human-Computer Studies*, *45*(4), 413–428.

Witmer, B. G., & Singer, M. J. (1998). Measuring Presence in Virtual Environments : A Presence. *Presence*, *7*(3), 225–240.

Wixted, J. T. (2007). Dual-process theory and signal-detection theory of recognition memory. *Psychological Review*, *114*(1), 152–176.

Wixted, J. T. (2009). Remember/Know judgments in cognitive neuroscience: An illustration of the underrepresented point of view. In *Learning and Memory*.

Wixted, J. T., & Mickes, L. (2010). A continuous dual-process model of remember/know judgments. *Psychological Review*, *117*(4), 1025–1054.

Xu, M., & Bellezza, F. S. (2001). A comparison of the multimemory and detection theories of know and remember recognition judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*(5), 1197–1210.

Yassa, M. A., & Reagh, Z. M. (2013). Competitive Trace Theory: A Role for the Hippocampus in Contextual Interference during Retrieval. *Frontiers in Behavioral Neuroscience*, *7*(August), 107.

Yonelinas, A. P. (1994). Receiver-operating characteristics in recognition memory: evidence for a dual-process model. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*(6), 1341–1354.

Yonelinas, A. P. (1997). Recognition memory ROCs for item and associative information: The contribution of recollection and familiarity. *Memory and Cognition*, *25*(6), 747–763.

Yonelinas, A. P. (2001). Components of episodic memory: the contribution of recollection and familiarity. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *356*(1413), 1363–1374.

Yonelinas, A. P. (2002). The Nature of Recollection and Familiarity: A Review of 30 Years of Research. *Journal of Memory and Language*, *46*(3), 441–517.

Yonelinas, A. P., & Levy, B. J. (2002). Dissociating familiarity from recollection in human recognition memory: Different rates of forgetting over short retention intervals. *Psychonomic Bulletin & Review*, *9*(3), 575–582.

Yonelinas, A. P., Ranganath, C., Ekstrom, A. D., & Wiltgen, B. J. (2019). A contextual binding theory of episodic memory: systems consolidation reconsidered. *Nature Reviews Neuroscience*, *20*(6), 364–375.

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.

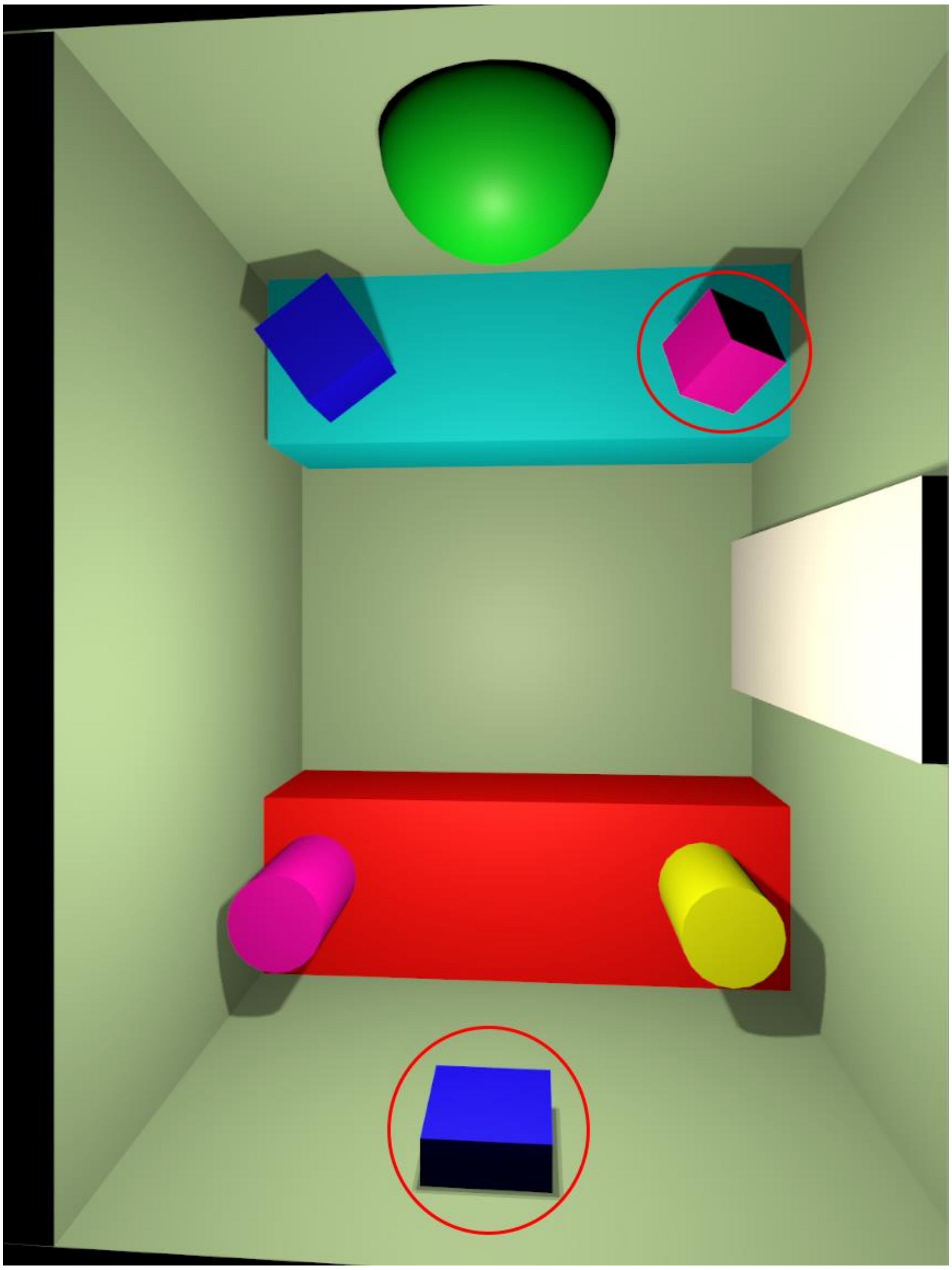
Zimmer, H. D., & Engelkamp, J. (1989). Does motor encoding enhance relational information? *Psychological Research*, *51*(4), 158–167.

Zlomuzica, A., Preusser, F., Totzeck, C., Dere, E., & Margraf, J. (2016). The impact of different emotional states on the memory for what, where and when features of specific events. *Behavioural Brain Research*, *298*, 181–187.

Appendix A: Virtual Environments for Experiments 1, 2 and 3

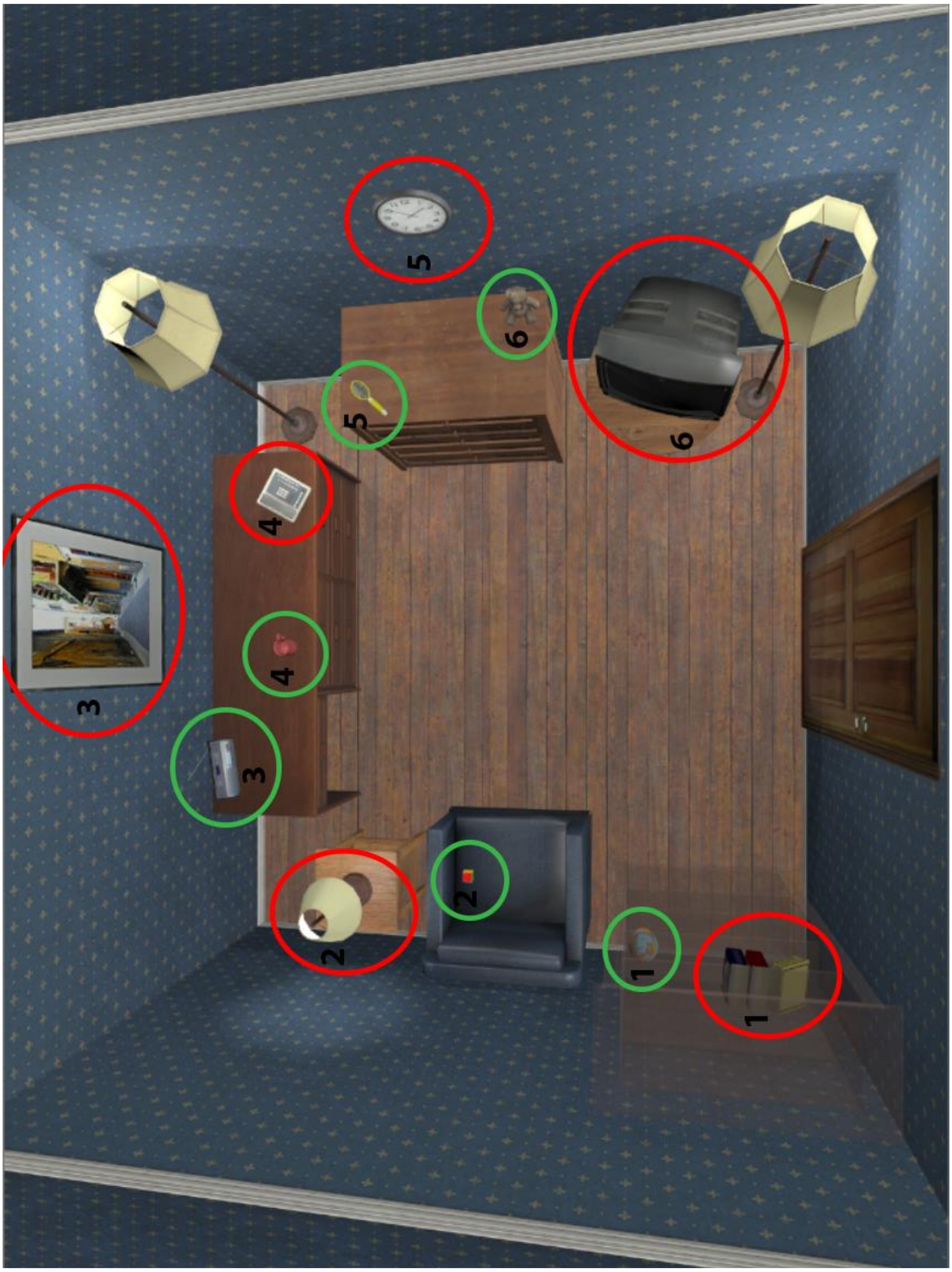
Objects circled in red are the Event objects. Objects circled in green are the non-Event objects. For a full list of objects and their events see Appendix II.

The Training room



Notes: The two event objects in this room are the pink cube and the blue square on the wall. The pink cube spins around its axis while making a creaking sound while the blue square falls down off the wall.

