Exploring the role of sleep on recognition memory and gist-based false memory

A thesis submitted to The University of Manchester for the degree of Doctor of Philosophy in the Faculty of Biology, Medicine and Health

2018

Hikaru Tsujimura

School of Biological Sciences

Contents

Total word count: 33,763

List of Figures

List of Tables

List of Abbreviations

iOtA	The Information Overlap To Abstract
MTL	Medial Temporal Lobe
CLS	Complementary Learning System
mPFC	Medial Prefrontal
PSG	Polysomnography
EEG	Electroencephalography
EOG	Electrooculography
EMG	Electromyography
REM	Rapid Eye Movement
NREM	Non-Rapid Eye Movement
SWS	Slow Wave Sleep
PGO	Ponto-geniculo-occipital
SO	Slow Oscillations
SWR	Sharp Wave Ripples
RE	Thalamic Reticular
CT	Corticothalamic
TC	Thalamocortical
ASC	The Active Systems Consolidation Theory
SHY	The Synaptic Homeostasis Hypothesis
SNR	Signal-to-noise-ratio
LTP	Long Term Potentiation
SSS	Stanford sleepiness Scale
KSS	Karolinska Sleepiness Scale
DRM	The Deese Roediger McDermott
fMRI	Functional Magnetic Resonance Imaging
SPM	Statistical Parametric Mapping
EPI	Echo Planar Imaging
TR	Repetition Time
TE	Echo Time
GLM	The General Linear Model
RSA	Representational Similarity Analysis

Abstract

Exploring the role of sleep on recognition memory and gist-based false memory

Hikaru Tsujimura, the University of Manchester

For the degree of Doctor of Philosophy (PhD), May 2018

To date, there is extensive evidence showing that sleep is important for memory consolidation. Specifically, a night of sleep seems to either strengthened many forms of memory or protect them against decay. The information overlap to abstract (iOtA) model suggests that when memories are replayed in an overlapping way the combination of potentiation for overlapped areas and downscaling for everything else leads to extraction of gist. In this thesis, I set out to test the above proposal.

In order to control the extent to which memories overlap or remain isolated, I developed a novel paradigm. Namely, I morphed facial images along both age and gender, creating a 2D 'face space'. I conducted Experiment 1 to determine how far apart images need to be in this face space in order to be remembered as distinct (not merged). I was then able to train participants on selected images from this space. These learned images were either close together (densely packed representations) or far apart (sparsely packed representations). Participants could then be tested for both veridical memory of the learned (old) images, and false memory of unstudied items in any part of the face space.

In a series of experiments, I then examined the impact of sleep on both veridical and false memory for items falling in both dense and sparsely populated areas of the face space. As expected, I consistently found that veridical memories were protected by sleep. However, surprisingly, Experiment 2 showed an overnight increase in false recognition for the face images in sparsely populated areas of this space, with no change in false recognition for images in densely populated areas. Experiment 3 replicated this finding even when I controlled spatial features of facial images, except face density. Polysomnography in Experiment 3 also showed a strong correlation between slow wave sleep and the extent of increased false recognition in the 'sparse' area. Next, I conducted Experiment 4 as a control study using inverted faces, and revealed that the sleep effect was not specific to facial images. This suggests that the finding generalises at least as far as these control images. In Experiment 5 I used a finer grained map of the face-space by using smaller morphing steps. This study suggested that, across a night of sleep, compared to wake, the representation of a learned face actually shifts away from other learned face representations, . In Experiment 6, I conducted a neuroimaging study in order to determine the neural correlates of these shifts, and found that activity in face selective areas shifts in parallel with the behavioural shift, while activity in the hippocampus shifts in the opposite direction.

Overall, my findings do not support the iOtA model, as memory representations in densely populated areas were not strengthened over sleep. Instead, my work adds to the growing literature suggesting that sleep is not only important for strengthening memories, but also for selectively weakening some aspects of memory. Specifically, my work could suggest that sleep plays a role in pattern separation, as it seems to force the representations apart, even when this increases false alarms.

Declaration

No portion of the work referred to in this thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Copyright Statement

i. The author of this thesis (including any appendices and/or schedules to this thesis) owns certain copyright or related rights in it (the “Copyright”) and s/he has given The University of Manchester certain rights to use such Copyright, including for administrative purposes.

ii. Copies of this thesis, either in full or in extracts and whether in hard or electronic copy, may be made only in accordance with the Copyright, Designs and Patents Act 1988 (as amended) and regulations issued under it or, where appropriate, in accordance with licensing agreements which the University has from time to time. This page must form part of any such copies made.

iii. The ownership of certain Copyright, patents, designs, trade marks and other intellectual property (the “Intellectual Property”) and any reproductions of copyright works in the thesis, for example graphs and tables (“Reproductions”), which may be described in this thesis, may not be owned by the author and may be owned by third parties. Such Intellectual Property and Reproductions cannot and must not be made available for use without the prior written permission of the owner(s) of the relevant Intellectual Property and/or Reproductions.

iv. Further information on the conditions under which disclosure, publication and commercialisation of this thesis, the Copyright and any Intellectual Property and/or Reproductions described in it may take place is available in the University IP Policy (see http://documents.manchester.ac.uk/DocuInfo.aspx?DocID=24420), in any relevant Thesis restriction declarations deposited in the University Library, The University Library’s regulations (see http://www.library.manchester.ac.uk/about/regulations/) and in The University’s policy on Presentation of Theses.

Acknowledgements

I wish to give my greatest thanks to Prof. Penny Lewis, the one who spent the most of time encouraging, commenting, and directing me for improving every aspect of my PhD works. Big thanks as well to Prof. Sonja Kotz, Prof. Matthew Lambon Ralph, Dr. Karen Lander and Dr. Andrew Stewart, who gave a number of helpful comments on my works, insightful inputs to elaborate my works and advice on advanced statistical analysis.

I would like to give my thanks to all of my NAPS colleagues, past and present, especially Jules Schneider, Tia Tsimpanouli, and Isabel Hutchison, whom I spent the most of time at work, remembering moments to exchange foreign language, to keep updating locations of nice restaurants and more.

I also appreciated to my families, who often visited to where I was so that I did not get bothered to go back to my home country and my girlfriend who made me relaxed when spending times with her.

This work was funded by a School of Psychological Sciences (SPS) PhD Studentship 2013 from Psychology Department of the University of Manchester.

The author

Hikaru Tsujimura completed an undergraduate degree in Psychology at the Saint Louis University, USA, graduating with grades equivalent to First Class Honours (i.e. 3.62). He was then completed a Master of Science (MSc) degree in Cognitive Neuroscience with merit from Institute of Cognitive Neuroscience, University College London. After additional three years of research assistant work at Goldsmiths, University of London, he started a PhD degree at the University of Manchester.

Rationale for submitting the thesis in an alternative format

The work in this thesis forms the basis of three articles that were prepared for submission to scientific journals and are at various stages of preparation for publication. These articles form chapters 2-5. Chapter 1 will provide a broad introduction to all of the research undertaken in this thesis and outline the key research questions and aims.

The author was the primary investigator for all work presented in this thesis, performing the majority of experiment conceptualisation, design, data collection, analysis, interpretation and writing. Contributions from co-authors included sleep scoring (where scoring by two experimenters is a requirement), supervisor guidance for project design, analysis and interpretation, and assistance with collecting data.

Chapter 1:

General introduction

Section 1.0 Brief review about background of memory

1.0.1. “Re-constructive” nature of memory over time

Our memory is not a mere copy of past experiences, but instead dynamically varies over time (Bartlett, 1932; Loftus & Pickrell, 1995; Schacter, 1999). In fact, some information persists, some is forgotten, and some is replaced with new information. Thus, it has been said that our memory is inherently “re-constructive” with time (Schacter, Guerin, & Jacques, 2011).

1.0.2. Major taxonomy

To date, it has been widely accepted that memory can be divided into “declarative memory”, which is memory for concepts and autobiographical episodes, and “non-declarative memory” (for critical studies see Milner, 1962; Graf & Schacter, 1985 and for a review of see Squire, 2004; Squire & Dede, 2015). Initially, this major distinction comes from a number of amnesic patient studies which show that losing the ability to recall new concepts and autobiographical events after damaging to the medial temporal lobe (MTL), especially the hippocampus, leaves other types of memory, such as acquiring procedural memory of a mirror tracing task, intact (e.g. Milner, 1962).

Yet, subsequent studies (e.g. Graf & Schacter, 1985) have shown intact capacity of amnesic patients in priming effects, in which initial training of word pairs (target-cue) benefitted memory performance on a later word completion task with the trained

cues while participants were not consciously aware of remembering the trained cue. Thus, such studies suggest why hippocampus-dependent (declarative) memory is also referred as explicit memory as one recalls facts or autobiographical events explicitly (e.g. Graf & Schacter, 1985; Smith & Squire, 2008), whereas non-declarative memory (e.g. acquiring procedural memory of a mirror trace task in the work of Milner, 1962) is often measured using a change in task performance (e.g. faster reaction speed) compared to baseline performance.

Declarative memory is further separated into “semantic memory” and “episodic memory” depending on whether one is recalling facts/concepts/generalised patterns or episodic/autobiographical/single event, respectively (for a review read Tulving, 1972; Squire & Zola, 1998). See **Figure 1.1** for a summary of the memory taxonomy.

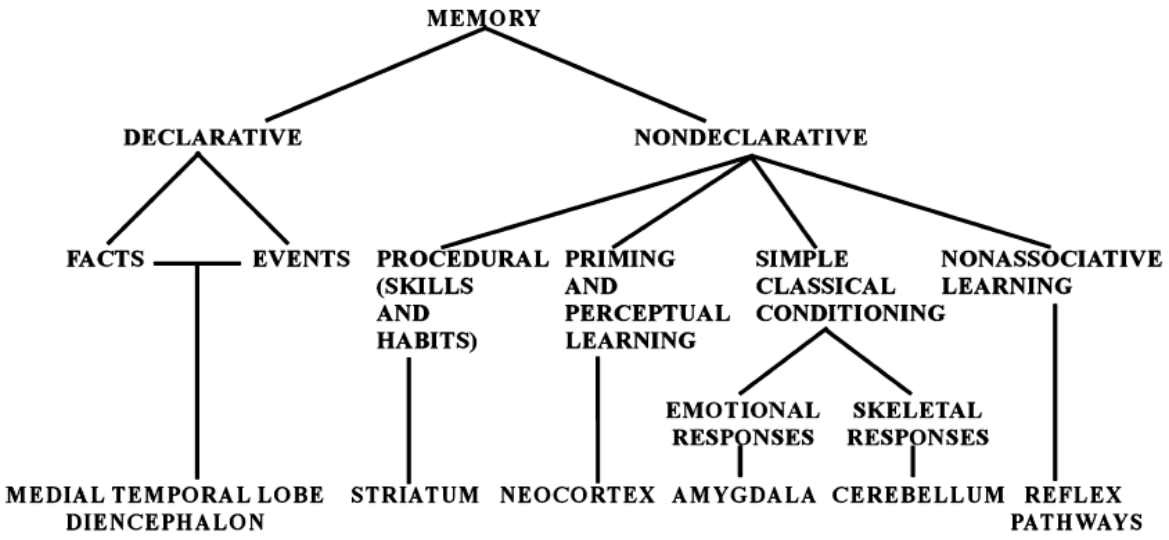

Figure 1.1. The taxonomy of memory. At first, memory is divided into declarative and non-declarative memory depending on presence or absence of one’s consciousness to access to the acquired information, respectively. Then, declarative memory is further divided into semantic memory (memory of facts) and episodic memory (memory of events). The figure is retrieved from Squire (2004).

Section 1.1. Memory models

1.1.1. Memory consolidation

Contrary to popular opinion, memory, especially declarative memory, does not remain static over time. Instead it evolves through the three major stages, i.e., (1) encoding, (2) storage and (3) retrieval (Melton, 1963). Encoding is the first stage in which new information is initially acquired. Next, the acquired information is maintained internally during an intermediate stage of storage. Finally, the internally maintained information is re-accessed at a stage of retrieval.

Among them, storage (later thought of as “consolidation”) of (declarative) memory is, at first, thought as a period to purely stabilize newly obtained memory over time (McGaugh, 2000). An initial form of memory is now thought to actually transform into a reconstructed form over time, integrating new information into pre-existing knowledge, with corresponding neural changes in the memory system (Frankland, & Bontempi, 2005; Nadel & Moscovitch, 1997; Scoville, & Milner, 1957). This ‘systems’ model of memory consolidation is mainly based on evidence from clinical studies of amnesic patients and has been elaborated further by studies simulating models of neural network-like structures called the complementary learning system (CLS; developed by McClelland, McNaughton, & O’Reilly, 1995).

1.1.2 Origins of memory consolidation theory: Clinical evidence

The later introduced proposition, such that a certain type of memory (which is subsequently recognised as declarative memory) is not just simply stabilised but evolves into a “reconstructed” form by integrating new information into pre-existing knowledge and changing corresponding neural structures with time, started based on findings in a series of amnesic patient studies (e.g. Scoville, & Milner, 1957). These clinical studies revealed that learning new individual words, autobiographical events and general facts, are impaired after hippocampus damage. In contrast, in these amnesic patients, old knowledge constructed before hippocampal damage was kept intact, even after the hippocampal damage. Besides this, these works revealed that the degree to which old memories were impaired was linked to the range of the damage around the hippocampus and extending areas (Squire, 1992), and that the new hippocampus-dependent memory of amnesic patients was intact, only holding within a short period of time (Baddeley & Warrington, 1970). According to these findings, it has been proposed that there is an intermediate shifting process that an initially hippocampus-dependent learning network, which temporarily represents initial learning for a short period of time, is gradually converted into a hippocampus-independent learning network, which integrates new learning into pre-existing knowledge represented at outside of hippocampal brain areas (i.e. the neocortex), which somehow consolidates newly acquired memory for a long period of time (Scoville et al. 1957). This is termed the ‘standard model’ of memory consolidation (see a review Nadel & Moscovitch, 1997).

1.1.3 The Complementary learning system (CLS) model

After the standard model of two neural learning networks was introduced, Marr (1970, 1971) proposed a model that suggests that neural representations of hippocampus-dependent learning are reactivated by the hippocampus during sleep, which eventually helps to consolidate its neural representations into the neocortex without support of the hippocampus. McClelland, McNaughton, & O'Reilly (1995) pointed out that these neural reactivations are caused offline (e.g. during sleep, imagery, and mental rehearsal) and/or online (e.g. a relevant task or cue to activate similar neural representations). McClelland, McNaughton, & O'Reilly (1995) suggested that the patterns of neural activations and reactivations and neural network strength developed from the neural reactivations can be theoretically modelled as two complementary learning systems (CLS). Accordingly, McClelland, McNaughton, & O'Reilly (1995) developed a computational model (the CLS model) with this structure in order to determine whether it captured the observed characteristics of memory.

Based on findings of the case studies mentioned above, the CLS model assumes that initial new learning is developed rapidly and transiently and represented in specific neocortical areas by integrating with hippocampus individually. As the representations are transient and individual, several immediate repetitions of activations in the hippocampo-neocortical network (hippocampus learning network) buffer the representations for a short period of time. Interestingly, learning in cortico-cortical network (neocortex learning network) is thought to develop gradually and interactively. In the CLS theory, representations of new learning in the neocortex network initially increase interference with representations of existing knowledge in the neocortex. However, by interleaving reactivations of existing knowledge and new

learning, new learning is eventually integrated into a neocortex-like learning structure. By applying these two assumptions, the CLS neuronal network learning model was able to become specialized for a specific task, for example, see the simulated neural network for learning concepts developed by Rumelhart (1995) where the CLS model examined simulated results in **Figure 1.2**. Results are changes of average network strength (i.e. absolute errors in the vertical axis of the two graphs in **Figure 1.2a** and **1.2b**) when multiple random activations (inputs) are applied to the simulated neural network. Absolute errors are determined by absolute differences between actual neural activations and estimated results. As estimated by the CLS model, new learning is facilitated by repeated training; existing knowledge is severely interfered by the repeated training, but relatively less so by interleaving new training (**Figure 1.2**). These results of the CLS model clearly show that acquisitions of new learning in the hippocampus seem to be developed rapidly by repeated neural (re)activations, but acquisitions of new learning in the neocortex develop slowly because neural activations need to be interleaving to avoid interfering neural representations of existing knowledge in the neocortex.

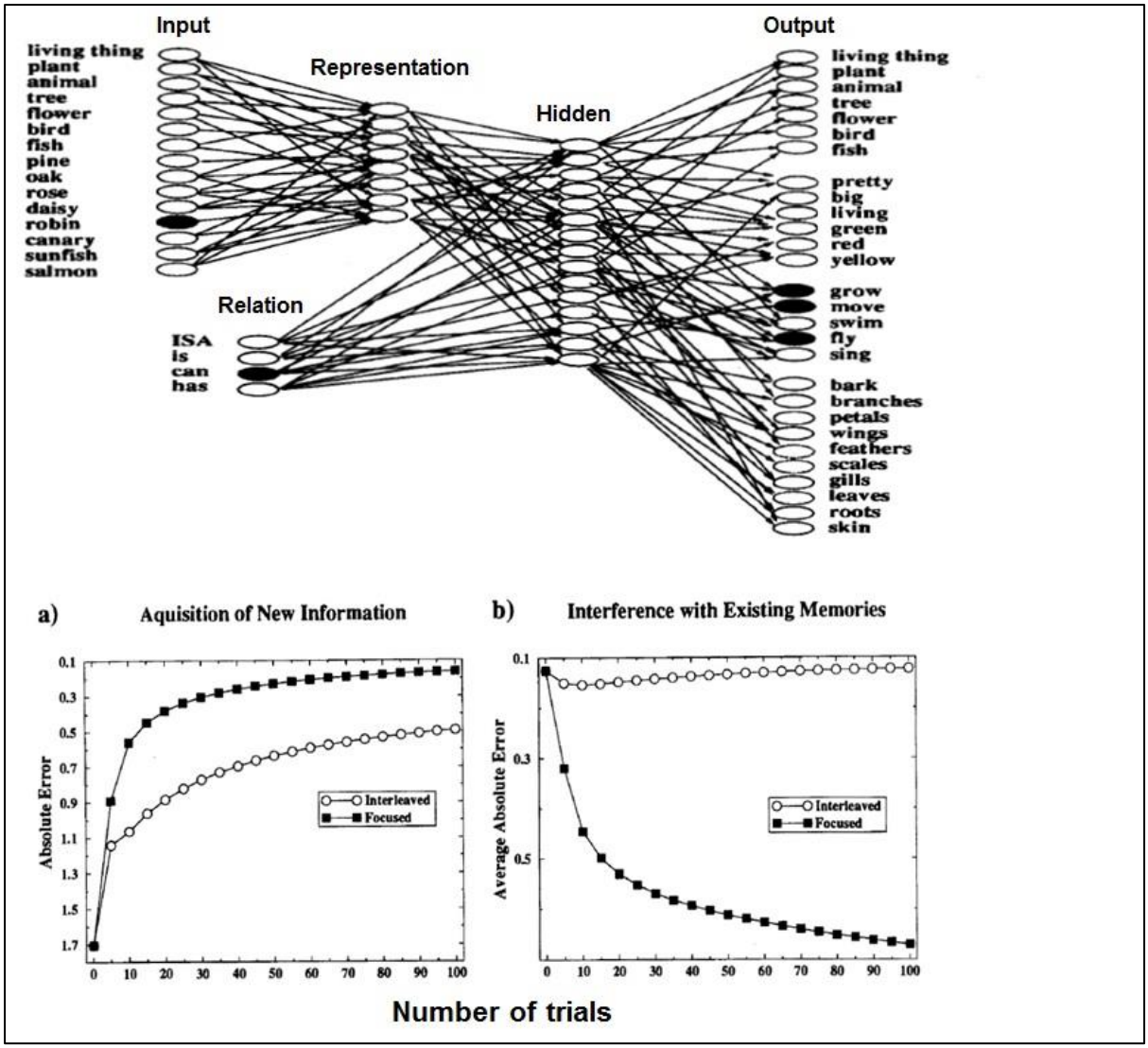

Figure 1.2. Simulated neural network changes in a concept learning task (top) and its results (bottom): At the top of **Figure 1.2**, the estimated neural structures relating to a concept learning task is simulated. On the left, lists are input and on the right, lists are the estimated output response for the corresponding input. Information of input is propagated from internal representations through a hidden environment towards output. Simulated neural connections across the path (i.e. lines) of propagation are activated. The two graphs at the bottom of **Figure 1.2** show the results for the concept learning task. The horizontal axis represents a number of trials. The vertical axis in 1.2a represents the absolute errors, which is a sum of absolute differences between estimated correct activations and actual activations. In Figure 1.2a, in which a higher (more errors) value means a higher interferences. Figure from McClelland, McNaughton, & O'Reilly (1995).

Section 1.2 Schemas and schema effects

1.2.1 Definition of a schema

A schema has been classically defined as generalized ideas of something, as a result of which has been encountered repeatedly in a form of everyday life events or through learning abstracted concepts, such as words and symbols (Piaget 1952; Bartlett, 1932; Mandler, 1984). For example, an (identity) schema of David Beckham would include him represented as a retired professional football player, who is married to the famous former singer Victoria Beckham, has several children, is friends with many celebrities and currently has a career on television and in magazines. It would also include information about his identity, as visually and genetically expressed by his ethnicity, race and nationality, as well as his appearance in his teens, 20s and 30s, his unique hairstyles, voice and tattoos. Even miniature size toy figures of him and drawings of his face, whose appearances may be rendered but have exaggerated features of his face, may be recognised as the identity of David Beckham. Nevertheless, it has to be noted that each of these specific characteristics of the David Beckham schema is not regarded as a schema in itself, but one specific characteristic of the schema. Instead, a schema is a generalized representation, which represents non-specific information, such as identities (e.g. David Beckham), concepts (e.g. a definition of a triangle as any polygon with three edges), facts (e.g. a category of the United Kingdom as an island country), or signs (e.g. ! as an exclamation mark).

Piaget (1952) proposed that a characteristic of schema is to assimilate new information into existing knowledge. What if David Beckham changes his career to be, a football team coach, or a baseball team coach? Because of pre-existing

knowledge about David Beckham as a previous professional football player, relevant new information (i.e. new career as a football team coach) is more likely to integrate into the existing knowledge than irrelevant new information (i.e. new career as a baseball team coach). This idea of schema boosting effects on the assimilation of relevant new learning into pre-existing knowledge has been referred to in recent findings mentioned below (e.g. Tse et al. 2007; Tse et al. 2011; Van Kesteren, Rijpkema, Ruiter & Fernandez, 2010; Van Kesteren et al. 2014). More specifically, these recent studies suggest that there are, at least, the following two conditions, either or both of which essentially induce the schema effects to facilitate new learning:

Condition 1. New learning of associations between two or more properties (e.g. learning to associate between knives, ovens, utensils, and aprons) that happen to match with context (in this example, kitchen) probably developed through real life experiences (see Van Kesteren, Fernandez, Norris & Hermans, 2010; Van Kesteren, et al., 2013)

Condition 2. New learning is closely related to pre-existing knowledge. (see Van Kesteren et al. 2014; Tse et al. 2007; Tse et al. 2011)

In addition, a review of schema studies suggests that schemas also promote false recognition, with participants falsely recognizing unstudied items when studied items relate contextually to the unstudied items (Head & Holmes 1911; Ghosh & Gilboa, 2014). For instance, when memorizing a collection of contextually matching objects (e.g. kitchen hob, sink, chairs, table, plates, kitchen tools and microwaves) in a

photograph one at a time, images of contextually matching but new items, such as an image of kettle, were falsely categorized as studied items the next day (Koutstaal & Schacter, 1997). This is probably because schema effects strengthen associations between the studied items and helps to develop a contextual frame (i.e. a context of kitchen in this example), which subsequently extends to link with unstudied but contextually matching objects (e.g. a kettle, a cup, a jar and oven) either consciously or unconsciously. This ability of schemas to build links between studied and unstudied items could cause an increase in false recognition. To support this idea, positive correlations between a number of contextually matching studied items at encoding and a number of false recognitions at test were found in this study (Koutstaal & Schacter, 1997); thus, more contextual matches may develop more widespread and strong links with contextually matching unstudied objects. An elderly age group also had a higher false recognition rate than a young age group probably because elderly people have more experience, allowing them to develop more such contexts. Based on this logic, schema effects have another feature which allows them to build contexts and links with related familiar properties, and the effects are plausibly determined by parameters of false recognition rates.

Furthermore, fMRI readings of human brains and experimental ablation methods on animal brains have revealed that these schema effects are closely linked with specific neural networks. Functional MRI studies by van Kesteren and her colleagues consistently found that the onset of schema effects and enhancements of new learning concurred with (1) increased connections between the medial prefrontal cortex (mPFC) and task-related cortical areas; (2) reduction of interactions between mPFC and hippocampus and (3) positive correlations between the dynamic

changes of mPFC (increase) and hippocampus (decrease) activity over a training period. In contrast, schema independent new learning was dependent on hippocampal activity (Van Kesteren, Rijpkema, Ruiter & Fernandez, 2010; Van Kesteren, Fernandez, Norris & Hermans, 2010; Kesteren et al., 2013, Van Kesteren et al., 2014). Summarising these human results, there seems to be two distinct neural structures involved in new learning: a hippocampus dependent learning, which interacts with and is eventually replaced by a mPFC dependent learning, which is closely associated with the onset of schema effects.

A rat lesion study supports the two neural learning networks and their relation with schema effects. Lesions on a rat's hippocampus impaired acquisition of new place-flavour learning when hippocampus was surgically removed before and immediately after the new training. However, acquired space-flavour learning was successfully spared when lesions were delayed to 48 hours after the new training (Tse et al., 2007). This result clearly shows that new learning is, at first, stored in hippocampus, but the learning is ultimately transferred to another brain area within a delayed period (e.g. 48 hours).

Further, the result supports the idea that second place-flavour learning was quickly acquired only on the following day after the first place-flavour training was acquired. This suggests that knowledge of previous training stored outside of hippocampus critically links with onsets of schema effects. Another rat study reported that this schema effect is linked with genetic development (i.e. upgrading gene expressions) at mPFC (Tse et al., 2011). Therefore, both human and animal studies coherently show that learning is developed through two learning stages (i.e. the first stage:

based in the hippocampus learning network, then the second stage: based in mPFC learning network) and schema effects are induced when related new learning is directly assimilated to existing knowledge that has already been constructed in the second stage in the neural learning networks.

In summary, a schema is defined as a mentally and neurally constructed theme or gist of previous experience, which functions to promote assimilations of related new learning into existing knowledge and increase false recognitions. Based on the results of brain imaging and animal studies, the schema effect actually promotes direct integration of new information into the already constructed neural networks. However, this schema effect does not clearly demonstrate how the direct integration of 'related new learning' into 'existing knowledge' leads to facilitating effects in the enhancement of new learning, such as better recognition rates, rapid acquisition, and consolidation of the new learning, compared to enhancement of 'unrelated new learning'.

1.2.2 Refinement of CLS theory and applicability to the schema effect

According to the original CLS theory (McClelland, 1995), learning systems are classified into quick (i.e. hippocampus learning network) and slow (i.e. neocortex learning network) learning systems. Yet, recent works by Tse et al. (2007, 2011) show that neocortex learning network can be learned rapidly. Accordingly, McClelland (2013) tested whether CLS simulations match with their actual results adding considerations of schema-related new learning. Consistent with Tse (2007, 2011), the simulation results showed that with schema effects, facilitation of new

learning and reduction of interference with existing knowledge, were found only when the new learning was related to prior knowledge (see **Figure 1.3**). Therefore, McClelland (2013) has amended the theory of CLS by noting that features of slow changes in neocortex learning networks should be paraphrased with dependency on existing knowledge, but other principal features of CLS should remain unaltered. This finding of fast acquisition of new learning in the neocortex learning network, when new learning is related to existing knowledge, explains mechanisms of schema effects and facilitations of new learning by pre-existing knowledge. Simulations of the refined CLS model have shown that new learning related to pre-existing knowledge is more quickly acquired, interferes less with existing knowledge, and more successfully consolidates into the neocortex-like learning network compared to unrelated new learning (**Figure 1.3**). Accordingly, memory consolidation theory can be refined as well to have special patterns of schema-related quick learning acquisitions with unchanged two neural learning networks.

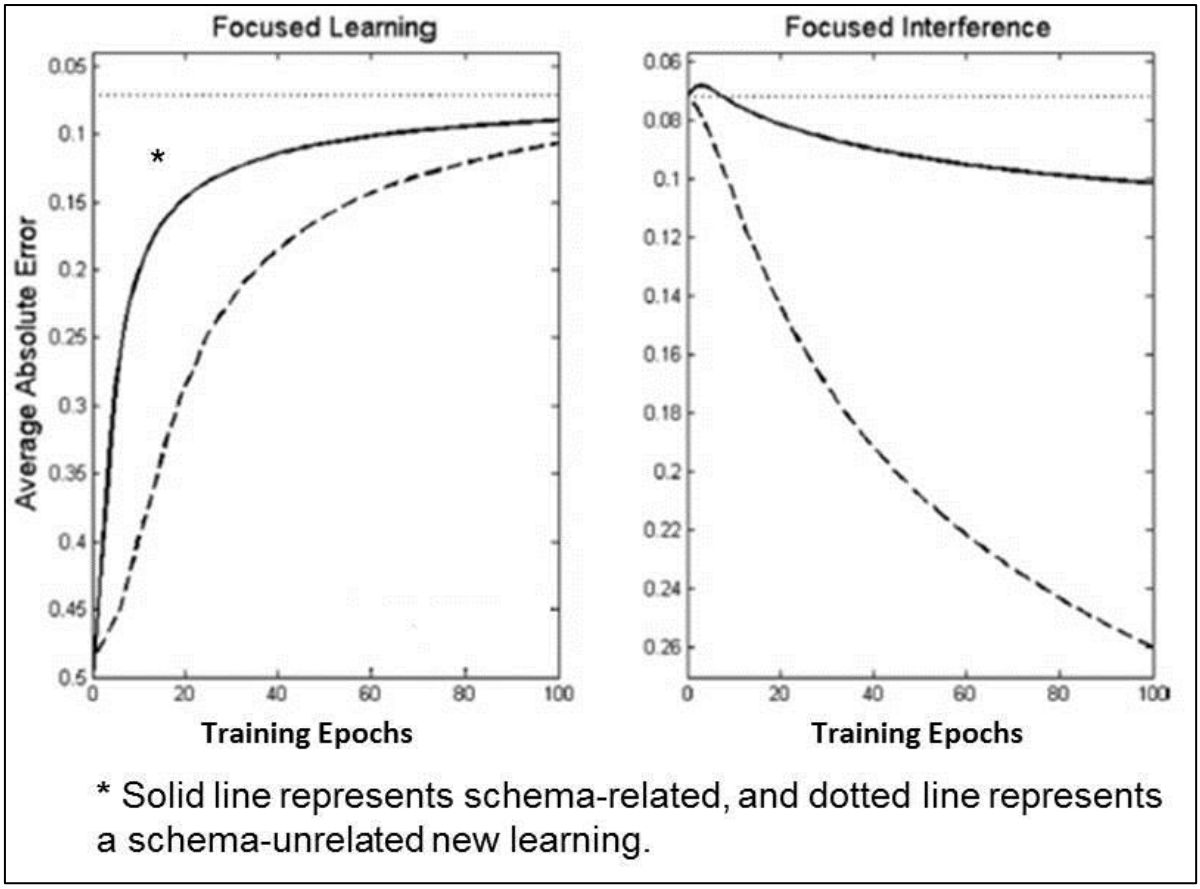

Figure 1.3. **Results after simulations for learning (Left) and Interference (Right):** The horizontal axis represents the number of training trials. The vertical axis represents the average absolute errors (estimated – actual network strength) per training trial. The lower the error scores the better learning performance is. Figure from McClelland (2013).