The Bedroom



Notes: The bookshelf in the bottom right corner has been made transparent in this image to better show the objects on it.

The Study

The Kitchen

Notes: The wall units and the extractor fan have been made transparent for this image to better show the objects under them.

Appendix B: Event and non-Event objects in Experiments 1, 2 and 3

Event objects with descriptions of their events

Bedroom

1. Books

The books fall on their sides making a thump sound.

2. Table lamp

The lightbulb inside the lamp pops while making a bulb shattering sound. The shade of the lamp turns darker and the lamp stops emitting light in the room.

3. Painting

The upper left corner of the painting gets loose and the whole painting slides to the right. This is accompanied by a thump sound.

4. Phone

The phone rings and a small red light on it flash for 3 seconds.

5. Wall clock

The clock falls of the wall onto the chest of drawers bellow it. This is accompanied with a crash sound.

6. TV

The TV comes on and plays static with the accompanying static sound. This lasts for 3 seconds after which the TV turns off.

Study

7. Photocopier

A blue light moves across the scanner part with an accompanying scanning sound. This lasts for 3 seconds.

8. PC

Windows ‘blue screen of death’ comes on the screen with an accompanying error sound. After staying of the screen for 3 seconds it disappears and a normal desktop is shown again.

9. Noticeboard

The board fall of the desk below it making a thump sound.

10. Office chair

The chair moves towards the corner of the room while producing the chair wheel noise.

11. Empty box

The box falls on its side with a thump sound.

12. PC case

The case slides down the wall and falls of the desk making a crashing sound.

Kitchen

13. Washing-up liquid

The bottle falls into the sink making plastic thump sound.

14. Cooking pot

The pot shakes as if it is boiling. This is accompanied with a sound of steam.

15. Fridge/Freezer

Both fridge and freezer doors open and close one after another. This is accompanied with appropriate sounds.

16. Chopping board

The board falls of its holder onto the worktop below it. The board hits the worktop with its thin side and then falls flat while making wooden thump sound.

17. Coffee maker

A red light comes on and the coffee maker plays a bubbling sound. This lasts for 3 seconds.

18. Microwave

The light inside the microwave comes on and it plays a microwave working sound. This is finished with a ping after which the light goes off. The whole event lasts 3 seconds.

Non-Event objects

Bedroom

1. Globe
2. Rubik's cube
3. Radio
4. Mug
5. Hair brush
6. Teddy

Study

7. Plant
8. Poster
9. File box
10. Desk lamp
11. Headphones
12. PC part

Kitchen

13. Plate
14. Water bottle
15. Pan
16. Apple
17. Toaster
18. Knife

Lure objects

- Bowl
- Ladle
- Dish drainer
- Sponge
- Dustpan
- Bottle of wine
- Screwdriver
- Clipboard
- Notepad
- Paper tray
- Battery
- Pencil
- Alarm clock
- Coat hanger
- Iron
- Glasses
- Camera
- Shoe
- Corkscrew
- Glass
- Clothes peg
- Fork
- Magnifying glass
- Pillow
- Ashtray
- Scissors
- Tin of fruit
- Teapot
- Key
- Fire extinguisher
- CD
- Hairdryer
- Ruler
- Marker pen
- Table mirror
- Sellotape

Appendix C: A free recall example from Experiment 1

Note: This is an example of a free recall given by one of the AM group participants in Experiment 1. The coloured parts of the text are the objects and their details which were counted and analysed.

Non-event object

Non-event object detail

Event Object

Event object detail

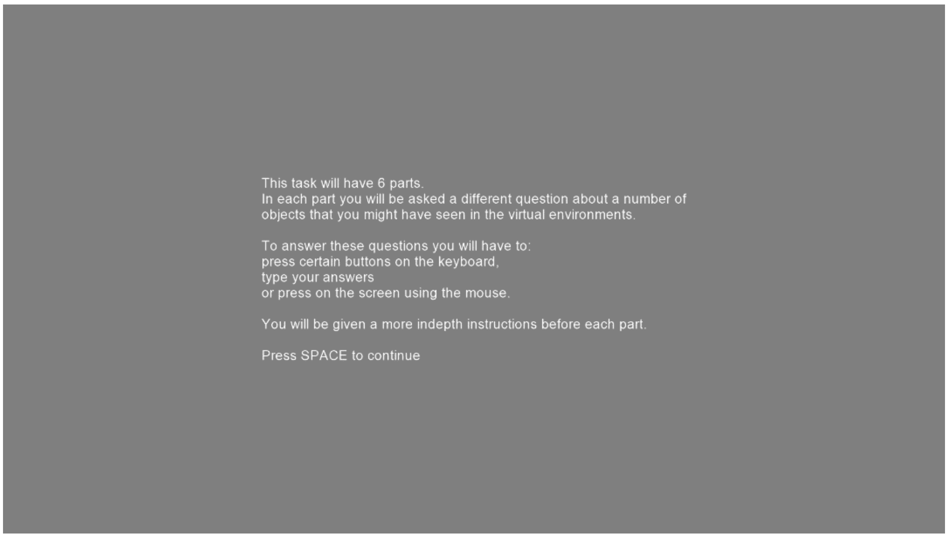
I went in a living room, there were 3 books, they were 3 different colours, one had stars on the binding, they were on shelves and they fell over. I saw a globe of the world on the shelves. There was a picture on the wall and the wallpaper was blue with stars. I saw a rubics cube on a chair, I could see the yellow and red sides of it. There was a television in the room. A clock on the wall fell down making a noise and the time was 1.50. The door opened when I was ready to leave. There were other noises and other things that moved and fell.

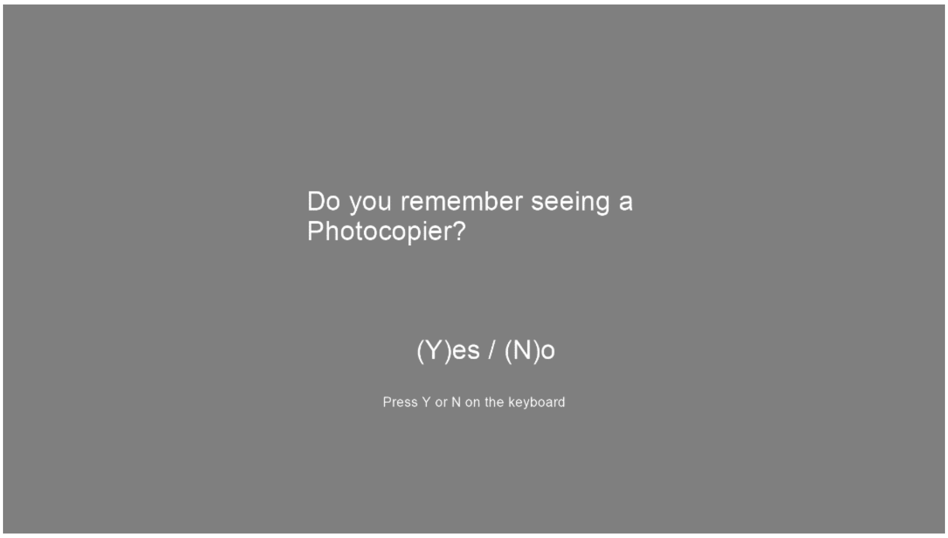
I went into an office, a photocopier made a noise. In the corner of the room there was a computer which switched on when I went near to it. There was a telephone that rang. There were filing cabinets with paperwork sticking out. There was a desk which had part of a computer on it which looked like it needed repairing and it fell over. There was a plant in the corner of the room. In the right hand corner of the room there was a computer chair which moved when I went close to it, this surprised me, as I did not expect it to move. There were other items which made noises and moved, the door opened when I was ready to leave.

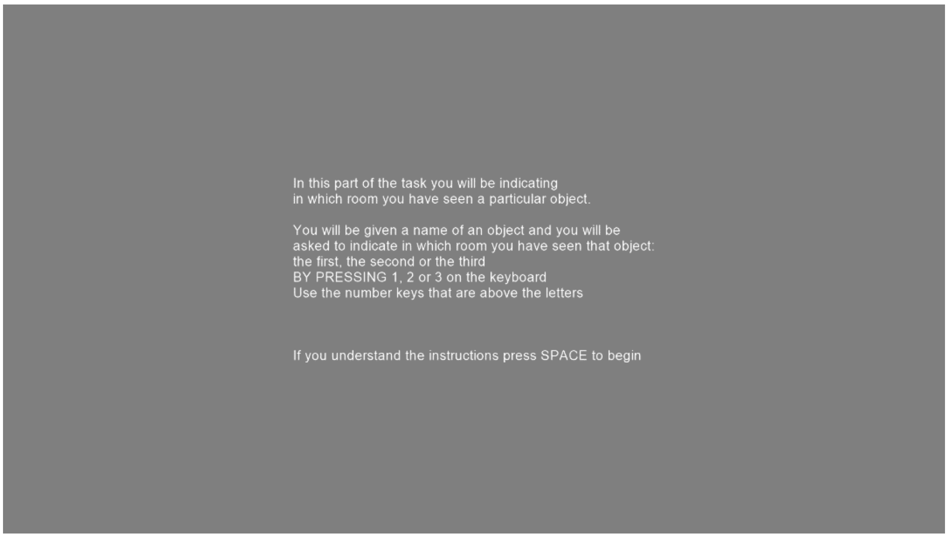
I went into a kitchen, there was a sink and the washing up liquid fell into it. There was a cooker with a saucepan on it, this made a noise when I went close to it. The fridge/freezer doors opened and closed. There was a coffee machine which switched on, this made a noise. There was a red toaster next to the fridge/freezer. Other items made sounds and moved and the door opened when I was ready to leave.