In short, memory consolidation theory describes how new learning is acquired and consolidated into the two types of neural learning network and how facilitations of new learning by pre-existing knowledge are induced. At the same time, this theory introduces the concept of neural reactivations during sleep for consolidating new learning.

Section 1.3 Sleep and learning

The standard model of the memory consolidation theory (Frankland, & Bontempi, 2005; Marr 1970, 1971; Nadel & Moscovitch, 1997) suggests that offline learning-related neural reactivations, potentially during post-learning sleep, are critical for consolidation of hippocampus-dependent learning. Yet, the standard model has not specified which aspects of sleep facilitate memory consolidation and how the mechanism of sleep-dependent memory consolidation works. Thus, this section will review architecture of sleep patterns and sleep-related neural activity, and then review the active systems consolidation theory (e.g. Diekelmann & Born, 2010) that proposes how learning-related neural reactivations (i.e. neural replays) at specific sleep stages can enhance hippocampus-dependent memory consolidation. In addition, I will introduce another theory, the synaptic homeostasis hypothesis (Tononi, & Cirelli, 2006), which advocates another possible function of sleep that is to boost a process of hippocampus-dependent memory consolidation through globally downscaling the strength of neural connectivity to mute irrelevant neural connectivity (noises) and improve signal (acquired knowledge) to noise (irrelevant information) ratio during specific sleep stages. Finally I will introduce a model of information overlap to abstract (iOtA) proposed by the work of Lewis and Durrant (2011), which suggests how those two models can actually be integrated each other to explain the mechanism of sleep-dependent memory consolidation.

1.3.1 Taxonomy of sleep structures

The major classification of sleep structure has been widely recognized since the work of Rechtschaffen and Kales (1968). With the multi-channel electrophysiological recording of polysomnography (PSG), which measures electrophysiological (i.e.

EEG) activity on the scalp and eyes and jaw movement in electrooculography (EOG) and electromyography (EMG), respectively, Rechtschaffen and Kales (1968) showed that sleep can be classified into two major stages: rapid eye movement (REM) sleep and Non-REM (NREM) sleep, which is further divided into sleep stage 1 (N1), stage 2 (N2) and stage 3 and 4, which is later on recognised as slow wave sleep (SWS). (Actual distinction between the REM and NREM was initially recognized by Kleitman and Aserinsky, 1953). REM sleep is characterised by the appearance of rapid eye movements, ponto-geniculo-occipital (PGO) wave and theta activity; N1 is a transitional stage between waking and sleeping; N2 is characterised by appearance of K-complex (irregular high amplitude sharp oscillation) and sleep spindle activity (burst of rapid oscillations within a short period with a fixed amplitude); SWS is evaluated as deep sleep as slow large oscillating wave is mainly observed in this period (see **Figure 1.4-b**). Sleep is cycled in 90 minutes in a relatively fixed order from N1 to SWS, then to REM sleep then back to shallow sleep and so on (see **Figure 1.4-a**).

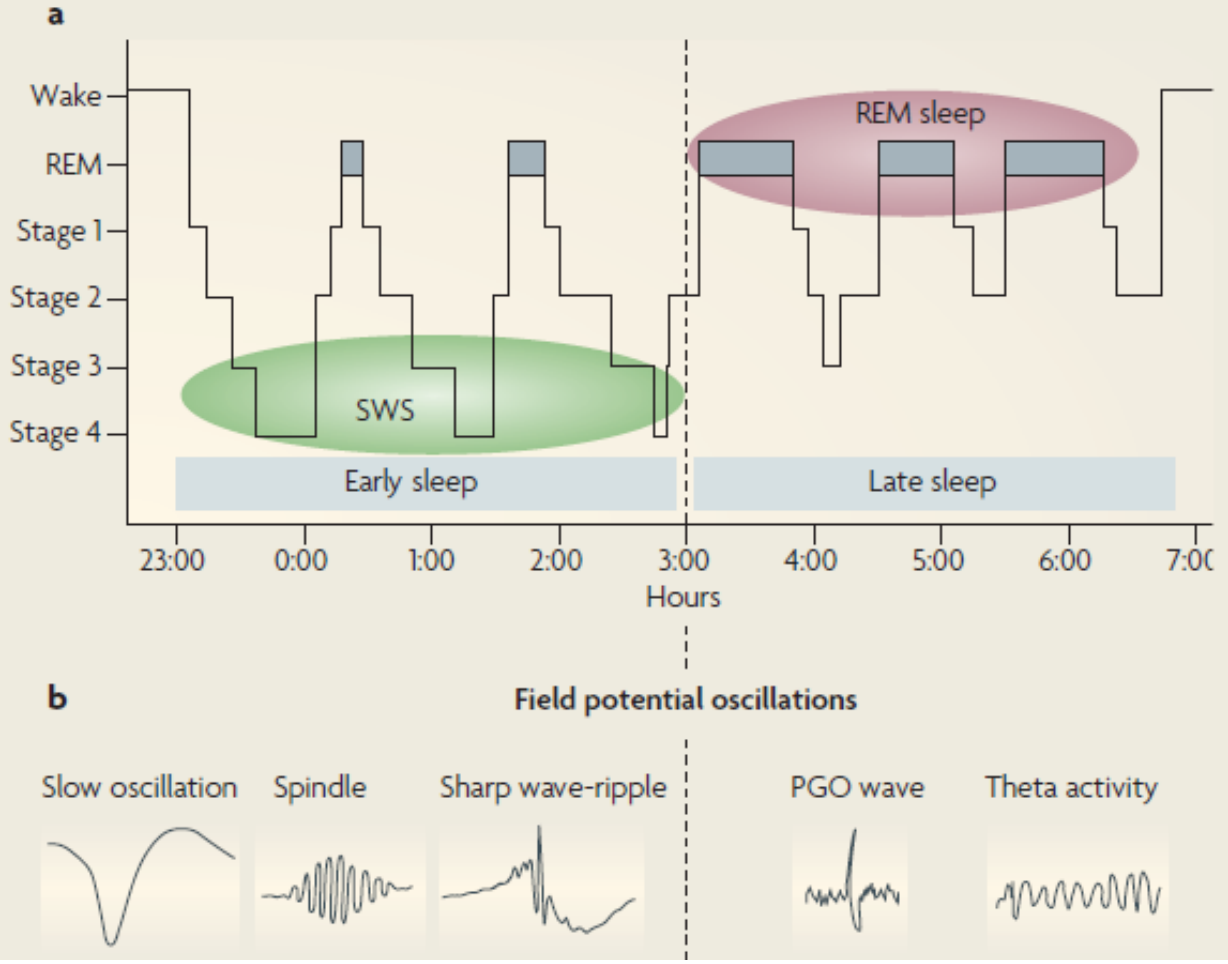

Figure 1.4. (a) Hypnogram (graph of sleep stage changes over a full night) and (b) types of sleep EEG oscillations. In the hypnogram, the horizontal axis is a time course across a full night; vertical axis indicates onsets of sleep stages. The hypnogram illustrates which sleep sage will be often seen in a specific time. 4-b shows specimen of specific sleep EEG oscillations. Figure from Diekelmann and Born (2010).

1.3.2 The neurophysiology of sleep

Animal and human studies with intracranial electrophysiological recording techniques suggest that specific sleep EEG activity originates from specific brain areas. For example, slow oscillations (SO) originate in the neocortex (Contreras & Steriade, 1995), sharp wave ripples (SWR) in the hippocampus (Eschenko et al. 2008) and spindles in the thalamus (De Gennaro & Ferrara, 2003). In detail, SWR are short-lived fast oscillations (around 200 Hz) synchronizing within the hippocampus

(especially CA1-CA3 areas of dentate gyrus), induced by intercellular depolarization (positive ward membrane potential change) that occur during wakefulness and NREM sleep (Buzsáki, et al. 1992). Sleep spindle oscillations (11-16 Hz) occur in the thalamic reticular (RE) neurons (surrounding around thalamus, Scholpp et al. 2009; Iber, Ancoli-Israel, Chesson, and Quan, 2007), induced by excitatory potential from the brainstem and corticothalamic (CT) neurons (see reviews, De Gennaro & Ferrara, 2003; Steriade 2005). RE neurons carry an inhibitory synaptic potential projecting inhibitory GABA neurotransmitters to thalamocortical (TC) neurons to hyperpolarise (negative ward membrane potential change) TC neurons; this results in a reduced responsiveness of TC neurons to stimuli of environment, while the TC neurons subsequently rebounds to depolarise and delivers excitatory discharge to cortical areas (see reviews, De Gennaro & Ferrara, 2003; Destexhe, & Sejnowski, 2003; Steriade 2005). SOs (<1 Hz) are cyclic changes of long-lasting intercellular depolarization (excitation) and hyperpolarisation (inhibition) synchronizing throughout the neocortex during NREM sleep (Steriade, McCormick, & Sejnowski, 1993; Sanchez-Vives & McCormick, 2000). As the neocortex receives electrophysiological inputs from subcortical areas, especially from the thalamus and hippocampus, the upward and downward oscillation of the neocortical SO is thought to be a reflection of globally synchronized depolarization and hyperpolarisation of the hippocampo-thalamo-neocortical neural network (Mölle, Marshall, Gais & Born, 2002; Mölle et al. 2006; Clemens et al. 2007). Hippocampal SWR, thalamic spindles and neocortical SO propagate their excitatory discharge to each other and strengthen their neural connectivity between them (e.g. Buzsáki 1996; Contreras, Timofeev & Steriade, 1996; Sirota, Csicsvari, Buhl & Buzsaki, 2003; Taxidis, Mizuseki, Mason, & Owen, 2013). Frequent exchanges of neural discharges and concurrent modulation of

neural connectivity are closely linked with average neural network weight simulated in the CLS described above. Consequently, finding onsets of specific sleep EEG oscillations (i.e. SWR, spindle and SO) during NREM sleep are good cues for estimating changes in neural network strength and related changes in learning.

1.3.3 A brief history of sleep role in learning

In parallel with studies on sleep neurophysiology, sleep's role on hippocampus-dependent memory consolidation has also been investigated. Since the work of Jenkins and Dallenbach (1924), surprisingly, there has been a consensus that sleep benefits the consolidation (strengthening) of hippocampus-dependent learning compared to an equivalent amount of wakefulness. For instance, Jenkins and Dallenbach (1924) employed a task involving learning nonsense syllables before a retention interval of sleep or wakefulness. They showed that those who slept after the task performed better on the memory retention than those who stayed awake. Similarly, this sleep advantage was seen in other behavioural tasks such as a word association task in Gais, Hüllemann, Hallschmid, & Born, (2006), spatial learning in Peigneux et al. (2004) and visual discrimination in Stickgold et al. (2000). These sleep advantages are closely associated with SWS and REM. Yet, the sleep advantage has only benefitted maintenance of acquired learning over time compared to deterioration in a non-sleep condition. Thus, there is a debate about whether sleep's role is mere passively protecting from interference and decay or actively improving learning (see a debate in Ellenbogen, Payne & Stickgold, 2006).

1.3.4 The active systems consolidation theory

One influential theory about how sleep influences memories is the Active Systems Consolidation Theory (ASC) (see a review of Marshall & Born, 2007). This theory draws heavily on the Standard Model of Systems Consolidation in combination with the observation that recently learned memories are actively reinstated or 'replayed' by the brain during sleep to explain how memories are altered across sleep.

Importantly, several studies of hippocampal place cells show that patterns of learning-related neural reactivation during wakefulness (e.g. three distinct neural locations: neural space A, B and C in this order) is reactivated during SWS, as one experiencing the activity during wakefulness (Pavlides & Winson, 1989; Wilson & McNaughton, 1994; Nadasdy et al. 1999; Ji & Wilson, 2007; Euston, Tatsuno & McNaughton, 2007; Lansink et al. 2008). With the evidence, the role of neural reactivation for memory consolidation in the standard memory consolidation theory and knowledge of sleep neurophysiology is established. The ASC claims that reactivation of learning-related neural activity during NREM sleep (i.e. SWS) helps to transform initially fragile learning into long-lasting learning (see reviews in Rasch & Born, 2008; Diekelmann & Born, 2010; Rasch and Born 2013). As an example of this proposed phenomenon, during initial spatial learning in rats, specific cells in the hippocampus fire vigorously (e.g., Suzuki 2006). Even after the spatial learning, the learning-related specific cell firing in the hippocampus is repeatedly replayed during NREM sleep (e.g., Wilson & McNaughton, 1994). This neural reactivation in the hippocampus also associates with acquisition of navigation tasks in virtual reality (Peigneux et al., 2004). This neural replay during a hippocampus-dependent task also synchronizes with the onset of hippocampal SWR activity during NREM sleep.

Disruption of this fast oscillation during learning corresponds to deficits of such learning (Girardeau et al. 2009; Ego-Stengel & Wilson, 2010; Jadhav, Kemere, German & Frank, 2012). In addition, synchronization between repeated neural replay of learning and hippocampal SWR is believed to play a key role in the facilitation of consolidating new learning into the neocortex because hippocampal SWR closely synchronizes with and actively exchanges synaptic discharge with neocortical SO during NREM sleep (Buzsáki 1996; Contreras, Timofeev & Steriade, 1996; Mölle, Marshall, Gais & Born, 2002; Sirota, Csicsvari, Buhl & Buzsaki, 2003; Buhl & Buzsáki, 2005; Mölle et al. 2006; Clemens et al. 2007; Taxidis, Mizuseki, Mason, & Owen, 2013). If there is repeated neural reactivation and communications between the hippocampus and neocortex strengthen during sleep, this is equivalent to the model of memory consolidation (Frankland, & Bontempi, 2005). Learning-related neural replay is also seen in the thalamus and neocortex during NREM sleep (Ji & Wilson, 2007; Lansink et al. 2008), which indicates that the neural replay is successfully shared among hippocampo-neocortico-thalamus networks guided by SWR-SO-spindles communication (Buzsáki 1996; Contreras, Timofeev & Steriade, 1996; Sirota, Csicsvari, Buhl & Buzsaki, 2003; Buhl & Buzsáki, 2005; Taxidis, Mizuseki, Mason, & Owen, 2013). Consequently, repeatedly reactivating and transferring encoded memory from the hippocampus to the neocortex through SWR-SO synchronization in the hippocampo-neocortial neural network is thought to be the mechanism for enhancing declarative learning during sleep (Marr, 1970, 1971; Buzsaki, 1989; McClelland, McNaughton, & O'Reilly, 1995; Gais & Born, 2004; Girardeau et al., 2011; Genzel et al., 2014).