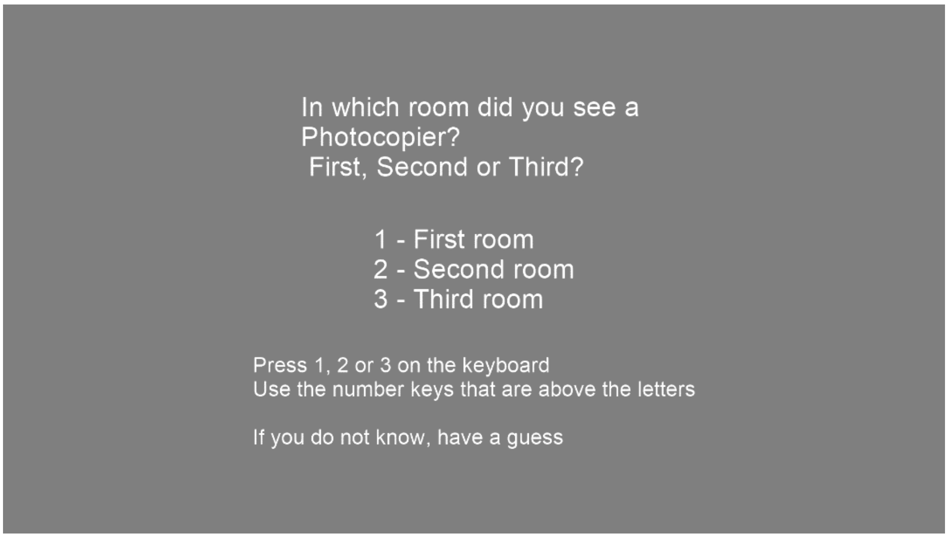
Appendix D: The What-When-Where and object recognition tasks used in Experiment 1

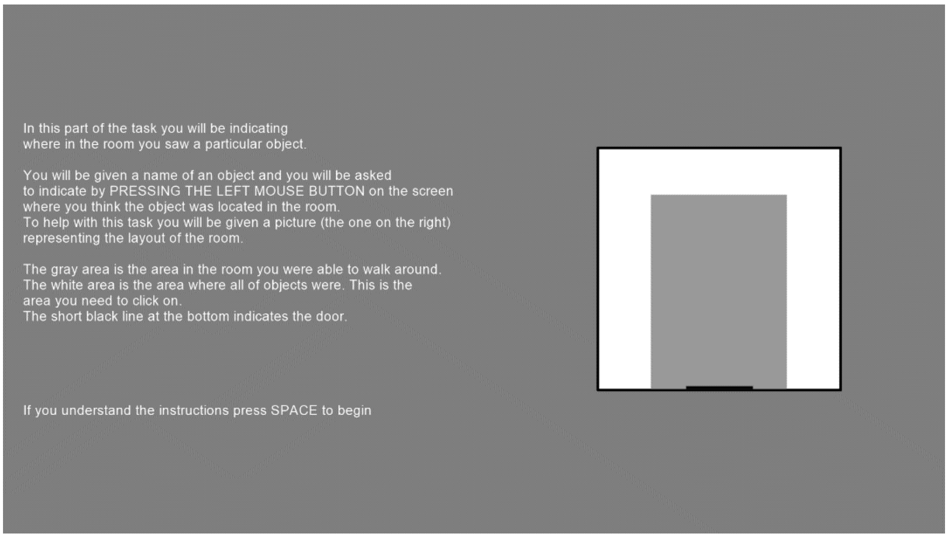
The What-When-Where task



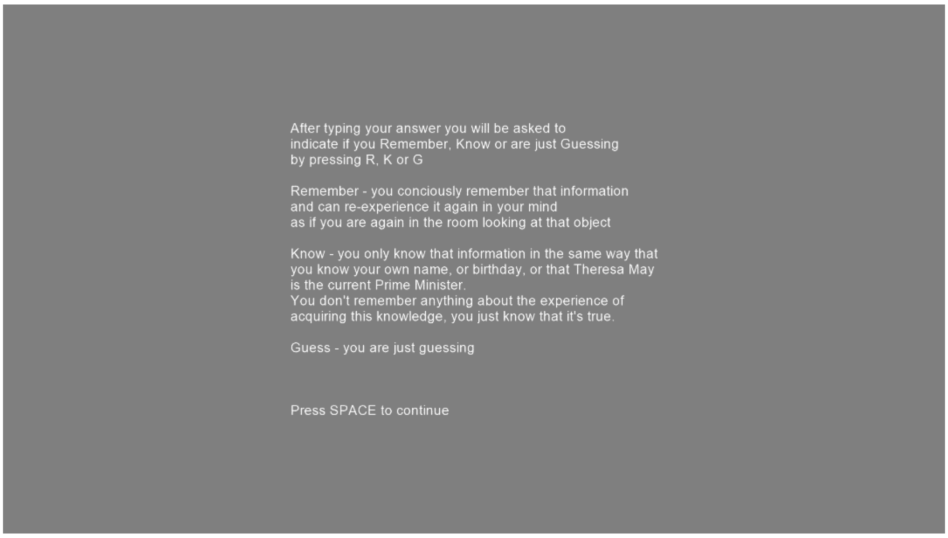

Do you remember seeing a
Photocopier?
(Y)es / (N)o
Press Y or N on the keyboard

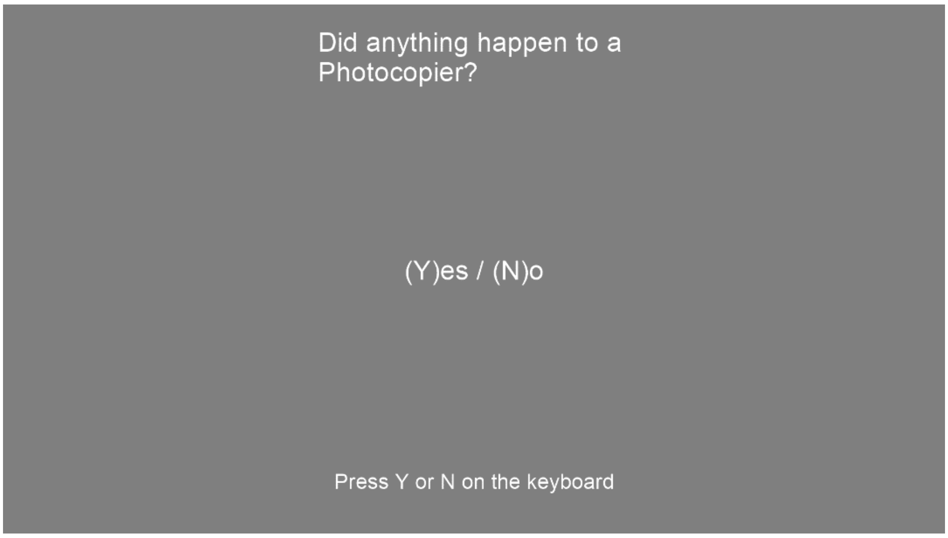

In this part of the task you will be indicating
in which room you have seen a particular object.
You will be given a name of an object and you will be
asked to indicate in which room you have seen that object:
the first, the second or the third
BY PRESSING 1, 2 or 3 on the keyboard
Use the number keys that are above the letters
If you understand the instructions press SPACE to begin


In which room did you see a
Photocopier?
First, Second or Third?
1 - First room
2 - Second room
3 - Third room
Press 1, 2 or 3 on the keyboard
Use the number keys that are above the letters
If you do not know, have a guess


In this part of the task you will be indicating
where in the room you saw a particular object.
You will be given a name of an object and you will be asked
to indicate by PRESSING THE LEFT MOUSE BUTTON on the screen
where you think the object was located in the room.
To help with this task you will be given a picture (the one on the right)
representing the layout of the room.
The gray area is the area in the room you were able to walk around.
The white area is the area where all of objects were. This is the
area you need to click on.
The short black line at the bottom indicates the door.
If you understand the instructions press SPACE to begin


In this task you will be asked to indicate
if anything happened to an object.
If you think something has happened - PRESS Y
If you think that nothing happened to that object - PRESS N
If you indicate that something happened to the object
you will be taken to a next screen in which you will
be asked to type your answer and press ENTER to submit.
Press SPACE to continue

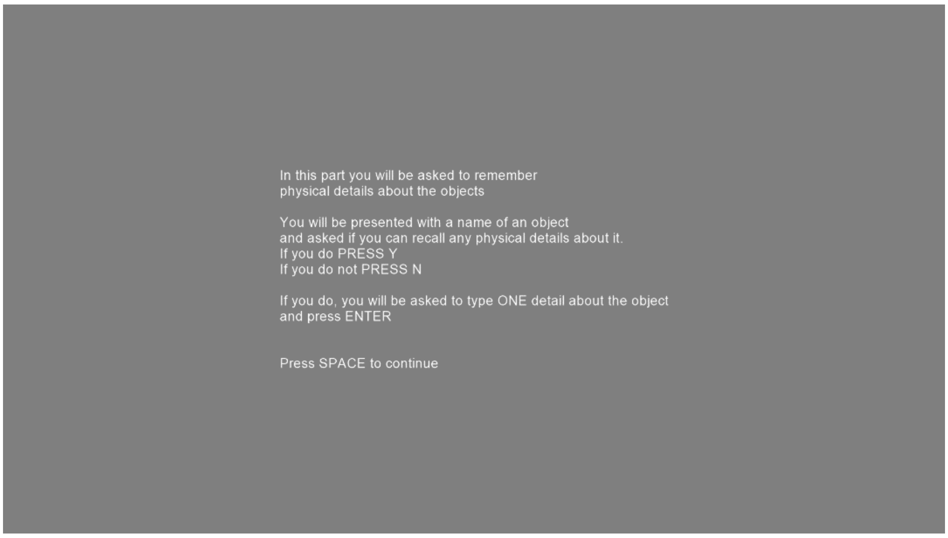

After typing your answer you will be asked to
indicate if you Remember, Know or are just Guessing
by pressing R, K or G
Remember - you conciously remember that information
and can re-experience it again in your mind
as if you are again in the room looking at that object
Know - you only know that information in the same way that
you know your own name, or birthday, or that Theresa May
is the current Prime Minister.
You don't remember anything about the experience of
acquiring this knowledge, you just know that it's true.
Guess - you are just guessing
Press SPACE to continue



Did anything happen to a
Photocopier?
(Y)es / (N)o
Press Y or N on the keyboard

What happened with a Photocopier?
Type your response here:
It started scanning
Press ENTER if you are done typing
or if you cannot remember anything