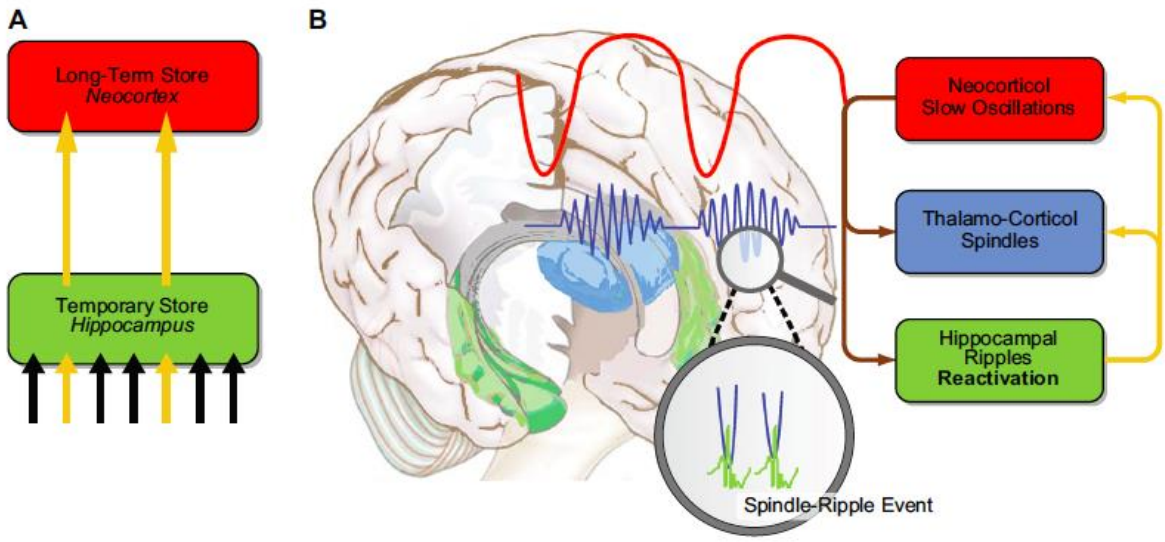

Figure 1.5. The active systems consolidation model. In the active theory, a process of memory consolidation is induced through repeated closed loop learning-related neural reactivations in loop in neural circuit between hippocampus, thalamus and neocortex. Figure from Rasch and Born (2013).

The enhancement of declarative learning by neural replay during NREM sleep has only been supported indirectly to date. For example, stronger hippocampus activation and increased learning are enhanced when a learning-related cue is administered during NREM (but not REM) sleep (Rasch, Buchel, Gais, & Born, 2007). This cue administration during NREM sleep accelerates the transfer of encoded hippocampus-dependent memory into the neocortex. Also, achievement of consolidation of new learning into neocortex during sleep is indicated by the connectivity reduction in the hippocampo-neocortical network and an increase in the neural connectivity within the neocortex during post-sleep waking (Takashima et al. 2009; Van Kesteren, Fernandez, Norris & Hermans, 2010; Van Kesteren et al. 2013).

Through the process of repeated reactivations during sleep, it is conceivable that there will be an overlap between related memories that have been reactivated at the

same time, or close together in time. This overlapping reactivation in the same cortical area will develop more stabilization of the shared area and thus lead to integration between or generalization of those memories guided by the well stabilized cues at the overlapped area (see **Figure 1.6** for schematic representation of this process). This integration process has been described in the information overlap to abstract (iOtA) model, which suggests why the onset of false memory and insightful ideas, memory combination, abstraction, and generalization are promoted during sleep as a side effect of the sleep-enhanced memory consolidation (see Lewis & Durrant, 2011; Stickgold, & Walker, 2013).

1.3.5 Synaptic homeostasis hypothesis (SHY)

In contrast to the active systems consolidation theory, the synaptic homeostasis hypothesis (SHY) suggests that sleep is a mere state which protects memories from forgetting by shutting down brain responses to external interference from the outside world and by downscaling encoded memories to filter noises and increase signal-to-noise-ratio (SNR) of acquired knowledge (Tononi, & Cirelli, 2006). There have been several reports that sleep maintains the quality of memories by protecting them from interference (Ellenbogen, Payne & Stickgold 2006; Ellenbogen, Hulbert, Jiang & Stickgold, 2009; Sheth, Varghese & Truong, 2012). The protective role of sleep in learning is also consistent with sleep disorder/deprivation studies where disturbance of sleep, for instance by sleep insomnia and sleep deprivation, causes loss or deterioration of memory (Grosvenor & Lack, 1984; Harrison & Espelid, 2004; Silvestri 2005; Palchykova et al. 2006; Griessenberger et al. 2013). One study reported that sleep deprivation even disrupts long term potentiation (LTP), (Kim,

Mahmoud, & Grover, 2005), which is a long-lasting synaptic plasticity and is thought of as a parameter of learning/memory at the neuronal level (see a review Bliss & Cooke, 2011). Based on a detailed understanding of the neurophysiological development during sleep (Tononi, & Cirelli, 2006; Spoormaker et al. 2010; Andrade et al. 2011; Bushey, Tononi & Cirelli, 2011; Van Alphen et al. 2013; Xie et al. 2013), such evidence supports the idea that sleep, especially SWS, is least likely to be influenced by interference of internal and external information. Information is maintained in local brain areas, unnecessary information is cleared, and no other information is added from the outside world (Tononi, & Cirelli, 2006; Spoormaker et al. 2010; Andrade et al. 2011; Alphen et al. 2013; Xie et al. 2013).

Moreover, a recent theory proposed that this sleep protection also plays a role in the integration and generalization of encoded memories by employing hierarchical downscaling (Lewis & Durrant, 2011; Hashmi, Nere & Tononi, 2013; Nere, Hashmi, Cirelli, & Tononi, 2013). When newly encoded memories share congruent information with consolidated prior memories, they share connectivity and net synaptic strength of the congruent information in the same topographic map; this results in greater connectivity and net synaptic strength of the overlapped information than with non-overlapping information. This means that congruent information is preserved whilst incongruent information is downscaled (Hashmi, Nere & Tononi, 2013; Nere, Hashmi, Cirelli, & Tononi, 2013). Based on this theory, SNR of surviving integrated and normalized information after downscaling is increased.

1.3.6 The function of sleep beyond mere memory strengthening

The role of sleep is not limited to the strengthening of learning which has been learned previously, but also to variety of tasks and memory performance beyond that. For example, sleep promotes learning of statistical patterns in visual (Djonlagic et al. 2009) and auditory modalities (Durrant, Cairney & Lewis, 2011), understanding grammatical rules (Nieuwenhuis et al. 2013), hierarchical orders between stimuli (Ellenbogen et al. 2007; Werchan & Gómez, 2013), and abstracting series of specific conceptual learning into general concepts (Lau, Alger & Fishbein, 2011). Based on these findings, sleep seems to arrange single experiences into generally coherent structures, which helps develop contexts or meaningfulness in a collection of experiences. To support such a role for sleep in learning, for example, learning a general concept of "female" in Chinese characters was acquired from learning individual Chinese characters of "sister," "mother" and "wife" (Lau, Alger & Fishbein, 2011). Sleep also enhances to integrate new learning into existing knowledge (Tamminen et al., 2010; Tamminen, Ralph & Lewis, 2013), finds hidden rules from unsupervised statistical pattern training (Wagner et al., 2004), and increases false recollections (McKeon, Pace-Schott & Spencer, 2012). Even after building such contexts in previous learning, the constructed contexts keep being modified externally (e.g. assimilating new contexts into old knowledge) and internally (e.g. strengthening contextual relations and increasing false recollections to unstudied but contextually fitting cues) by sleep.

In addition, sleep plays a role beyond strengthening of hippocampus-dependent learning that depends on specific sleep stages. For instance, NREM sleep, especially spindle activity, which is mainly seen in N2, is closely linked with improved associations between pre-existing knowledge and new learning (e.g. Groch et al.,

2017; Hennies, et al., 2016; Tamminen, et al., 2010) and this spindle-dependent association improvement is particularly associated with reduction of hippocampal activity (Hennies, et al., 2016). Another NREM sleep type, SWS (deep sleep), is important for abstraction (Durrant, Cairney & Lewis, 2013; Lau, Tucker & Fishbein, 2010), transformation of implicit to explicit knowledge (Cousins, et al., 2014; Diekelmann, Born & Rasch, 2016), gist extraction (Konrad, Herbert, Schneider & Seehagen, 2016; Lutz, et al., 2017), and acquisition of statistical patterns (Durrant, Cairney & Lewis, 2016; Durrant, Taylor, Cairney & Lewis, 2011). Classically, REM sleep has been known as a sleep stage that is key to enhancing procedural and emotional memories (see a review Genzel, Spoormaker, Konrad & Dresler, 2015; Rasch & Born, 2013), but recent reviews have suggested that this sleep stage may instead be important for amygdala-related memories (Genzel, Spoormaker, Konrad & Dresler, 2015) and it has also been suggested to be critical for memory reorganisation, in order to make a space for subsequent new learning after a sleep (Poe, 2017). Yet, this REM sleep dependent memory reorganisation requires more elaboration of the ideas and supporting evidences.

Putting all of this together, NREM sleep, especially SWS and maybe in combination with REM sleep, seems to play a critical role for reorganisation of overall knowledge.

1.3.7 iOtA model

The Information overlap to abstract (iOtA) model (Lewis & Durrant, 2011; see **Figure 1.6**) was proposed by integrating ideas of the ASC model and the SHY to explain the versatile consequences of sleep-dependent memory consolidation introduced above.

This model suggests that schema effects and the facilitation of new learning are probably induced because overlapping areas develop relatively strong connections. Based on a combination of this model and growing knowledge of schema-effects, it seems plausible that new learning which falls into the overlapping area would be learned quickly thanks to the strong relations within the area developed by overlapping. Still, it is unclear how to represent this overlapping area. Word (or semantic) cues may be problematic since it is difficult to represent words in the same two-dimensional space. Rather, mental representations of word relations are arbitrary and strongly influenced by individual differences (e.g. Cain, Oakhill & Lemmon, 2004; Mülberger, 2017; St George, Mannes & Hoffman, 1997). For example, when thinking of a relationship between words of “leisure,” ‘outside’ and ‘people,’ one can infer a context of either ‘picnic,’ ‘shopping,’ ‘traveling,’ ‘watching or playing football’ depending on the person’s perspective. Based on this perspective, relations between those three words can be varied. Therefore, some types of visual cues, which can be systematically represented in a two-dimensional space, are required to be introduced to test ideas of the iOtA model.

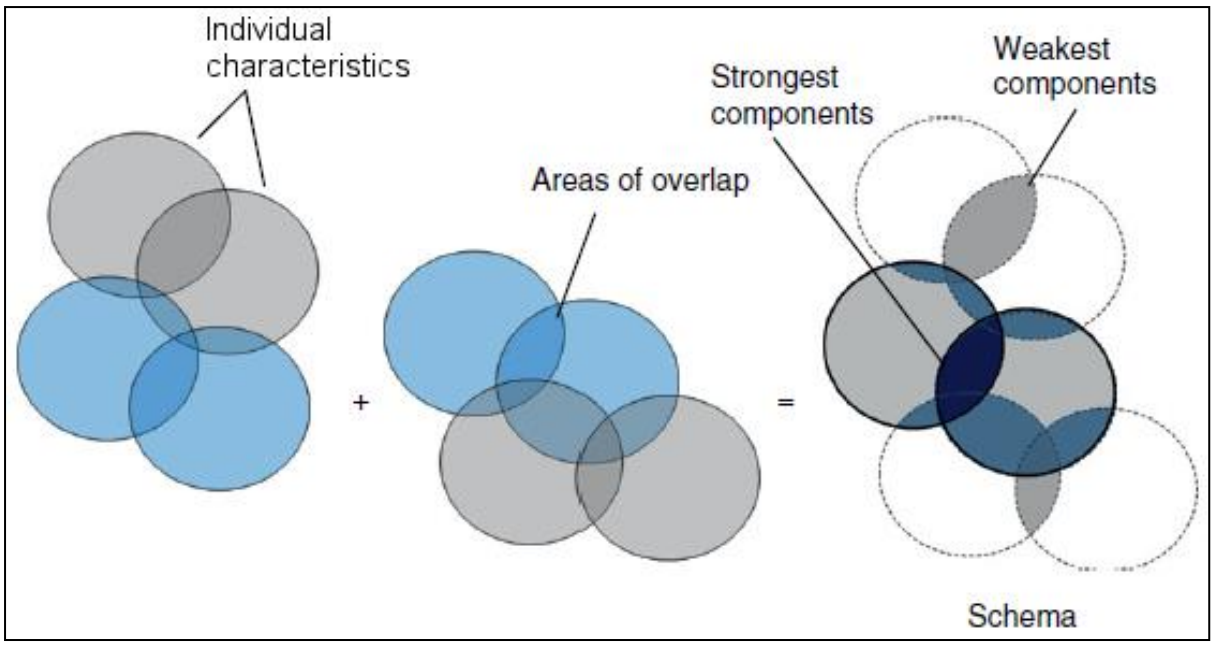

Figure 1.6. Information overlap to abstract (iOtA): Each circle represents individual characteristics essential for representing a schema. To form a schema, these characteristics need to be close enough to overlap (interrelate). Through the overlapped space, all characteristics are linked each other to form an associative network to represent a schema (Lewis & Durrant, 2011). Because of the overlapping patterns, there can be the strongest (the most overlapping) and weakest (the least overlapping) schema components. Figure from (Lewis & Durrant, 2011).

Section 1.4 Integration of all the above theories

Schema effects, CLS model in the memory consolidation theory, the Active Systems Consolidation theory, the iOtA model, and the synaptic homeostasis hypothesis are all closely linked to the establishment of interconnections between different memories. Schema effects and the facilitation of new learning occur when new learning relates to existing knowledge. The CLS model demonstrates simulated schema effects in which new neural activation relating to an existing neural network quickly integrates into the existing network. The iOtA model suggests that the overlapping of neural reactivations between neural representations facilitates

consolidation of new learning. Within the synaptic homeostasis hypothesis, the concept of hierarchical downscaling illustrates that sleep preferentially preserves overlapped neural representations by specifically downscaling, to filter irrelevant noise neural activity. Yet, although related, these theories do not explain clearly what kind of relations and overlaps are vital for those proposed theories. Specifically, existing research has not yet examined:

- How much detail is sufficient to represent the relations/overlaps between visual cues? This question is interesting both behaviourally and neurally, and in consideration of both densely overlapped areas and sparsely overlapped areas.
- How do neural or behavioural memory representations of these overlaps alter across a retention interval of wake or sleep?

Interestingly, a rat's place cell firing patterns corresponding to a real space map (**Figure 1.7**, O'Neill, Pleydell-Bouverie, Dupret, & Csicsvari, 2010) hints a possible method to track overlapped neural areas related with established learning relations in behavioural terms. Each place cell fires only when the rat runs through a specific area (e.g. cell number 16 and 5 firing at the central area in the actual space map). Also, it is apparent in this figure that each shape of firing pattern of place cell (i.e. the shape of small circles in the blue real space map) in a specific cell number is different each other. That is to say, depending on where on the real space map behavioural activity is focused (e.g. running on a central area of the real space map), only corresponding neural activity (activating cell number 16 and 5) will be detected, which implies those neural activities are probably overlapped (in this case cell number 16 and 5) because both neural areas are activated at the same time for a

specific behaviour. In this way, it is possible to say which neural area is probably overlapped when learning is overlapped behaviourally.