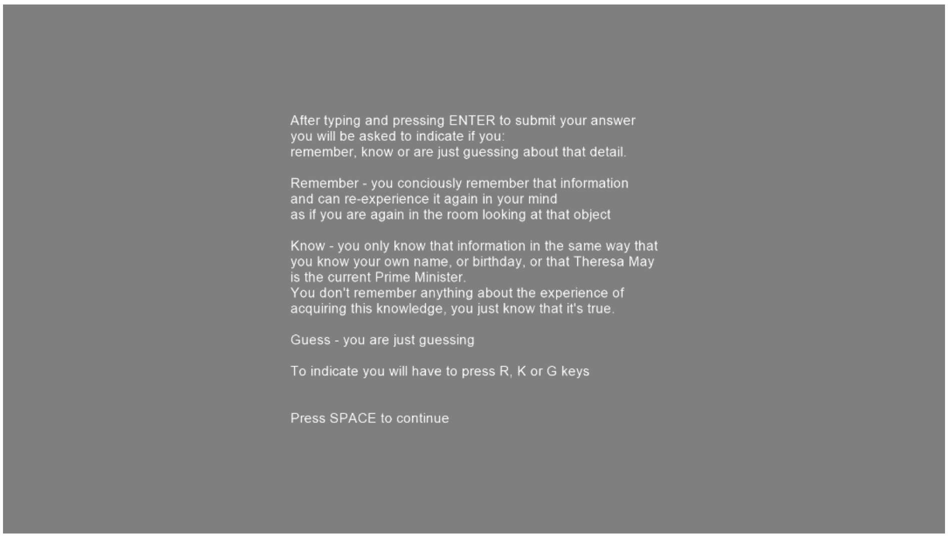
Do you Remember, Know or are just
guessing?
The object was a Photocopier
It started scanning
(R)emember (K)now (G)uess
Remember - you conciously remember that information
and can re-experience it again in your mind
as if you are again in the room looking at that object
Know - you only know that information in the same way that
you know your own name, or birthday, or that Theresa May
is the current Prime Minister.
You don't remember anything about the experience of
acquiring this knowledge, you just know that it's true.
Guess - you are just guessing


In this part you will be asked to remember
physical details about the objects

You will be presented with a name of an object
and asked if you can recall any physical details about it.
If you do PRESS Y
If you do not PRESS N

If you do, you will be asked to type ONE detail about the object
and press ENTER

Press SPACE to continue


After typing and pressing ENTER to submit your answer
you will be asked to indicate if you:
remember, know or are just guessing about that detail.

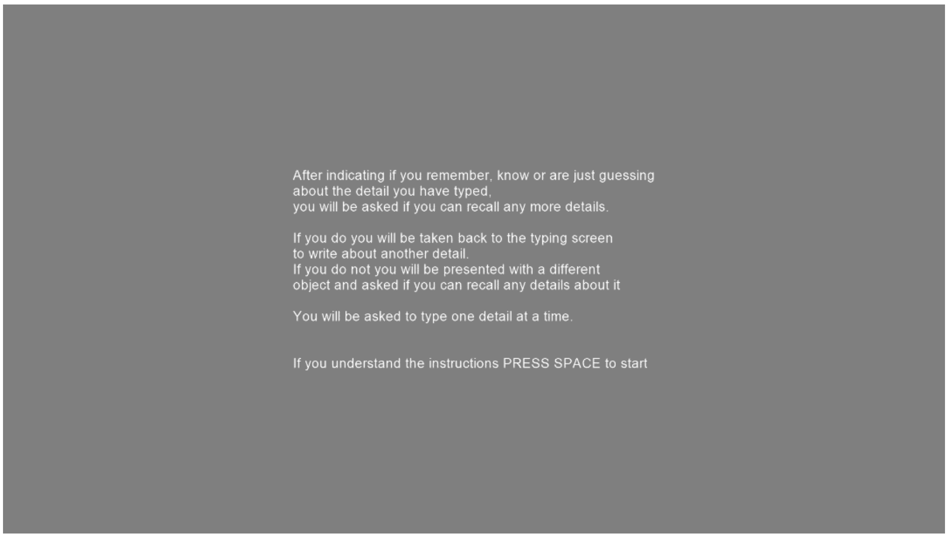
Remember - you conciously remember that information
and can re-experience it again in your mind
as if you are again in the room looking at that object

Know - you only know that information in the same way that
you know your own name, or birthday, or that Theresa May
is the current Prime Minister.
You don't remember anything about the experience of
acquiring this knowledge, you just know that it's true.

Guess - you are just guessing

To indicate you will have to press R, K or G keys

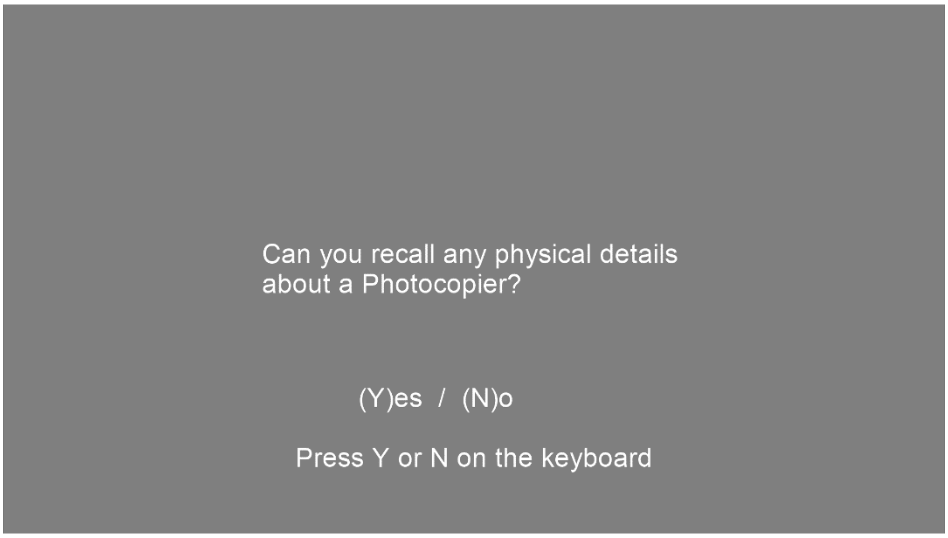
Press SPACE to continue


After indicating if you remember, know or are just guessing
about the detail you have typed,
you will be asked if you can recall any more details.

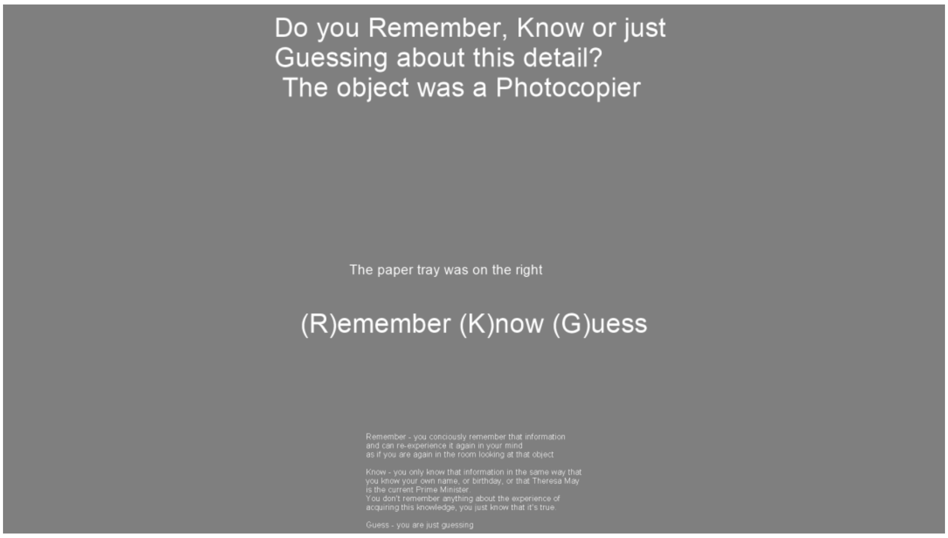
If you do you will be taken back to the typing screen
to write about another detail.
If you do not you will be presented with a different
object and asked if you can recall any details about it

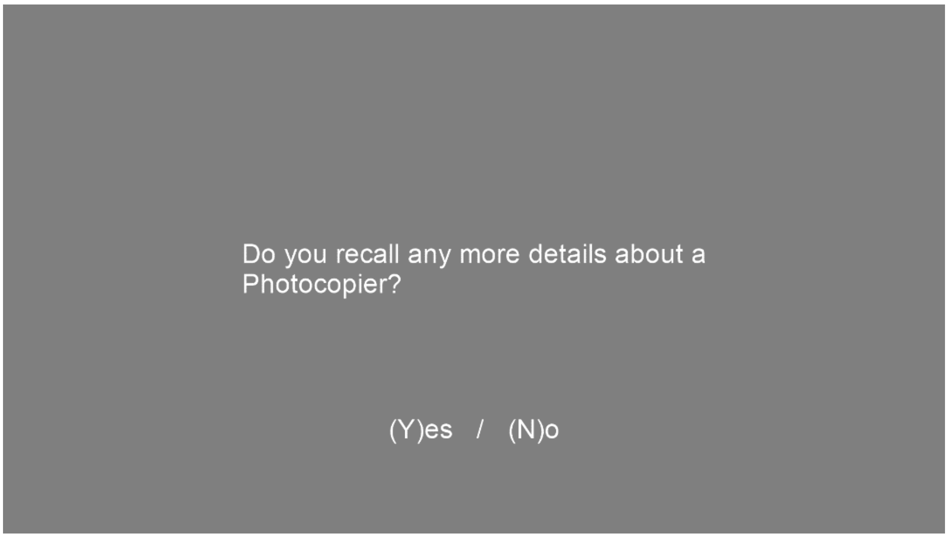
You will be asked to type one detail at a time.