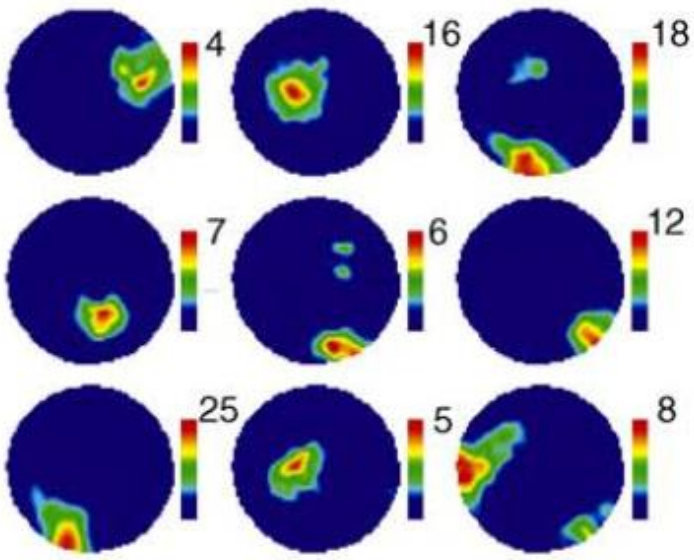

Figure 1.7. Mean Firing Pattern of Place Cells in Real Space Map. Real space map and each rat's place cell firing at a specific brain area. Red indicates higher cell activity. Blue indicates lower activity. The number on top right of each real space map indicates a specific place cell number. Figure from (O'Neill, Pleydell-Bouverie, Dupret, & Csicsvari, 2010).

Based on the above, I hypothesize that using a carefully designed declarative learning task it will be possible to construct areas of densely overlapped and sparsely overlapped learning. In addition, I expect that sleep would influence these representations, and the way new knowledge is integrated into both dense and sparse memory spaces.

1.4.1 Aims

As there are few research paradigms to systematically represent semantic or episodic memory figuratively in space, I one of my first aims was to develop a new

paradigm to achieve this. Recent technology allows for the gradual morphing of face images along specific families of features such as age (young to old), gender (feminine to masculine) and race (African to Asian to Caucasian). Thus, by performing age and gender morphing in x- and y-axis, I was able to build up a two-dimensional facial morph space (**Figure 1.8 & 1.9**). By using this this 2D face morphing face space in combination with a face learning task, I was able to determine how multiple episodes of face learning in this 2D space combine to form a more general knowledge of studied faces. I can then test how sleep influences the shape of the knowledge map in this 2D space. Please see Chapter 2 for more detail on this paradigm.

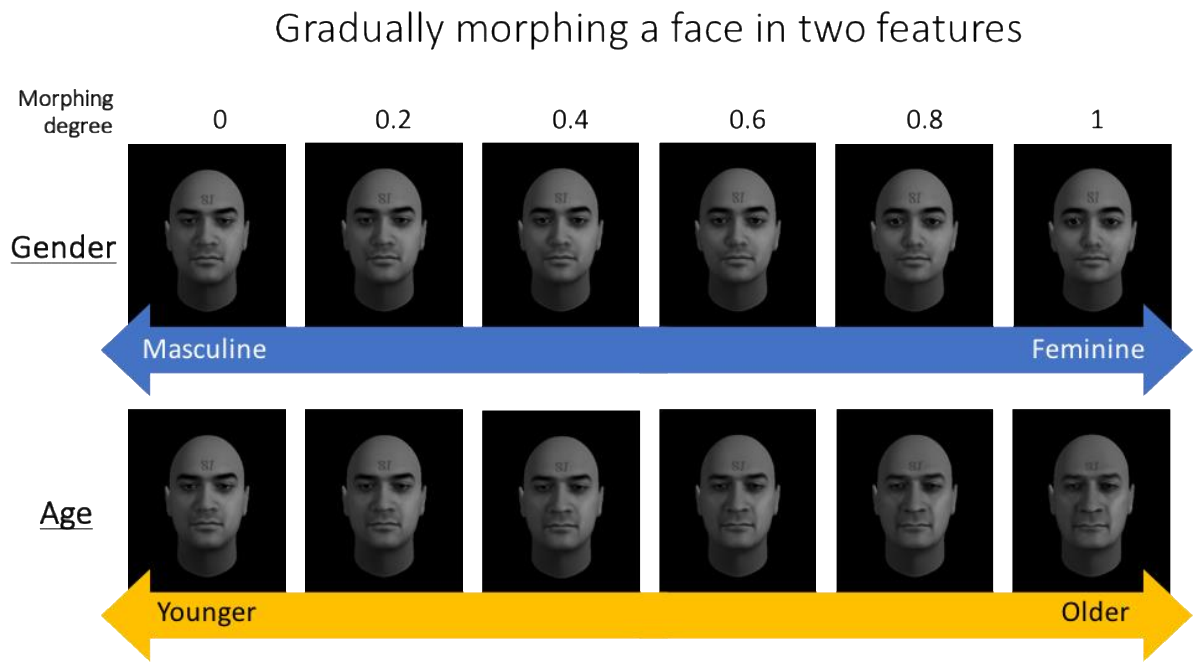

Figure 1.8. Gradual Face Morphing Scale. Specific aspects (i.e. age and gender) of a face is morphed across one dimension separately. The assigned end of morphed face continuum is defined as image 0 (in this case the most masculine and young face as 0) and the other end as 1 (in this case the most feminine and old face as 1).

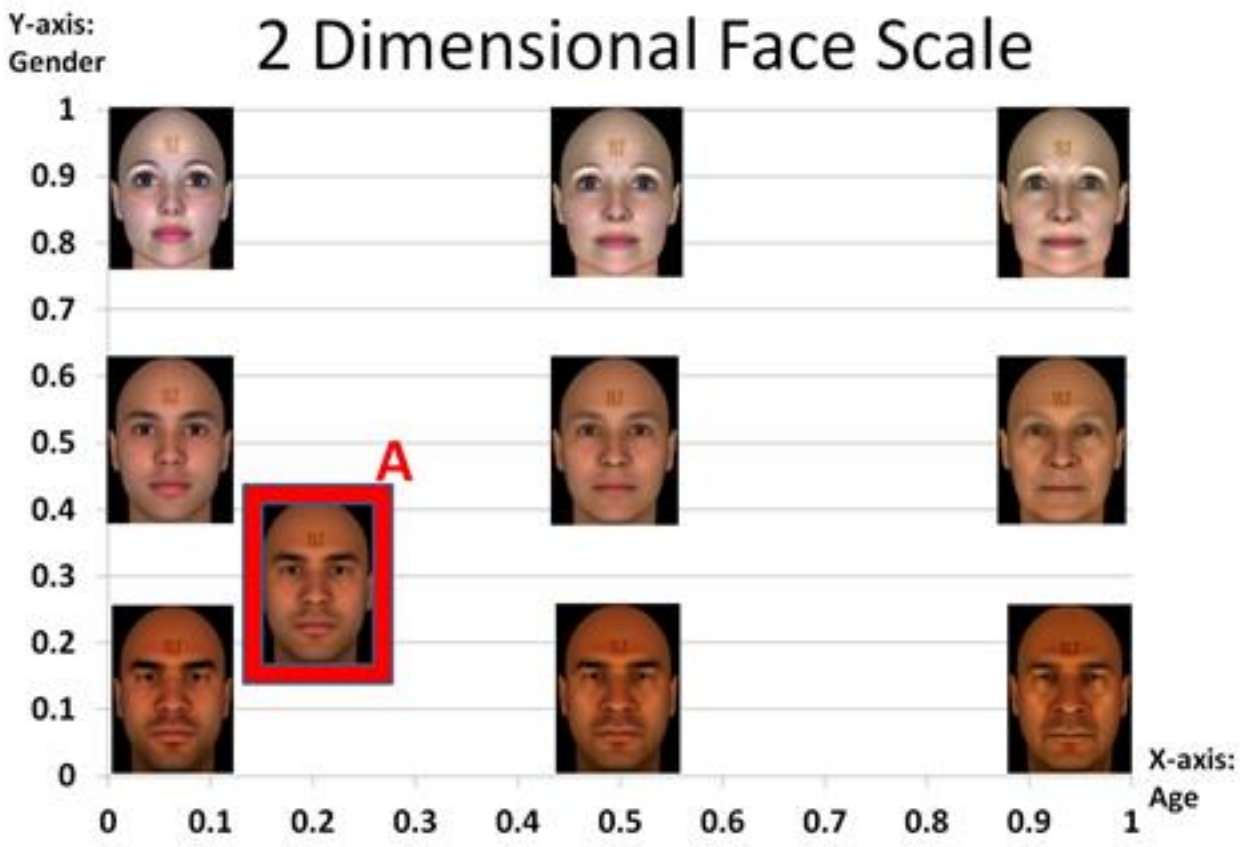

Figure 1.9. 2-Dimensional Face Scale. Each facial feature is assigned in to horizontal and vertical axis, which creates a two-dimensional scale.

By employing this new face space paradigm (**Figure 1.8 & 1.9**), I hope to understand how the shape, size, and density (in the 2D face space) of learned face representations are developed across sleep, and to determine which neural activity and sleep stages play a role for constructing these spaces.

My specific research questions are:

(1) Determine how the density of learned stimuli represented in this 2D-space influences the shape of the resulting knowledge space within the 2D map.

(2) Determine how sleep alters the shape and size of such knowledge space.

(3) Determine which neural activity and sleep stage relates to constructed knowledge space and updating.

Chapter 2

Investigating the role of sleep on the gist-based face recognition memory

Hikaru Tsujimura, University of Manchester, UK

Penelope A. Lewis, University of Manchester, UK

2.1 Abstract

Memories generalise across sleep. One framework, the Information overlap to abstract (iOtA) model, suggests that abstraction across sleep occurs as a result of overlapping replay of memories followed by downscaling. If this is correct, then greater degrees of overlap should be associated with a higher probability of abstraction. We tested this proposal using a face-morphing paradigm. In this paradigm, a face of an actor was morphed in age and gender and those morphing scales were assigned to a vertical and a horizontal axis, as a result, inventing a 2D face morph scale. Then, participants were trained on a series of facial images selected from the invented 2D scale. These learned images were either close together (densely packed representations) or far apart (sparsely packed representations). Participants were then tested for both veridical memory of the learned (old) images, and false memory of unstudied items in any part of the face space after a full night of sleep or an equivalent amount of daytime wakefulness. As expected, veridical memories were protected by sleep compared to wakefulness. However, contrary to expectation, we showed that sleep promoted generation of false memory for items in sparsely populated areas of this space compared to items in densely populated areas. We interpret these findings within a false memory framework.

2.2 Introduction

Sleep is critical for the consolidation (strengthening) of newly acquired memories (Born & Wilhelm, 2012; Rasch & Born, 2013; Stickgold, 2005, and more), as well as for other memory processes, such as integration of new information into pre-existing knowledge (Tamminen et al., 2010), generalisation (Lewis & Durrant, 2011) and gist abstraction of memories (Durrant, Taylor, Cairney, & Lewis, 2011; Durrant, Cairney, & Lewis, 2016; Durrant & Lewis, 2009; Durrant, Cairney, & Lewis, 2013) and even resultant formation of false memories, in which one falsely remembers items/events that have never been shown before (e.g. McKeon et al. 2012; Pardilla-Delgado & Payne, 2017; Payne et al., 2009; Lutz et al. 2017). However, the mechanisms by which sleep plays a role in gist extraction and resultant false memory formation has not been fully determined.

One framework, the information overlap to abstract (iOtA) model (Lewis and Durrant, 2011), has suggested that gist abstraction and resultant false memory based on the gist extraction occurs as a result of replaying overlapped memories during sleep. If this is the case, then greater degrees of memory overlap should be associated with more strongly abstracted gist and gist-based false memory after a period of sleep.

The current work set out to test this hypothesis using a paradigm in which recognition memory of facial images that were densely and sparsely grouped in an artificially constructed face memory space was compared in two groups who either slept over night or stayed awake for an equivalent amount during daytime.

To test this hypothesis, a main experiment (Experiment 2) was run to invent a novel study paradigm. In the invented study paradigm, face images were morphed along two axes (age and gender), creating an artificial 2D 'face space'. Participants were trained to remember specific facial images from this space, carefully choosing the position of these images in the face space in order to create both dense and sparsely populated sub-areas of the face memory space (see **Figure 2.4**). In a subsequent recognition memory test after a period of over-night sleep (sleep group) or equivalent daytime wakefulness (wake group), participants were shown both 'old' (learned) faces and 'new' (unlearned) faces from the created 2D face-space and they were asked to make old/new judgements. Then, this current work tested the iOtA's specific hypothesis about the role of sleep in gist-based false memory formation, by comparing false recognition memory performance (i.e. incorrect 'old' responses) for 'new' faces in the dense and sparsely populated areas of space in sleep group.

Before investigating the sleep role in gist-based false recognition memory in the main experiment (Experiment 2), a pilot experiment (Experiment 1) was run to determine two critical aspects about participants' perceptual discrimination ability on the created face images. First, Experiment 1 checked the minimal face morphing degree needed by participants to detect a change of face identity (e.g. the sensitivity index). This was critical, because it was necessary to ensure that the different stimuli which were presented were separated by sufficient perceptual space to be recognised as distinct. This was tested for both face morphing dimensions (age and gender). Second, Experiment 1 also examined difference of sensitivity in mix and single face discrimination task. In the planned study paradigm in Experiment 2, there

was a case that a face was morphed in both age and gender at the same time, which was depicted as a separation of a face morph to the other in a diagonal line in the invented 2D face morph space. Thus, Experiment 1 tested whether morphing diagonally showed relatively superior perceptual sensitivity relative to morphing on a single dimension. The information derived from this was used to determine how far apart faces needed to be on the diagonal axis in Experiment 2 in order to be perceived as distinct faces. Thus, if the diagonal dimension proved more sensitive, face morphs would be separated by a smaller face morphing degree diagonally than the defined minimum face morphing degree in single dimension. In other words, if the minimal distinguishable face morphing degree was 10 or 20 units in single dimension, then faces would be separated diagonally by 5 or 10 face morphing degrees, respectively.

2.3. Materials & Methods (Experiment 1)

Participants

12 healthy Caucasian participants (six males: age 29±3.52 (mean±SD), range 25–35 years, six females: age 26.33±4.08 (mean±SD), range 23–34 years) with normal or corrected-to-normal vision were recruited from colleagues of School of Psychological Sciences at University of Manchester. One participant was excluded due to poor performance in the discrimination tasks. Thus, data from 11 participants (six males: age 29±3.52 (mean±SD), range 25–35 years, five females: age 24.8±1.78 (mean±SD), range 23–27 years) was analysed.

Two-dimensional face morph scale

Colour photographs of 12 (6 females) Caucasian actors' faces were selected from Karolinska Directed Emotional Faces (KDEF) database (Lundqvist, Flykt & Ohman, 1998) and transformed into computerized facial images by FaceGen Modeller 3.5 (http://www.facegen.com) face morphing software (see **Figure 2.1A**). All faces were posed straight on to the camera, with neutral expressions and no obvious facial features, such as moles, beard, or unusual facial shapes of jaw, chin, nose and forehead. The computerized face of each actor was morphed along parameters of age and gender individually (see **Figure 2.1B**). The FaceGen software allows a face to be manipulated along a continuum of age, from a young to much older adult. It also can manipulate the gender of a face in a similar way. Accordingly, the FaceGen software was used to manipulate the age feature from the most young-look (0^{th} morph unit of age feature in FaceGen) to the most old-look face (its 50^{th} morph unit); gender feature was morphed from the most masculine-look (0^{th} morph unit of gender feature in Facegen) to the most feminine-look face (its 50^{th} morph unit). This morphing scale was relative rather than absolute. For instance, 0^{th} morph unit in age looked relatively younger than its 50^{th} morph unit, and vice versa. Allocating the age and the gender morphing scales in horizontal and vertical axis to put them together created a two-dimensional face morphing scale for an actor (see **Figure 2.1C**).

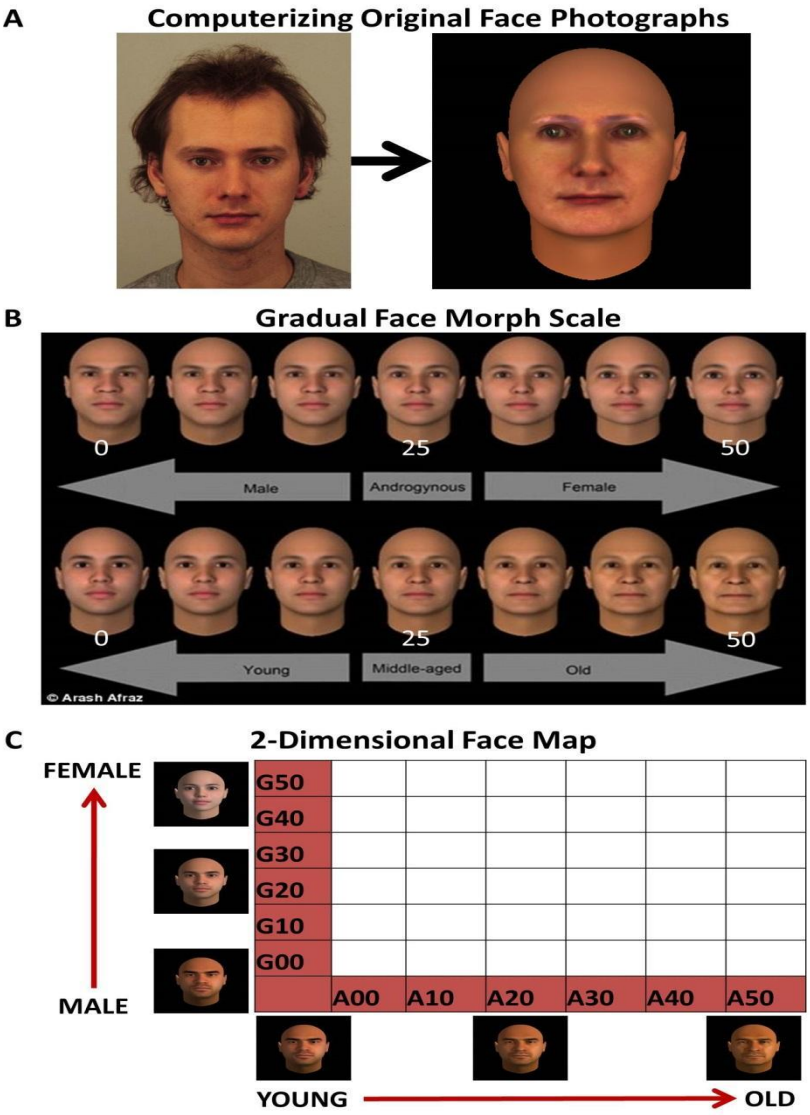

Figure 2.1. Face stimuli preparations: (A): An example of transformation of an original photograph (Left) into a computerized facial image (Right). (B): An example of face morphing in age and gender features. (C) An example of created two dimensional face morphing scale.

In total, 12 separate 2D face morphing scales corresponding to those selected 12 actors were created.

Procedure

In Experiment 1, participants were instructed to complete three types of discrimination task with facial images morphed along a 2D face morph scale; one

with faces morphing along features of age alone (age discrimination task), one with faces morphing along features of gender alone (gender discrimination task), and the other with faces morphing along features of age and gender together (mix discrimination task). In the age discrimination task, gender parameter was set to the most masculine-look (i.e. G00); age parameter was set to the youngest-look (i.e. A00) in the gender discrimination task. In the mix discrimination task, equivalent amounts of FaceGen morphing unit in age and gender parameters were morphed together. Between the age and gender discrimination tasks, there was a one minute break. The order of age and gender discrimination tasks was counterbalanced across participants. Please see a schematic diagram of experiment procedures in **Figure 2.2A** and **2.2C**.

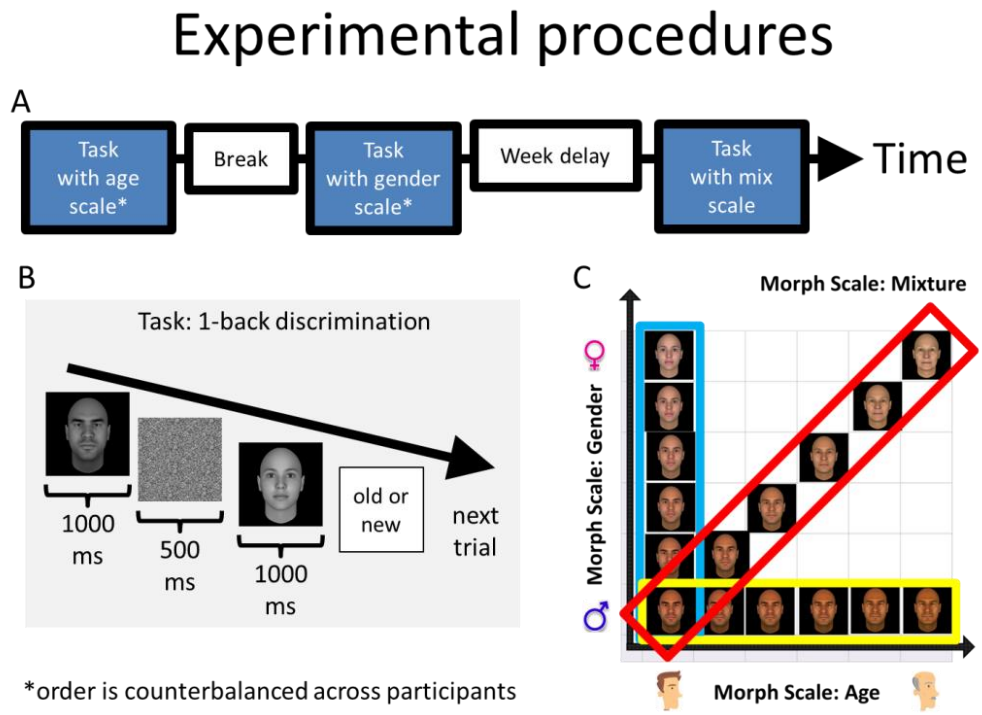

Figure 2.2. A diagram of experimental procedures: This diagram describes a time course of (A) experimental procedure and (B) 1-back face discrimination task and (C) a schematic diagram of three types (ie age alone, gender alone and mixture) of face morph scales. As indicated by icons on 2.2C, the young masculine faces on

the left bottom side of the square graph become older male faces on rightward, and the same (young male) faces become more feminine and young-look faces upward. Along the diagonal line, the same faces become older and feminine as the faces are morphed along the diagonal line.

The mix discrimination task was always run at least one week after completion of the other two discrimination tasks, so that performance of age and gender discrimination tasks was not influenced by mix discrimination task, and vice versa.

Discrimination task

Similar to a typical n-back task (e.g. Kirchner, 1958), a facial image with 0^{th} morph unit in both age and gender features, locating on the left bottom side on the **Figure 2.2C** graph, was shown at first as a reference face, and a face with either 0^{th} morph unit ('old' face, which was same as the reference face) or other morph units ('new' face, which was morphed from the reference face) was shown subsequently as a probe face. In each trial, participants viewed a reference face (for 1000 milliseconds), followed by a noise image, which was an image of random black and white dots (for 500 milliseconds). Next, a probe face was displayed (for 1000ms) and participants were asked to decide if the probe image was 'old' (an indication that participants believed that a probe face was same as the previous reference face) or 'new' (an indication that participants thought that a probe face was morphed from the reference face). All images were shown in grey scale in a size of 400 x 400 pixels on centre of computer screen. Once participants gave an old/new response, there was a 500 milliseconds delay before a next trial began and so on. In one block, 11 morph units (i.e. 0^{th}, 5^{th}, 10^{th}, 15^{th}, 20^{th}, 25^{th}, 30^{th}, 35^{th}, 40^{th}, 45^{th}, and 50^{th} morph units) of an actor's face were randomly shown as a probe face one at time for 10 repetitions,

which ended up 110 trials in total. In sum, there were 12 blocks; each block presented one of 12 actors' faces individually one at time and order of actor's face presentation was randomised across participants. Between blocks, there was a one-minute break. Please see **Figure 2.2B** for the visualised task procedure.

Statistical Analysis

In Experiment 1, 'old' responses to each probe face were collected and a rate of the 'old' responses over a total number of the probe face events (ie 10 events) was calculated. That meant that, if there were three 'old' responses on a probe face, its old response rate was calculated as 0.3. There were 11 types of probe faces morphing in a range of 0^{th} to 50^{th} morph units. Thus, I collected 11 different 'old' response rate of probe faces per block. Then, each of those 11 'old' response rate was averaged across 12 blocks individually. This calculation was repeated in the other two discrimination tasks as well. Then, 'old' response rates were normalised together in z-score across the 11 morph units of probe faces in three discrimination tasks. In order to find a point of subjective equality (PSE) or a ratio of 0.5 of 'old' response ratio, the normalised data was fixed to a scale of 0.0-1.0 by transforming the z-score into rate of standard normal cumulative distribution, which was used as normalised 'old' response ratios. Normalised 'old' response rates were used to exclude an outlier (ie, larger than two standard deviations, Olewuezi, 2011) across participants at any morph unit. The normalised 'old' response rates averaged across participants in each discrimination task at each morph unit detected that they surpassed a ratio of 0.5 at 15^{th} morph unit but were under 0.5 at the 25^{th} morph unit (see **Figure 2.3**). Thus, one sample t-tests of normalised 'old' response rates were

run to check if normalised 'old' response rates of three discrimination tasks was statistically different from PSE (i.e. a ratio of 0.5) at a point of 20th morph unit. In addition, a one-way repeated ANOVA (3 types of discrimination task: age, gender and mix) of normalised 'old' response rates was analysed if there was a difference of normalised 'old' response rates between the three tasks. Then, planned paired t-tests were conducted to test difference of normalised 'old' response rates between the mix and the age discrimination task and the mix and the gender discrimination task.

2.4 Results (Experiment 1)

One sample t-tests revealed that normalised 'old' response rates of neither age (0.54+/-0.14, (mean±SD)), gender (0.51+/-0.17) nor mix (0.43+/-0.14) discrimination tasks at the 20th face morph unit differed from 0.5 (age discrimination task: $t(10) = 0.92$, $p = .38$; gender discrimination task: $t(10) = 0.20$, $p = .84$; mix discrimination task: $t(10) = -1.55$, $p = .15$). This showed that normalised 'old' response rate of all three discrimination tasks did reach or were about to reach PSE at the 20th morph unit in this perceptual discrimination paradigm.

A one way repeated ANOVA revealed that there was a difference between the normalised 'old' response rate of the three discrimination tasks at the 20th face morph unit, $F(2, 20) = 7.54$, $p = .004$. Planned repeated t-tests revealed that the normalised 'old' response rate of the mix discrimination task (0.43+/-0.14, (mean±SD)) was lower than that of age (0.54+/-0.14) and gender (0.51+/-0.17) discrimination tasks [mix-age comparison, $t(10) = -3.33$, $p = .008$; mix-gender

comparison, t(10) = -3.71, *p* = .004], respectively (see **Figure 2.3**). In this context, a lower value meant that participants were able to detect a change from reference face more frequently. This suggests that participant's sensitivity to face morphing in the mix discrimination task was better than that in either the age or gender discrimination tasks at the 20th face morph unit in this perceptual discrimination paradigm.

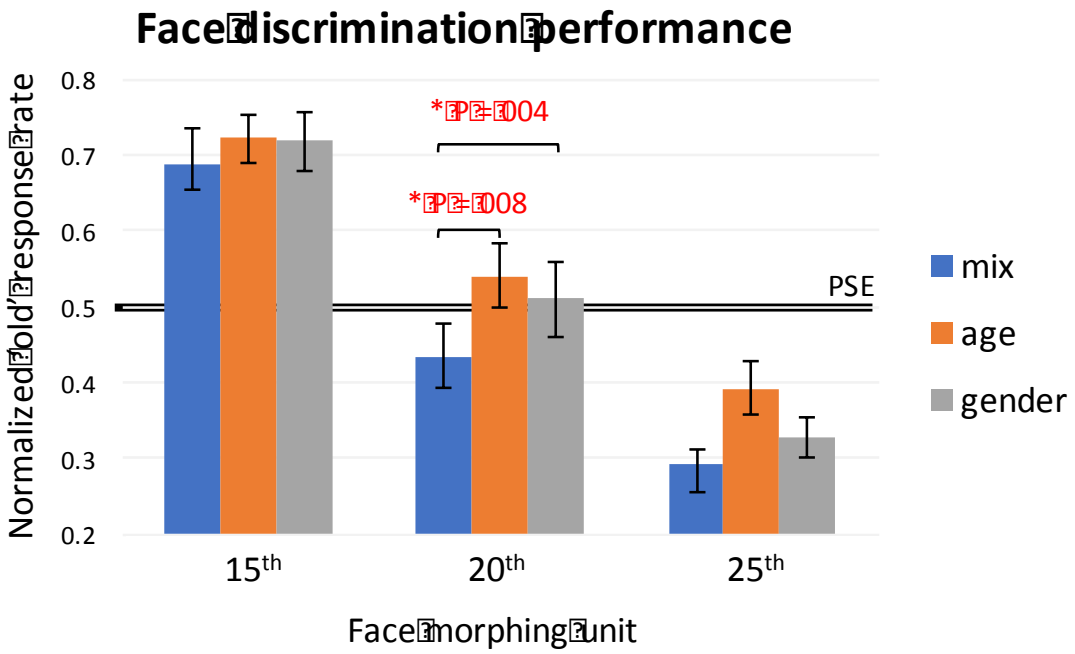

Figure 2.3. A bar chart of face discrimination performance at morphing unit of 15th, 20th and 25th: This graph drew normalised 'old' response rates on probe faces with specific morph units manipulated from the reference face. There were three types of face discrimination tasks: morphing in age alone, gender alone and mixture of morphing in age and gender. The higher the value, the more likely participants were to believe that presented probe faces were not morphed from a reference face. Point of subjective equality (PSE) was defined as normalised 'old' response rate at a rate of 0.5. An error bar was represented as a standard error of mean.

2.5 Discussion (Experiment 1)

Experiment 1 demonstrates that participants reach a threshold (i.e. PSE) for recognizing a change of identity from a reference face when morphing in a single

facial feature is 20 or larger face morph units. Based on the result, we determined that face images in Experiment 2 would be separated by at least 20 face morph units in a single facial feature in the 2D scale.

Another result of Experiment 1 shows that combination of two features of face morphing (i.e. age and gender) is relatively more perceptually detectable than a single change of either age or gender face morphing at the 20^{th} face morph unit. Accordingly, in Experiment 2, we decided to separate face morphs in a smaller distance diagonally (i.e. face morphing in age and gender at the same time). Due to the limitations of morphing a face in a range of 0^{th} to 50^{th} morphing point by the FaceGen face morphing software, there can be at most, three distinguishable faces (either a pattern of 0^{th}, 20^{th}, 40^{th}, or 10^{th}, 30^{th}, 50^{th}) vertically or horizontally in the invented 2D face morph space. Face morphs cannot be separated by 15 face morphing units, we chose to separate them by a distance of 10 face morphing units diagonally (i.e. a change in 10 face morphing units in both age *and* gender) in Experiment 2.

Notably, there are some discussion points regarding the methods of Experiment 1. One is that this discrimination task uses a short delay (i.e. 1 second) between a reference and a probe face. However, Experiment 2 will necessarily use a longer delay (i.e. 10 minutes and 12 hours) in order to investigate the effect of sleep. In addition, Experiment 1 only investigates participant's discriminating ability with respect to one specific face (i.e. 0^{th} morph unit) from the others. In Experiment 2, reference faces will be compared with probe faces (including old and new faces) around the entire invented 2D face morph scale. Furthermore, it was not tested

whether 20 face morph units in single face feature is equivalent to the 10 morph units in combinational facial features. Experiment 2 will examine if 20 morph units of a single face feature and 10 morph units of combining two facial features can be equally detected by participants.