If you understand the instructions PRESS SPACE to start

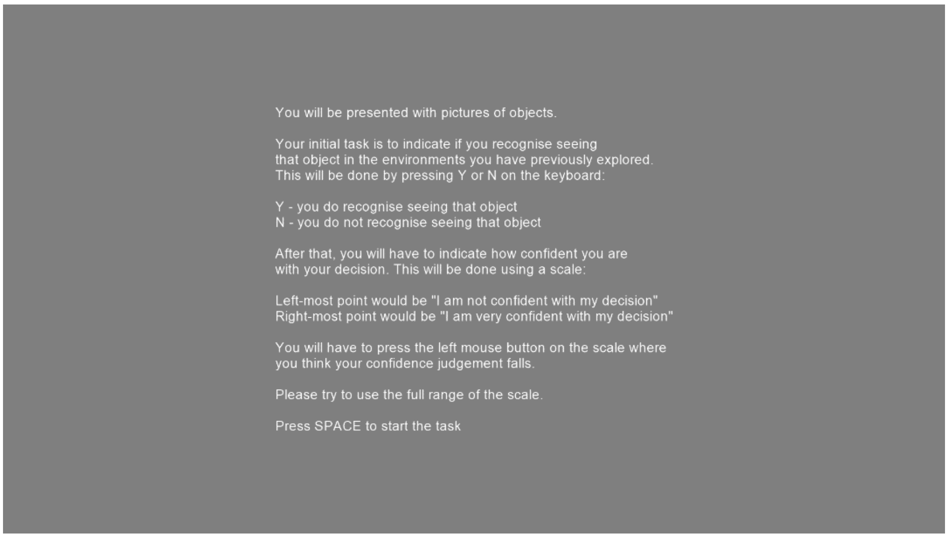

Can you recall any physical details
about a Photocopier?
(Y)es / (N)o
Press Y or N on the keyboard

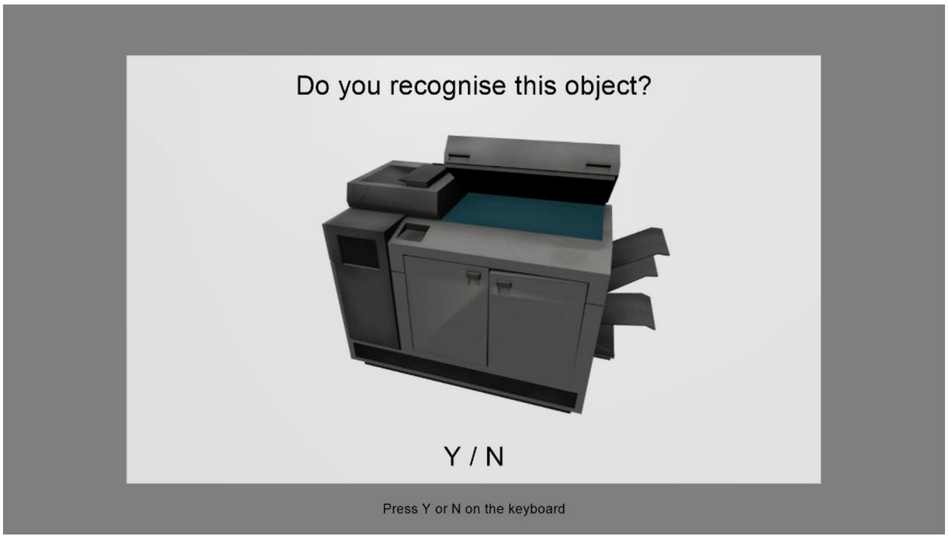

Describe ONE physical detail about a
Photocopier
Type your response here:
The paper tray was on the right
Press ENTER if you are done typing

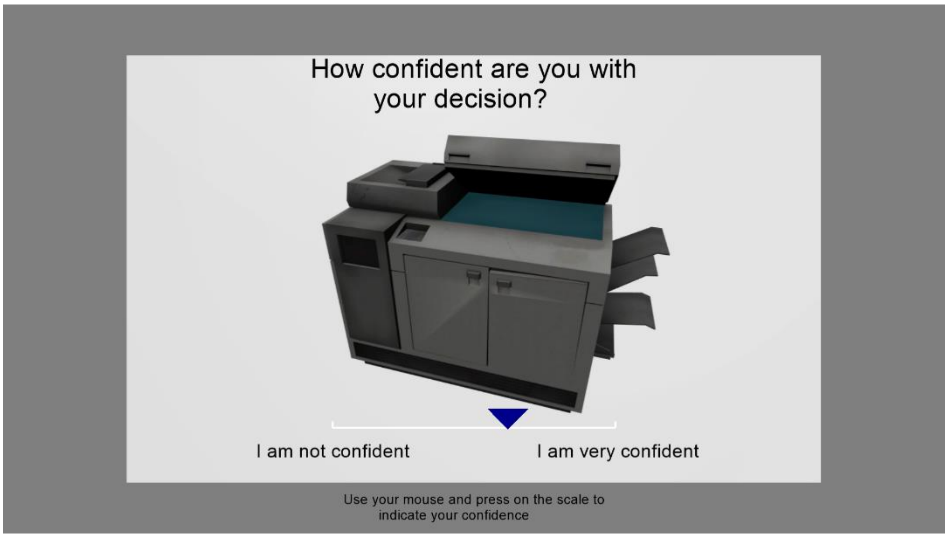

Do you Remember, Know or just
Guessing about this detail?
The object was a Photocopier
The paper tray was on the right
(R)emember (K)now (G)uess
Remember - you conciously remember that information
and can re-experience it again in your mind
as if you are again in the room looking at that object
Know - you only know that information in the same way that
you know your own name, or birthday, or that Theresa May
is the current Prime Minister.
You don't remember anything about the experience of
acquiring this knowledge, you just know that it's true.
Guess - you are just guessing


Do you recall any more details about a Photocopier?
(Y)es / (N)o

The object recognition task

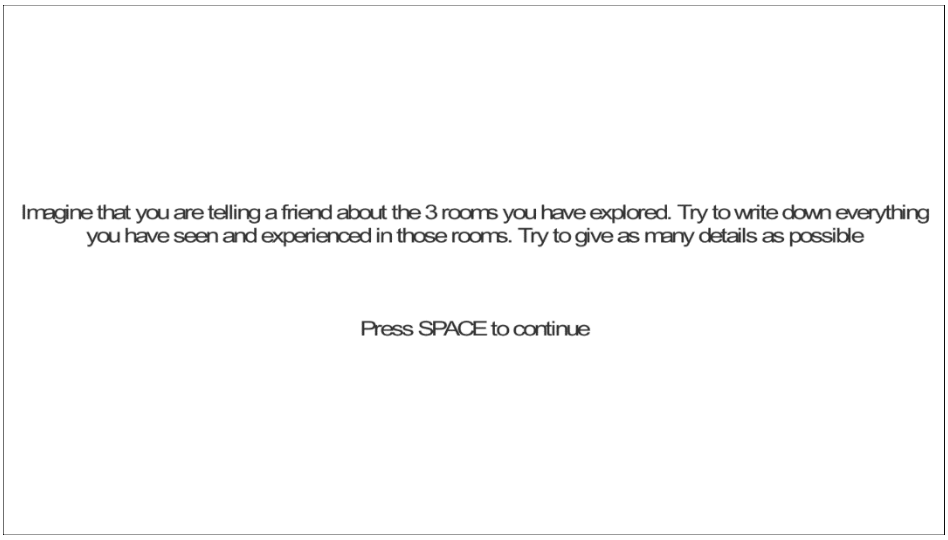


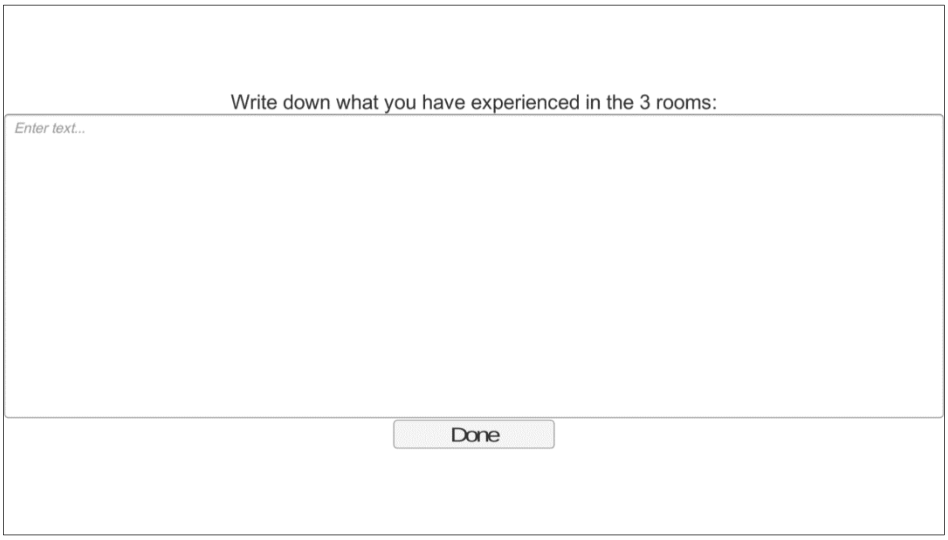




Appendix E: The memory tasks used in Experiments 2 and 3

The free recall task





The What-When-Where task

In this part of the task you will be shown names of objects
and you will be asked to indicate if you recognise those objects
from the environments you have explored.

Press Y if you recognise that object from the environments
Press N if you do not recognise that object

Press SPACE to continue

After giving your answer you will be asked to
indicate if you Remember, Know or are just Guessing
by pressing R, K or G

Remember - you conciously remember that information
and can re-experience it again in your mind
as if you are again in the room looking at that object

Know - you only know that information in the same way that
you know your own name, or birthday, or that Theresa May
is the current Prime Minister.
You don't remember anything about the experience of
acquiring this knowledge, you just know that it's true.

Guess - you are just guessing

Press SPACE to start

Do you remember seeing a Phone ?

(Y)es / (N)o

Press Y or N on the keyboard

a Phone

Do you -

(R)emember seeing this object
(K)now you saw this object
or are just (G)uessing

Press the corresponding key on the keyboard

In this part of the task you will be indicating
in which room you have seen a particular object.

You will be given a name of an object and you will be
asked to indicate in which room you have seen that object:
the first, the second or the third
BY PRESSING 1, 2 or 3 on the keyboard
Use the number keys that are above the letters

Press SPACE to continue

After giving your answer you will be asked to
indicate if you Remember, Know or are just Guessing
by pressing R, K or G

Remember - you conciously remember that information
and can re-experience it again in your mind
as if you are again in the room looking at that object

Know - you only know that information in the same way that
you know your own name, or birthday, or that Theresa May
is the current Prime Minister.
You don't remember anything about the experience of
acquiring this knowledge, you just know that it's true.

Guess - you are just guessing

In which room have you seen a Phone ?

1st 2nd 3rd

Press 1, 2 or 3 on the keyboard

---

a Phone

In the first room

Do you -

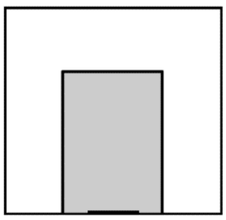
(R)emember seeing it in that room
(K)now that it was in that room
or are just (G)uessing

Press the corresponding key on the keyboard

In this part of the task you will be indicating
where in the room you saw a particular object.

You will be given a name of an object and you will be asked
to indicate by PRESSING THE LEFT MOUSE BUTTON on the screen
where you think the object was located in the room.
To help with this task you will be given a picture (the one on the right)
representing the layout of the room.

The gray area is the area in the room you were able to walk around.
The white area is the area where all of objects were. This is the
area you need to click on.
The short black line at the bottom indicates the door.