2.6 Materials & Method (Experiment 2)

Participants

In the main experiment (Experiment 2), 52 healthy Caucasian participants (28 males: age 22.50±3.92 (mean±SD), range 18-35 years old, 24 females: age 23.88±5.62 (mean±SD), range 19-36 years old) were recruited from posters, e-mail and online advertisements. Participants were randomly assigned to either a sleep (n = 26, 13 males: age 23.46±4.58 (mean±SD), range 18-35 years old, 13 females: age 24.92±6.03 (mean±SD), range 19-36 years old) or a wake group (n = 26, 15 males: age 21.67±3.18 (mean±SD), range 18-30 years old, 11 females: age 22.64±5.08 (mean±SD), range 19-34 years old). They were screened for any history of neurological and psychiatric diseases and sleep disorders. They were asked to abstain from caffeine and alcohol 24 hours prior to experiment. This study was approved by the University of Manchester Research Ethics Committee. All participants were informed of the experiment details and signed consent forms.

Face stimuli in the invented two-dimensional face morph scale

The two findings from Experiment 1 determined a pattern of face stimuli selection in Experiment 2. First, a result from Experiment 1 showed that participants required

separating each face morph at least 20 FaceGen morphing units in a single face feature (either age or gender) to distinguish one face morph from the other. The result led to three distinguishable face morphs (either in a pattern at 0^{th}, 20^{th} and 40^{th} or at 10^{th}, 30^{th} and 50^{th} FaceGen morphing point) that can be recruited from each row array (ie three face morphs in vertical axis in **Figure 2.4A**) of the invented 2D face morph scale and same applied to its column array (i.e. three face morphs in horizontal axis in **Figure 2.4A**). That meant, 3x3 or 9 distinguishable faces can be recruited from the invented 2D face morph scale based on the first result. Second, it was also found in Experiment 1 that a combinational face morphing of age and gender features was perceptually more tangible for participants to distinguish between face morphs than a single face morphing. Based on the second finding, combinational face morphing of age and gender, which was represented as a diagonal or a combination of vertical and horizontal face morphing units in the 2D scale in **Figure 2.4A**, required a smaller morphing distance between face morphs to become distinguishable each other than single feature face morphing. Based on the first finding, face morphs were separated, either in a pattern at 0^{th}, 20^{th} and 40^{th} or 10^{th}, 30^{th} and 50^{th} FaceGen morphing points vertically and horizontally in the invented 2D space. The pattern of face selection did not leave an option to separate face morphs in a distance of 15 face morphing units in the 2D scale, as can be seen in **Figure 2.4A**. Therefore, face morphs are separated diagonally in a distance of 10 face morphing units in the 2D scale. Altogether, this ended up selecting 18 distinguishable face morphs from the 2D scale as a set of face stimuli in Experiment 2 (Please see illustration of selected 18 face morphs in **Figure 2.4A**).

Types of stimuli

Within the selection of 18 face morphs, they were designed in a way to have five types of facial stimuli: an old stimuli (i.e. O) memorized in the initial encoding phase, and four types of new stimuli (i.e. N0, N1, N2 and N3), which were presented at the subsequent test phase together with old stimuli and were adjacent to the old stimuli at a minimal face morphing distance (i.e. 20 units for single face morphing and 10 units for combinational face morphing) in the 2D array. The number next to the N indicated the number of old stimuli adjacent to the new stimuli. N0 was the only new stimulus, which was at distance of four spaces away from any old stimuli so that a difference between appearance of O and N0 reached the largest compared to the other O-N differences. Creation of the N0 helped in data analysis to calculate corrected hits, which is a formula of hits(O)-false alarms(N0). One diagonal distance was counted as two spaces apart based on one space apart from gender and the other from age axis. This stimuli arrangement was counterbalanced by rotating the pattern 90 degrees across the 2D face morph scale. Therefore, 2D face morph scales of four different actors (2 females) were used for each session. In total, for an entire experiment with two experimental sessions, 2D face morph scales of eight (4 female) actors were used. See **Figure 2.4B** to visualize these patterns.

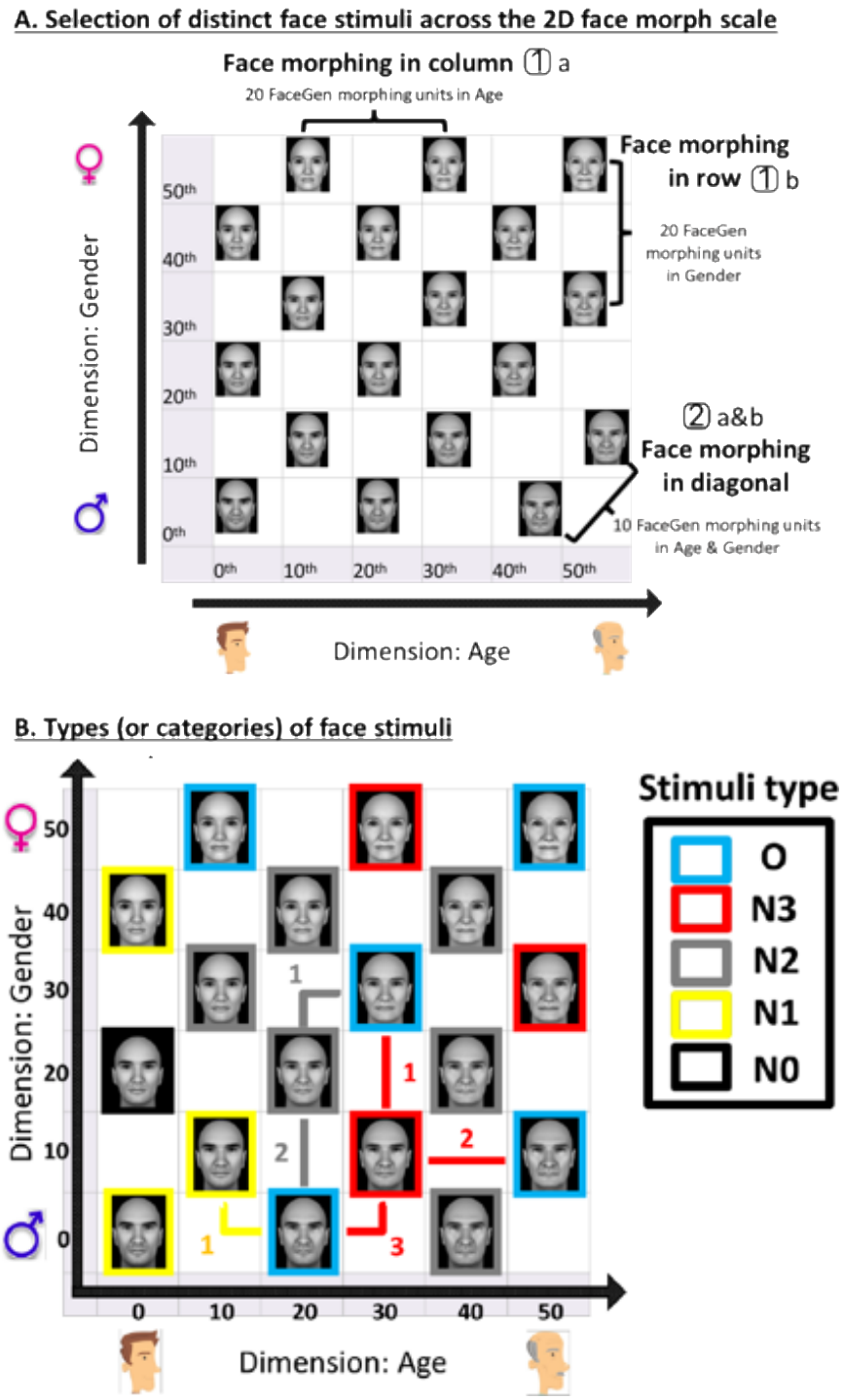

Figure 2.4. Selecting face stimuli from the invented 2D Face Morph Scale. There were two steps to select face stimuli from the invented 2D face morph scale. An example 2D face morph scale morphing a face in a range from the young to the old (the vertical dimension) and in a range from the masculine to the feminine (the horizontal dimension) looks.

Procedure

In Experiment 2, participants were randomly assigned into a wake or a sleep group. Participants of both groups completed two experimental sessions. In each session, five face morphs selected from the invented 2D face morph spaces were trained at the encoding phase. The encoded face morphs and new face morphs (never presented at encoding) were shown in the retrieval phase (i.e. taking an old/new recognition test) after a delay. In one of the two sessions, the post-encoding delay lasted for 10 minutes (the short delay session), and in the other session the delay lasted for 12 hours (the long delay session). The order of the short delay (sd) and the long delay (ld) sessions in the experiment was counterbalanced so that in both sleep and wake groups half of the participants started with the sd session and the other half started with the ld session, and vice versa.

The wake group started at 9:00. This meant that a half of the wake group participants (n = 13) completed the entire sd session and the encoding of the ld session in the morning. These participants were then free to go back to their daytime activity and were asked to come back to lab again at 21:00 to complete an old/new recognition test of the ld session. The other half of the wake group participants (n = 13) completed the experimental sessions in the opposite order; completing the encoding of the ld session in the morning and then coming back to lab to complete a subsequent old/new recognition test of the ld session and the entire sd session in the evening.

The sleep group followed same procedures as the wake group, but started at 21:00 and they were instructed to go back home to sleep at their own bed over night after

completing the encoding of the ld session. Participants in the sleep group were instructed not to deprive their sleep during the long delay, otherwise, they were instructed to report it and they were excluded from later data analysis. Similarly, the wake group participants were asked not to take a nap during the long delay, otherwise, they were instructed to report it and they were excluded from later data analysis. There were no reports of sleep deprivation and nap taking from either the sleep and the wake group, respectively. At the beginning of every lab visit in the morning and at night, participants completed questionnaires of Stanford Sleepiness Scale (SSS; Hoddes et al., 1973) and Karolinska Sleepiness Scale (KSS; Glenville, Broughton, Wing, & Wilkinson, 1978) to check their vigilance before the cognitive tasks were completed. A schematic illustration of a time course of the entire procedures is shown in **Figure 2.5A**.

In each block of encoding phase, participants were trained to memorise selected face morphs distinctively so that they formed an idiosyncratic representation of each face morph in their mind. In this phase, tasks were divided into face- association task and zero-delayed test. At the face- association task, each of five O (old) stimuli (i.e. five blue boxes in the encoding phase in the **Figure 2.5**) was shown on the centre of computer screen for 1000ms, then a corresponding number (i.e. either one of 1-5) was shown for 1000ms after a 500ms blank image between the gap. This was repeated five times.

At the retrieval phase after either a short or a long delay, participants were asked to complete the old/new recognition memory test. In this test, they were shown both old stimuli memorised at the encoding phase and new stimuli that were not previously

shown one at a time for 1000ms on the centre of computer screen. Here, participants were instructed to press “o” button when they thought presented stimuli were old stimuli.

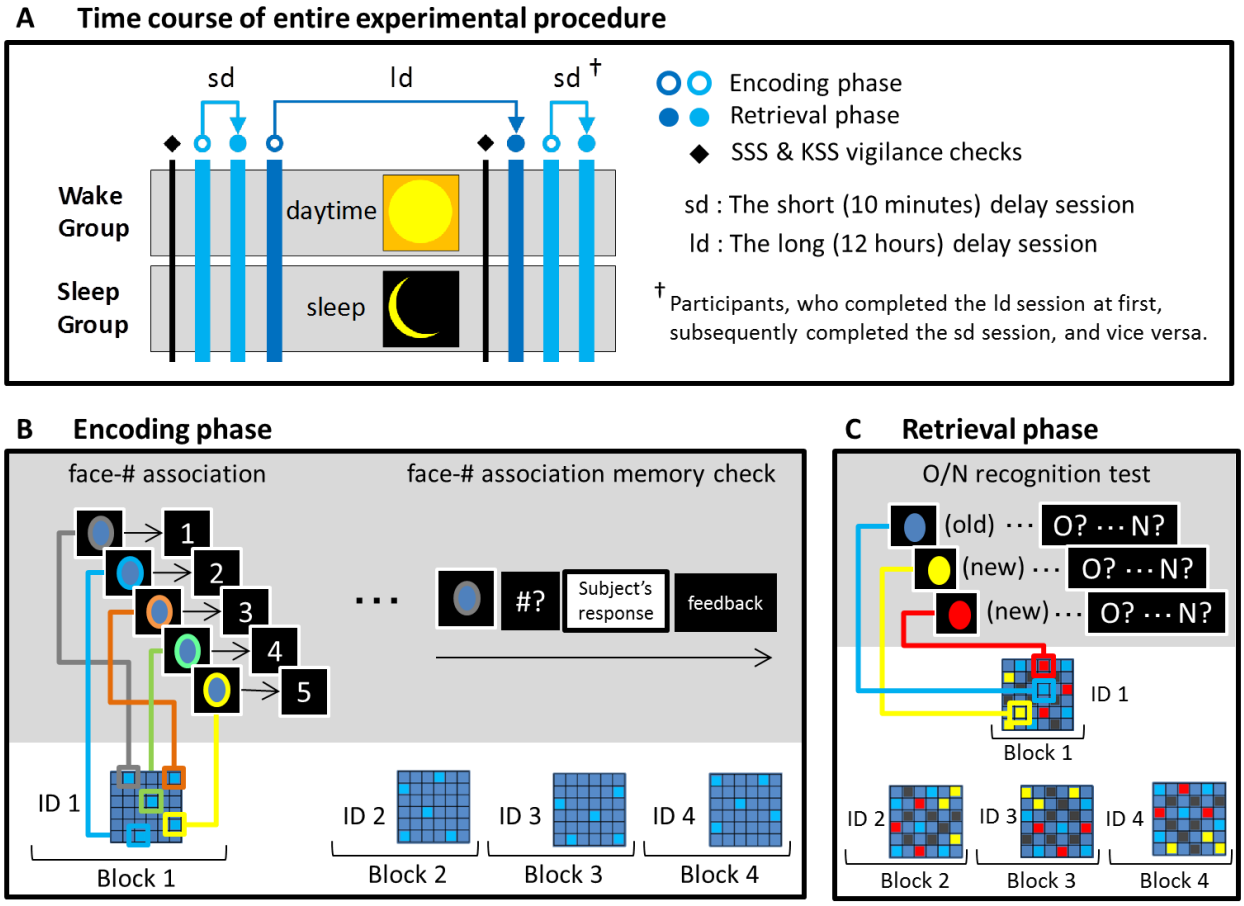

Figure 2.5. **A Schematic Illustration of Experimental Procedures of Experiment 2.** Here drew an illustration of (A) a time course of the entire experiment procedure of Experiment 2. The experimental procedure is mainly divided into (B) the encoding phase and (C) the retrieval phase. During the encoding phase, participants were trained to memorise O stimuli (blue boxes) by taking the face-# association task and zero-delayed test. This repeated for four identities (ID1 to 4). During the retrieval phase, participants took an old/new face recognition test.

In initial category learning task, participants were instructed to memorize association between five “O” facial stimuli and numbers of 1-5. A collection of face stimuli were

composed of five different versions of morphed faces of an actor within the 2D face morph map established in the pilot study. In another session, another set of four other actors were used. An exact face morphing pattern is shown as "O" on the 2D face morph map in **Figure 2.4**. In each trial of the learning task, each morphed facial image was shown on computer screen for 1 second followed by a number (i.e. either one of 1-5) shown for another 1 second and participants had to associate the face and the number. After the face-number presentations for the five morphed faces for twice, participants were shown each of the five facial stimuli again for a 1.5 second and had to respond which number was corresponding to the presented facial stimuli (for five times for each of the five facial stimuli). There was no time limit for participants to decide their responses then a correct number was shown after their decision for a 1.5 second. There were four trials, each of which used five morphed faces of different actors. In total, participants had to memorize 5 (distinctly morphed facial images) x 4 (actors) facial images for each session; another session used facial image set of four other actors. There was a short break between trials. An aim of the learning task was to make participants to distinctly memorize each morphed stimulus so that the corresponding numbers were not asked in later tasks. Between study and testing tasks, participants had either a 12-hour delay across a sleep or a 10-minute delay. During the 12-hour delay, they went home to sleep after completing the study task and came back to take the test task the next morning. Patterns of the delays between study and test tasks were counterbalanced across participants. That is, some participants had a 10-minute delay in the first session between study and test tasks and had a 12-hour delay in the second session, and vice versa. Different sets of facial stimuli were used for each session. See **Figure 2.5** for the detailed instructions.