Press SPACE to continue

After giving your answer you will be asked to
indicate if you Remember, Know or are just Guessing
by pressing R, K or G

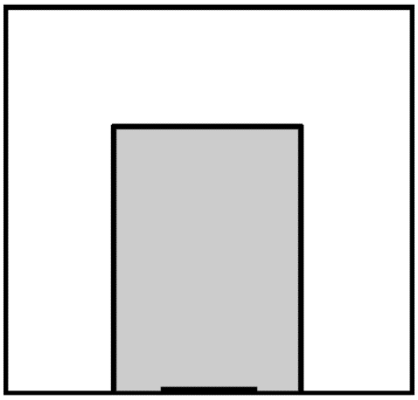
Remember - you conciously remember that information
and can re-experience it again in your mind
as if you are again in the room looking at that object

Know - you only know that information in the same way that
you know your own name, or birthday, or that Theresa May
is the current Prime Minister.
You don't remember anything about the experience of
acquiring this knowledge, you just know that it's true.

Guess - you are just guessing

Press SPACE to start

Where in the room have you seen a Phone ?



The gray area is the
area you were
able to walk around

The white area is the
remaining room
where the objects were.
This is where you need
to press.

Please use the mouse and press on the white part of the image where you think the object was in the room.

a Phone

Do you -

(R)emember the object being there
(K)now that it was there
or are just (G)uessing

Press the corresponding key on the keyboard

In this task you will be asked to indicate
if anything happened to an object.

If you think something has happened - PRESS Y
If you think that nothing happened to that object - PRESS N

If you indicate that something happened to the object
you will be taken to a next screen in which you will
be asked to type your answer and press DONE to submit.

Press SPACE to continue

After typing your answer you will be asked to
indicate if you Remember, Know or are just Guessing
by pressing R, K or G

Remember - you conciously remember that information
and can re-experience it again in your mind
as if you are again in the room looking at that object

Know - you only know that information in the same way that
you know your own name, or birthday, or that Theresa May
is the current Prime Minister.
You don't remember anything about the experience of
acquiring this knowledge, you just know that it's true.

Guess - you are just guessing

Press SPACE to start

Did anything happened to a Phone ?

(Y)es / (N)o

Press Y or N on the keyboard


Did anything happened to a Phone ?
It started ringing
Done


a Phone
It started ringing
Do you -
(R)emember (K)now or are just (G)uessing
Press the corresponding key on the keyboard


In this part you will be asked to remember
physical details about the objects
You will be presented with a name of an object
and asked if you can recall any physical details about it.
If you do PRESS Y
If you do not PRESS N
If you do, you will be asked to type ONE detail about the object
and press DONE
Press SPACE to continue

After typing and pressing NEXT to submit your answer
you will be asked to indicate if you:
remember, know or are just guessing about that detail.

Remember - you conciously remember that information
and can re-experience it again in your mind
as if you are again in the room looking at that object

Know - you only know that information in the same way that
you know your own name, or birthday, or that Theresa May
is the current Prime Minister.
You don't remember anything about the experience of
acquiring this knowledge, you just know that it's true.

Guess - you are just guessing

To indicate you will have to press R, K or G keys

Press SPACE to continue

After indicating if you remember, know or are just guessing
about the detail you have typed,
you will be asked if you can recall any more details.

If you do you will be taken back to the typing screen
to write about another detail.
If you do not you will be presented with a different
object and asked if you can recall any details about it

You will be asked to type one detail at a time.

PRESS SPACE to start

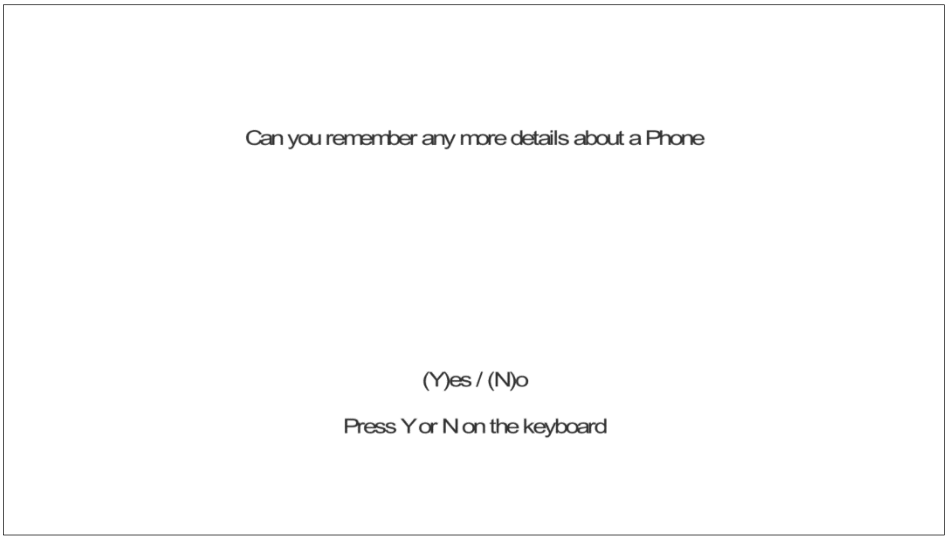
Can you remember any details about a Phone

(Y)es / (N)o

Press Y or N on the keyboard


Type ONE detail about a Phone
It was grey
Done


a Phone
It was grey
Do you -
(R)emember this detail
(K)now it
or are just (G)uessing
Press the corresponding key on the keyboard


Can you remember any more details about a Phone
(Y)es / (N)o
Press Y or N on the keyboard

The object recognition task

You will be presented with pictures of objects.

Your initial task is to indicate if you recognise seeing
that object in the environments you have previously explored.
This will be done by pressing Y or N on the keyboard:

Y - you do recognise seeing that object
N - you do not recognise seeing that object

After that, you will have to indicate how confident you are
with your decision. This will be done using a scale:

Left-most point would be "I am not confident with my decision"
Right-most point would be "I am very confident with my decision"

You will have to press the left mouse button on the scale where
you think your confidence judgement falls.

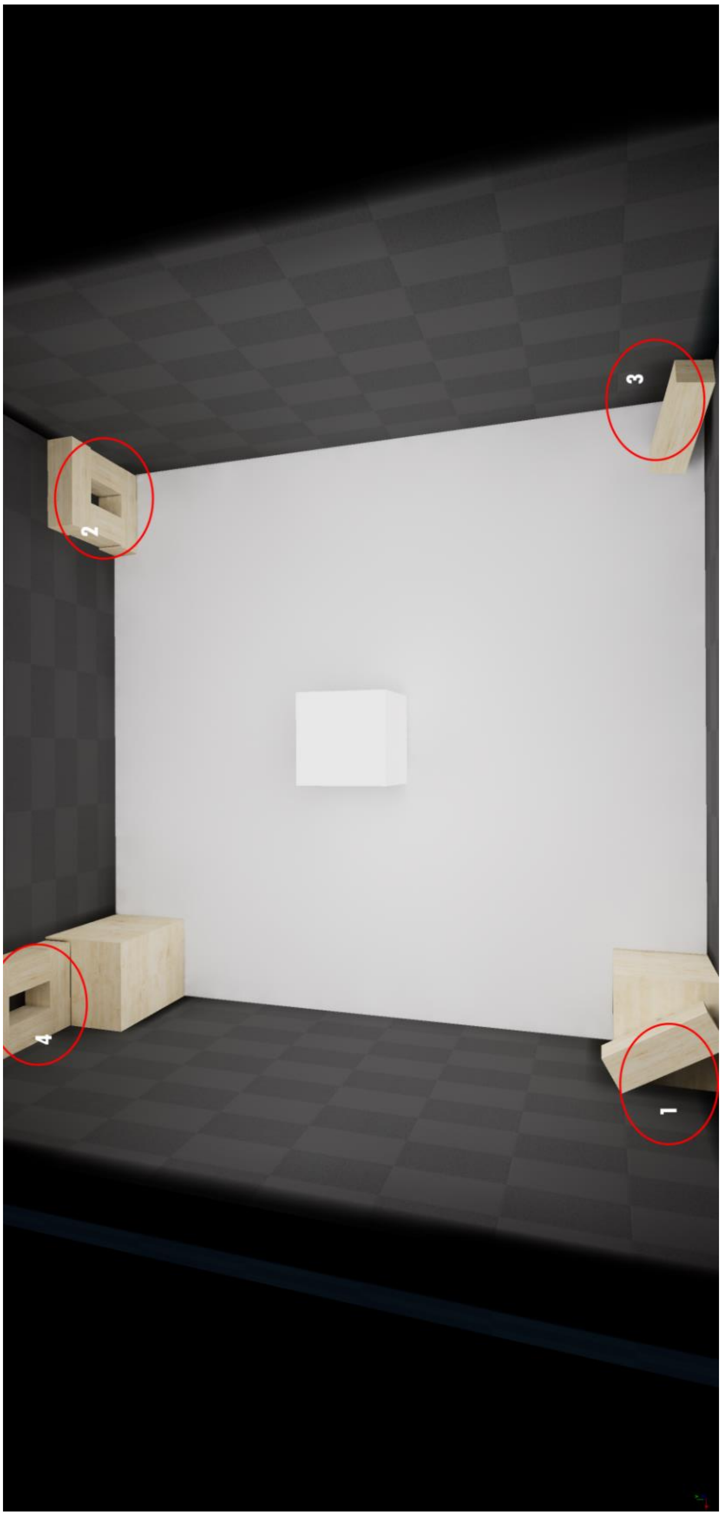
Please try to use the full range of the scale.

Press SPACE to start the task

Appendix F: Environments used in Experiment 4

Red circles indicate the locations in which participants had to hide objects. The objects for the participants to hide were initially placed on a box or table in the middle of each room. The numbers indicate the order in which the objects had to be hidden.

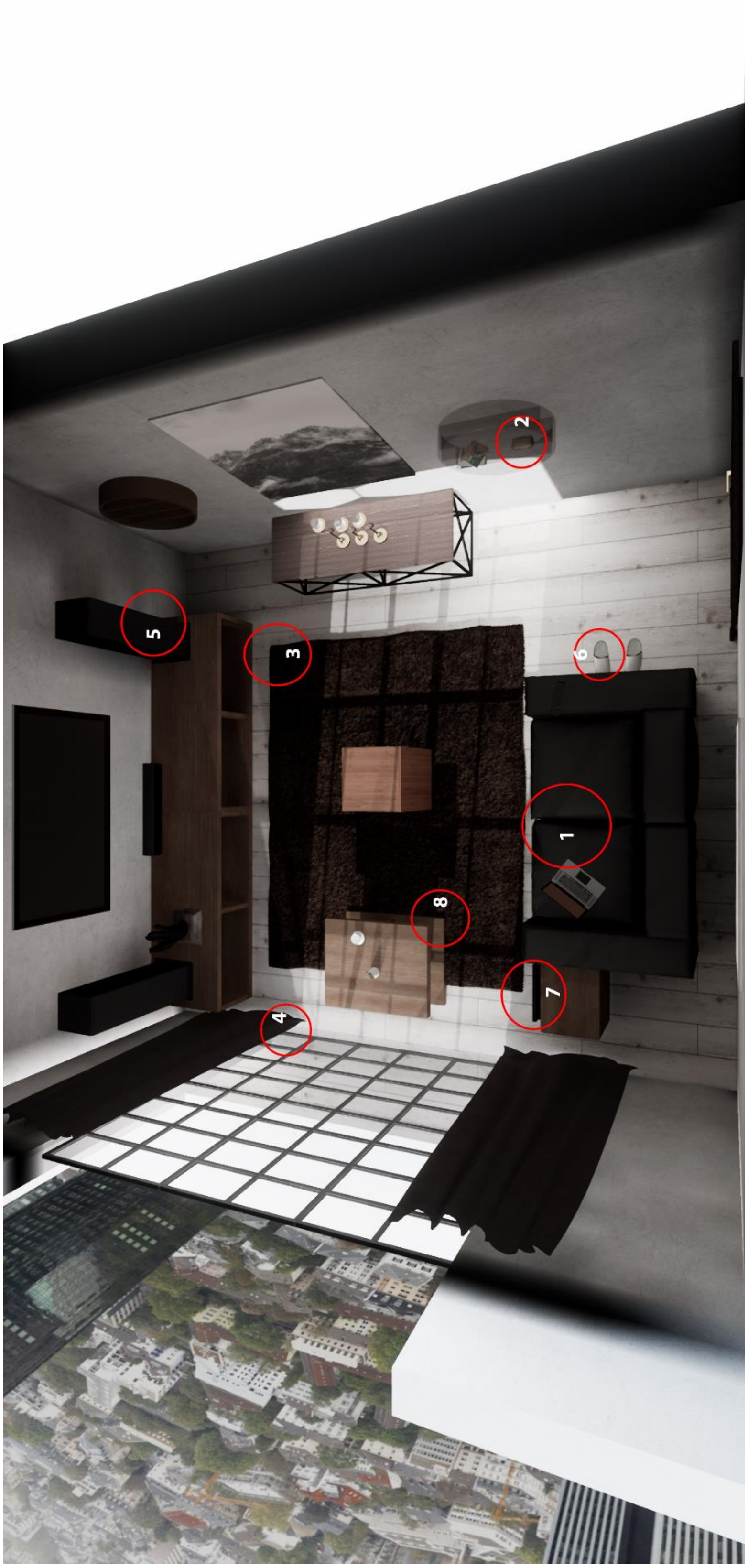
Desktop-VR and HMD-VR training room



HMD-VR room

Notes: The bookshelf between the two windows was made transparent for this image to better show the hiding locations.

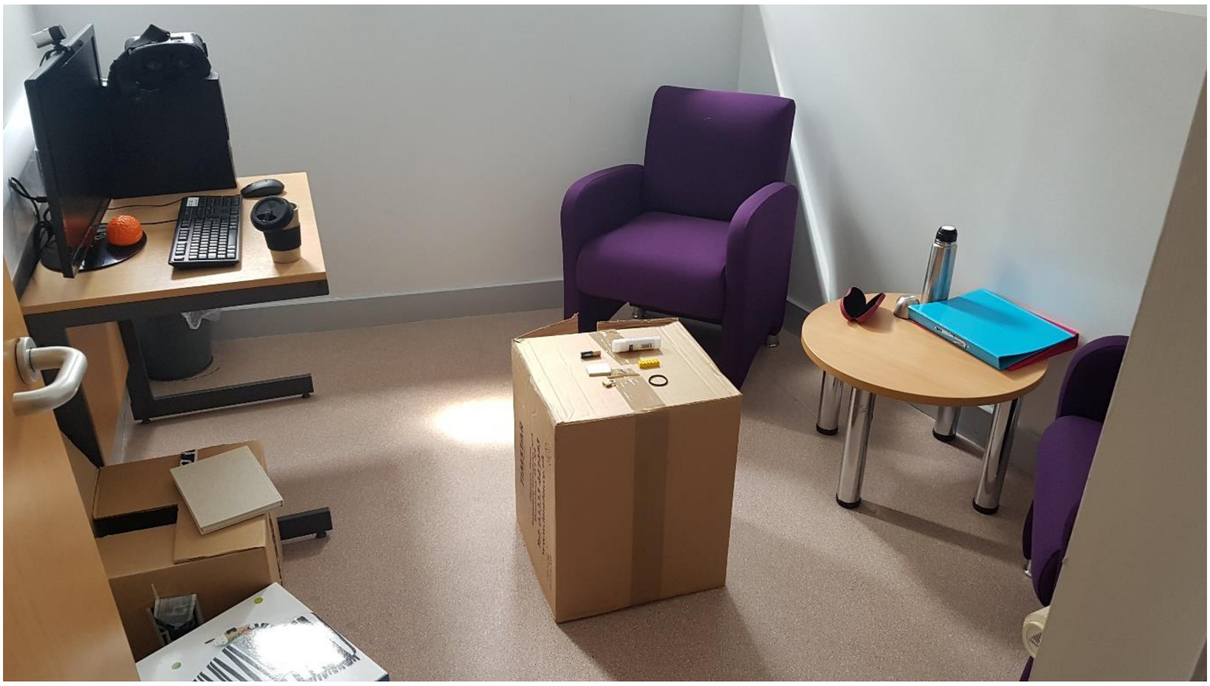
Desktop-VR room



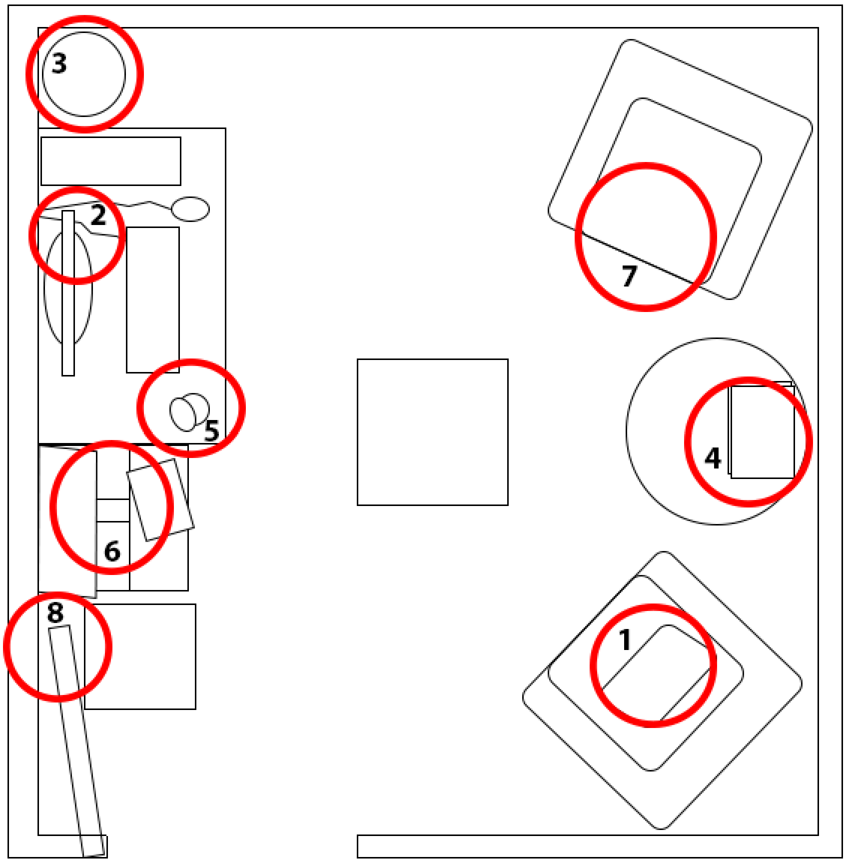
Notes: The round shelf for the second hiding place was made transparent for this image to better show the hiding location.

Real-life room

Photo of the room



Top down map of the room



Appendix G: Objects and locations used in Experiment 4

Notes: The object numbering corresponds to the location numbers in Appendix VII. Objects numbered 1 to 4 were in the first group of objects while objects numbered 5 to 8 were in the second group of objects. Non-numbered objects in each group represent objects the participants did not had to hide but were still mixed in with the main objects.

Desktop-VR

Group 1

1) Clothes peg – in the middle of the sofa, between the cushions.
2) Tea-light – in a pencil case, on a shelf.
3) Pen – under the carpet corner.
4) Toy car – Behind the curtain.

- Star
- Empty bottle

Group 2

5) Bow tie – Behind a speaker.
6) Key – in a slipper.
7) Nail clippers – in a drawer.
8) Bracelet – under a table corner.

- Bolt
- Ruler

HMD-VR

Group 1

1) Lego figure – under a book.
2) Dice – behind the mirror.
3) Ring – under a pillow on the sofa.
4) Bottle cap – on a shelf, behind a statue of a lion.

- USB stick
- Comb

Group 2

5) Chess piece – on a chair under a cloth.
6) Lighter – in a plant pot.
7) Tape – on a shelf, between the books.
8) Crayon – in a vase.

- Eraser
- Spoon

Real-life

Group 1

1) Battery – in a backpack
2) Padlock – under the computer monitor
3) Hair bobble – in the rubbish bin.
4) Post-it notes – in a folder on the table.

- Lego brick
- Glue stick

Group 2

5) Paperclip – in a coffee cup.
6) Trolley token – in a cardboard box.
7) Hook – under a chair.
8) Cotton earbud – behind the door.

- Sharpener
- Pen top

Lures used in the object recognition task

Eye drops	Floppy disk
Domino	Screwdriver
Button	Matchbox
Stopwatch	Tube
Marker	Flashlight
Diamond	Pin
Décor	Tin opener
Spinner	Compass
Nut	Hair pin
Toothbrush	Watch
Plug	Hand grip
Earpod	Fork
Scissors	Heart
Cigarette	Golf ball
Saltshaker	Duck
Lipstick	
Swiss knife	
Carabine hook	
Bank card	
Playing card	
Dart	

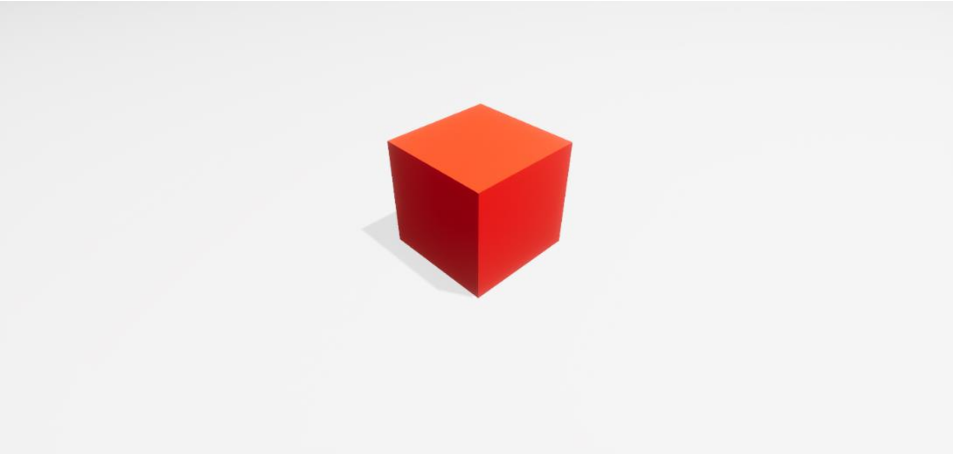
2.

Appendix H: Object lists given to participants in Experiment 4

Desktop-VR and HMD-VR training room

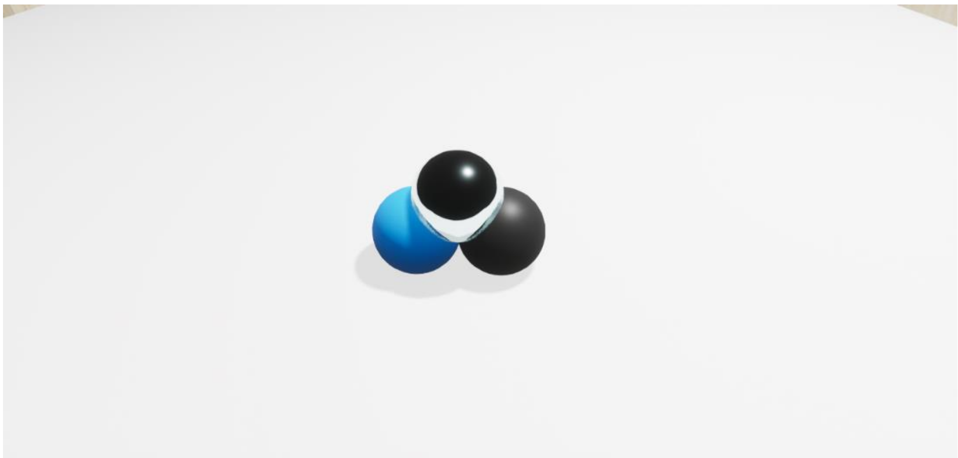
Notes: Same lists were given to both Desktop-VR and HMD-VR participant groups. In the HMD-VR group the lists were shown in the VE after a participant pressed a button on their controller. In the Desktop-VR group the lists were printed on paper and placed next to the computer.

Group 1

1.


2.


Group 2

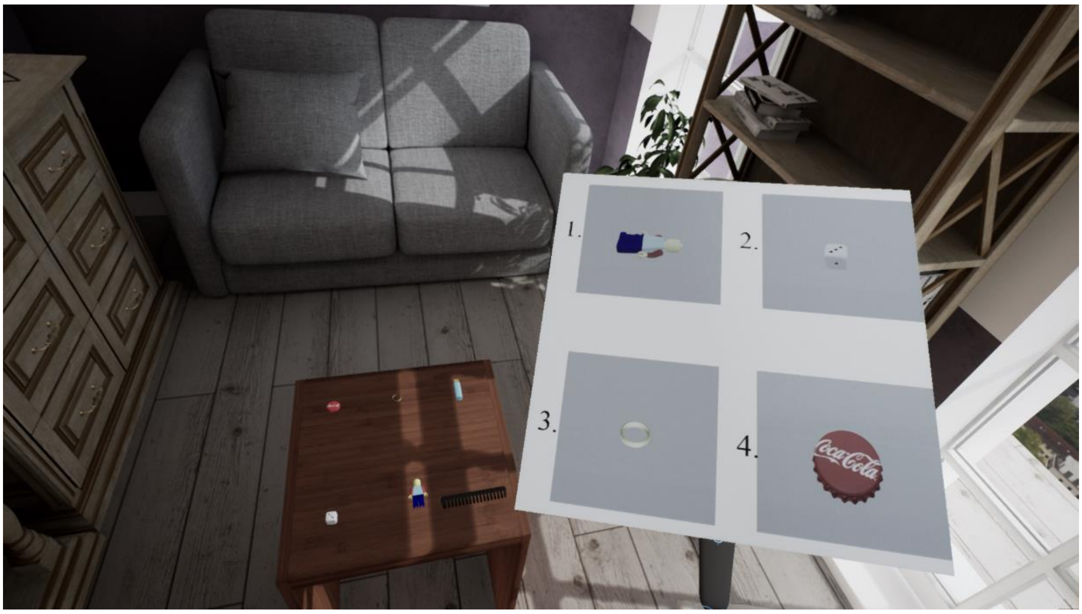
1. 

2. 

HMD-VR room

Notes: The object lists were shown in the VE after a participant pressed a button on their controller.

Group 1



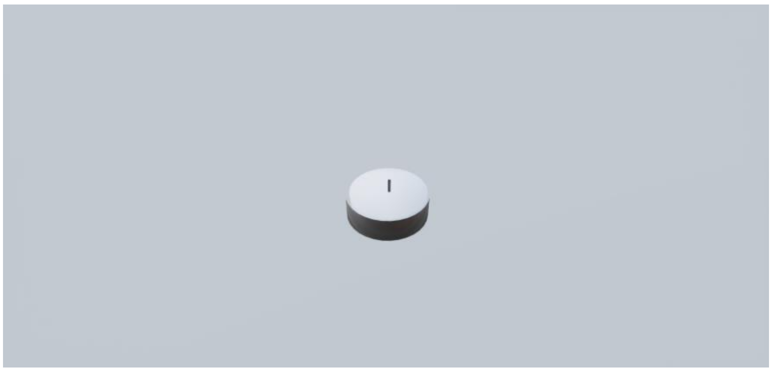
Group 2

Desktop-VR main room

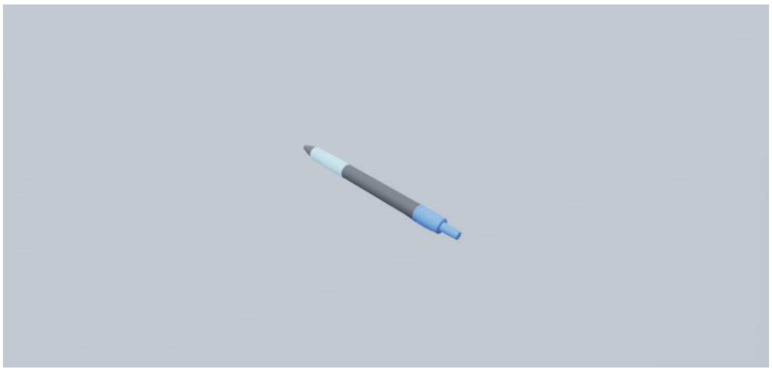
Group 1

1.

2.



3.

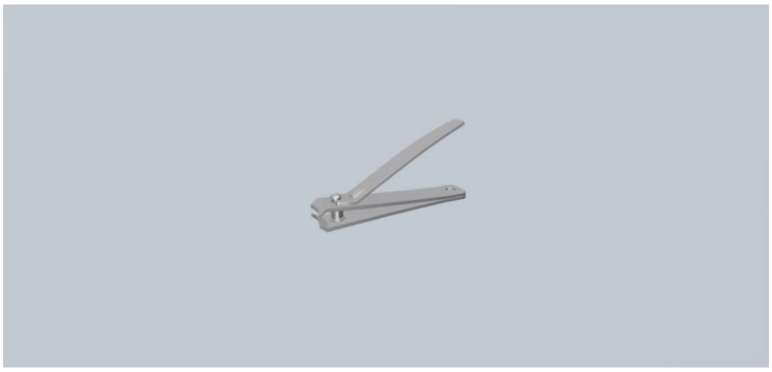


4.

Group 2

1.


2.


3.


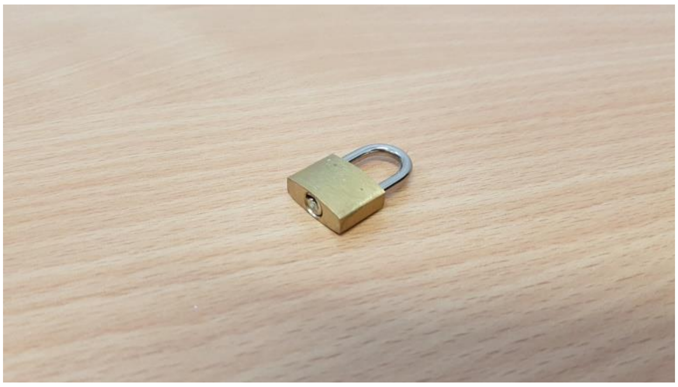
4.


Real-life room

Group 1

1.

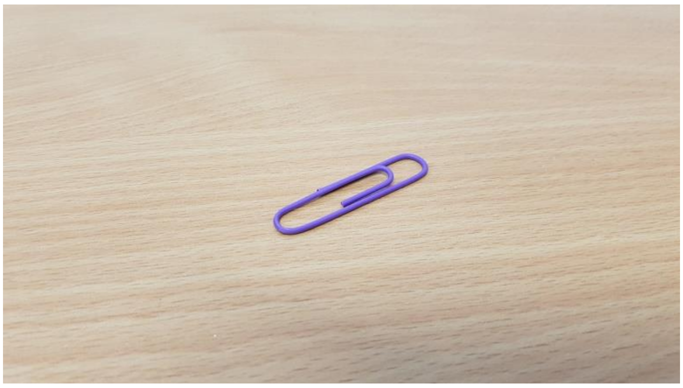
2.



3.

4.

Group2

1.


2.


3.


4.