2.7 Results (Experiment 2)

True Recognition Memory

Veridical recognition of old stimuli was analysed using a two-way mixed ANOVA (within-participant: 2 delays; between-participant: 2 groups). This revealed that there were no main effects of delay or group (Delay: $F(1,50)$ = 1.08, p = .31 Group: $F(1,50)$ = .18, p =.68, respectively). In contrast, there was a significant interaction between delay x group, $F(1, 50)$ = 6.26, p = .016. A post hoc t-test revealed that this was driven by a difference between the change in veridical memory performance across sleep (ld-sd) in the sleep group (m = .11, sd = .64) and wake group (m = -.27, sd = .44), $t(50)$ = 2.50, p = .016. Further post hoc t-tests showed that performance at long and short delays was significantly different in the wake group: sd (m = .64, sd = .35), ld (m = .37, sd = .31), $t(25)$ = -3.11, p = .005 (after Bonferroni correction, p = .01), and not the sleep group (m = -.27, sd = .44) and ld session (m = -.27, sd = .44), $t(25)$ = .89, p = .38 (see **Figure 2.6 B** for summary of this analysis). This result suggests that memory performance change over time differed between the wake and sleep groups due to deterioration of memory performance in wake group.

False Recognition Memory

False recognition of new stimuli was analysed using a three-way mixed ANOVA (within-participant: 2 delays x 2 stimuli; between-participant: 2 groups). This revealed

that there were no main effects of delay, stimuli or group (Delay: *F(1, 49)* = .83, *p* = .37, Stimuli: *F(1, 49)* = 3.72, *p* = .06, Group: *F(1, 49)* = .59, *p* = .44, respectively). However, there were interactions of delay x group *F(1, 49)* = 9.91, *p* = .003, and delay x stimuli, *F(1,49)* = 8.94, p = .004, while the interaction between stimuli x group was not significant, *F(1, 49)* = 2.28, *p* = .14. A post hoc t test for the interaction of delay x group revealed that change in false memory across sleep (ld-sd) differed between the sleep group(m = .43, sd = .75) and the wake group (m = -.21, sd = .71), *t*(50) = 3.15, *p* = .003. Similarly, a post hoc t test for the interaction of delay x stimuli revealed that the change of false recognition across sleep (ld-sd) in N1 stimuli (m = .33, sd = 1.04) differed significantly from the equivalent change in N3 (m = -.11, sd = .86), *t*(51) = 3.00, *p* = .004.

Confound checks

In order to ensure that the specific location of N1 and N3 items in the 2D face space was not responsible for my findings, I compared false memory results for single n- vs n+ items that were only separated by 2 spaces. False recognition was analysed using a three-way mixed ANOVA (within-subject: 2 delays x 2 stimuli; between-subject: 2 groups). This revealed that there were no main effects of delay, stimuli and group (Delay: *F(1, 50)* = 1.53, *p* = .22, Stimuli: *F(1, 50)* = 2.42, *p* = .13, Group: *F(1, 50)* = .09, *p* = .76, respectively). However, there were significant two way interactions of delay x group, *F(1, 50)* = 6.82, *p* = .012, and delay x stimuli, *F(1,50)* = 4.43, *p* = .04. Yet, the two way interaction effects of stimuli x group and three way interaction effect of delay x stimuli x group were not significant, (stimuli x group: *F(1, 50)* = .63, *p* = .43, Delay x Stimuli x Group: *F(1, 50)* = .83, *p* = .37, respectively). Post

hoc t test for the interaction of delay x group investigated differences of overall false recognition (the target N1 and the target N3 combined) across ld-sd changes between groups(i.e. wake and sleep), and revealed that overall increase of false recognition (the target N1 and the target N3 together) in the sleep group across ld-sd change (m = .50, sd = 1.05) differed significantly from decrease of false recognition (the target N1 and the target N3 together) in the wake group across ld-sd change (m = -.18, sd = .83), $t(50) = 2.61$, $p = .012$. Similarly, post hoc t tests for the interaction of delay x stimuli investigated difference of false recognition across ld-sd changes between stimuli (i.e. the target N1 vs the target N3), and that revealed that overall increase of false recognition in one N1 stimulus across ld-sd change (m = .35, sd = 1.21) differed significantly from decrease of false recognition in one N3 stimulus across ld-sd change (m = -.03, sd = 1.18), $t(51) = 2.11$, $p = .04$.

In order to determine whether there was a difference between diagonal and vertical or horizontal spacing I performed a confound check comparing false memory for a set of N- items that were separated by a single vh (vertical or horizontal) space to similar n- items separated by a diagonal space: False recognition was analysed using a three way mixed ANOVA (within-subject: 2 delays x 2 space; between-subject: 2 groups). This revealed that there was no main effect of spacing type (vh vs diagonal), $F(1, 50) = .19$, $p = .67$, indicating that these spacing types can be treated as equivalent. Similarly, the absence of two way interaction effects of space x group and space x delay and three way interaction effect of space x group x delay supported the above notion, (space x group: $F(1, 50) = .57$, $p = .46$, space x delay: $F(1, 50) = .06$, $p = .81$, space x delay x group: $F(1, 50) = .002$, $p = .97$, respectively). As expected, there was a main effect of delay alone, $F(1,50) = 6.24$, $p = .016$. In

addition, there was an interaction of delay x group, *F(1, 50)* = 7.91, *p* = .007, while there was no main effect of group, *F(1, 50)* = .87, *p* = .35. A post hoc t test on the difference of false recognition between delays (sd vs ld, all the other variables combined) showed that overall false recognition (m = -.01, sd = .79) in the ld session was smaller than overall false recognition (m = -.38, sd = .79) in the sd session, *t*(51) = 2.34, *p* = .023. Post hoc t tests for the interaction of delay x group investigated the change in false recognition across the night and revealed that the change in false recognition from short to long delay (ld-sd) differed between sleep group (m = .50, sd = 1.05) and wake group (m = -.18, sd = .83), *t*(50) = 2.61, *p* = .012.

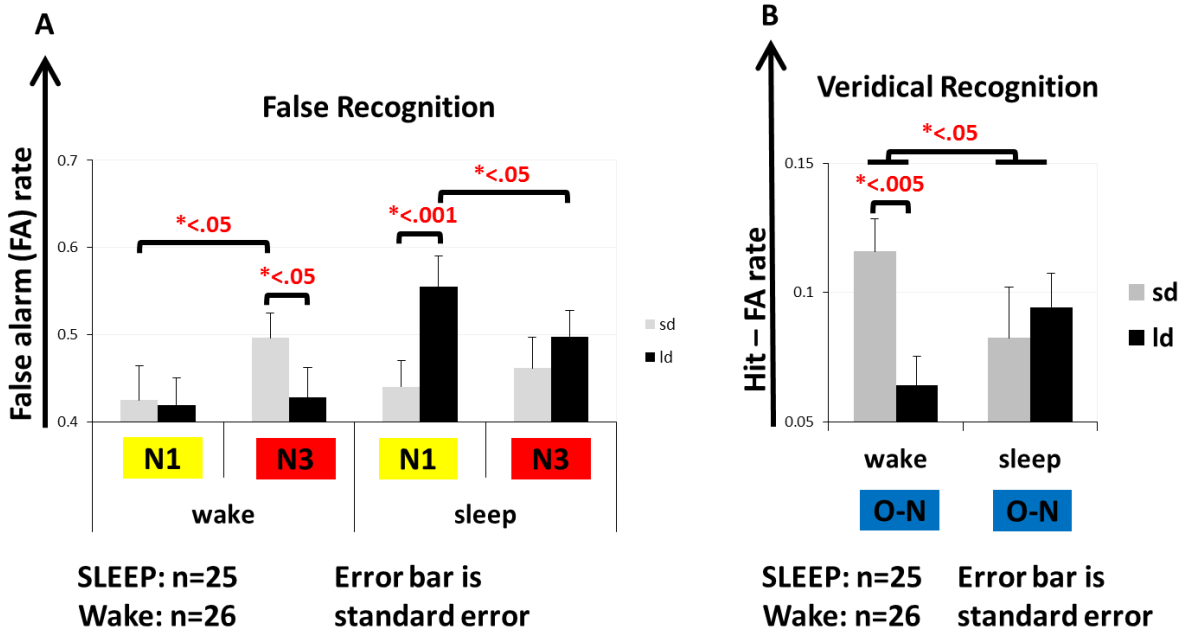

Figure 2.6. Subject's false and veridical recognition memory in the wake and the sleep groups. Here are results of participant's false and veridical recognition memory in the sd and the ld session. A) On the false recognition, there are two types of stimuli, N1(surrounded by only one O nearby) and N3(surrounded by three Os nearby). Y-axis is a value of old response ratios for new stimuli. B) On the veridical recognition, Y-axis is a value of hits – false alarms. N stands for all n stimuli.

2.8 Discussion (Experiment 2)

Contrary to the predictions of the iOtA model, our findings indicated that sleep increased the likelihood that participants would endorse false memories that fell within a sparsely populated memory space, as compared to the densely populated memory space. This observation does not match the predictions of the model, which would instead lead us to expect greater gist abstraction (and thus presumably false memory) for items falling within densely populated areas of memory space.

One possible explanation for this difference is the structure of the task, in which we attempted to test gist abstraction through the use of a false memory paradigm. In our task we examined responses to 'new' faces that had not been presented previously rather than examining whether memory for faces that had been presented ('old') was stronger in dense rather than sparse areas of memory.

The generation of gist-based false memories can be shown in perceptual or image list learning paradigms (e.g. Slotnick & Schacter, 2004), in addition to other forms of memory such as word lists. In visual list learning, faces are morphed by gradually altering from a starting face to a completely different face (Wilford and Wells, 2010). Morphing a stranger's face with the observer's own face or his or her acquaintances' face enhances subsequent recognition memory of the stranger's face (e.g. Keenan et al., 2003; Nevi, Cicali & Caudek, 2016) and even induces false recognition of new facial stimuli (e.g. Tomita, Yamamoto, Matsushita & Morikawa, 2014; Iidaka, Harada, Kawaguchi & Sadato, 2012). This is thought to occur because a face is informative,

containing multi-information about age, gender, race, eyes/hair/lip/skin colour, physical fitness and more, but is processed holistically to obtain gist of those contexts generally (Bruce &Young, 1986; Haxby, Hoffman & Gobbini, 2000). Therefore, merely being familiar with one face will allow us to build up the gist or a generalised idea of that face such that slightly morphing it would induce false memories.

Evidence has been inconsistent when determining whether sleep also enhances false memory, which is to the illusion of recalling events that have actually never been experienced before. Some work (e.g. McKeon et al. 2012; Pardilla-Delgado & Payne, 2017; Payne et al., 2009) has shown that sleep enhances false memory after learning word lists in the Deese Roediger McDermott (DRM) false memory paradigm (Deese, 1959; Roediger & McDermott, 1995). In the DRM paradigm, participants initially learn a list of items (e.g. a list of words of “winter”, “snow”, “hot” and “breeze”) that are all semantically or perceptually highly associated with a “lure” (e.g. a word of “cold”) which is the theme or gist of the learned list items. Participants are then asked to recall learned items either with no clue (i.e., free recall test) or from some available choices, such as whether a presented item is old or new (i.e. old/new recognition test). In the DRM paradigm, liability to falsely recall the gist or lure at a later test is used as an index of generating false memories; therefore, this type of false memory has also been called as gist-based false memory (see Schacter, Guerin and Jacques, 2011 for details). In the DRM paradigm with free recall tests, post-learning sleep has been shown to generate more gist-based false memory than equivalent amounts of wake delay (McKeon et al. 2012; Pardilla-Delgado & Payne, 2017; Payne et al. 2009). In contrast, other research has shown no sleep effect or

even shown a sleep-dependent decrease of the gist-based false memory when employing the DRM paradigm with recognition tests (Fenn et al. 2009; Lo et al., 2014).

Although there have been inconsistent findings, the research groups of McKeon et al. (2012) and Pardilla-Delgado and Payne (2017) have pointed out that the role of sleep in gist-based false memory might be small and dependent on the task demands of a participant's internal representation of memories. Indeed, these research teams have suggested that tasks such as a free recall task require more mental efforts to recollect own internal representation compared with a (e.g. old/new) recognition memory task, where the task is to decide between physically presented options. Thus, differing task demands may explain some of the reason why sleep benefited the gist-based false memory in some more 'difficult' free recall tasks but not 'easier' recognition tasks.

Importantly, the iOtA model does not make any prediction about false memories. Instead, it suggests that information shared between multiple true memories will be abstracted across sleep, creating a 'true gist' for these experiences. Still, it is reasonable to expect that strong gist representation might also lead to false memory for items that are consistent with this gist (Lewis & Durrant, 2011). This expectation did not match with our results. Instead, our results introduce a new insight that competing new information hinders such gist abstraction and reduction of false memories across sleep. This will be an interesting notion that can lead to modification of the iOtA model itself. Yet, further evidence requires to test this.

References

Born, J., & Wilhelm, I. (2012). System consolidation of memory during sleep. *Psychological Research*. https://doi.org/10.1007/s00426-011-0335-6

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. https://doi.org/10.1093/cercor/bhs244

Durrant, S. J., Cairney, S. A., & Lewis, P. A. (2016). Cross-modal transfer of statistical information benefits from sleep. *Cortex*, *78*, 85–99. https://doi.org/10.1016/j.cortex.2016.02.011

Durrant, S. J., & Lewis, P. A. (2009). Memory Consolidation: Tracking Transfer with Functional Connectivity. *Current Biology*. https://doi.org/10.1016/j.cub.2009.08.019

Durrant, S. J., Taylor, C., Cairney, S., & Lewis, P. A. (2011). Sleep-dependent consolidation of statistical learning. *Neuropsychologia*, *49*(5), 1322–1331. https://doi.org/10.1016/j.neuropsychologia.2011.02.015

Glenville, M., Broughton, R., Wing, a M., & Wilkinson, R. T. (1978). Effects of sleep deprivation on short duration performance measures compared to the Wilkinson auditory vigilance task. *Sleep*, *1*(2), 169–76. https://doi.org/10.1093/sleep/1.2.169

Hoddes, E., Zarcone, V., Smythe, H., Phillips, R., & Dement, W. C. (1973). Quantification of Sleepiness: A New Approach. *Psychophysiology*, *10*(4), 431–436. https://doi.org/10.1111/j.1469-8986.1973.tb00801.x

Kirchner, W. K. (1958). Age differences in short-term retention of rapidly changing information. *Journal of Experimental Psychology*, *55*(4), 352–358. https://doi.org/10.1037/h0043688

Lewis, P. A., & Durrant, S. J. (2011). Overlapping memory replay during sleep builds cognitive schemata. *Trends in Cognitive Sciences*. https://doi.org/10.1016/j.tics.2011.06.004

Lundqvist, D., Flykt, A., & Ohman, A. (1998). The Karolinska directed emotional faces (KDEF). *CD ROM from Department of Clinical Neuroscience, Psychology Section, Karolinska Institutet*. https://doi.org/10.1017/S0048577299971664

Lutz, N. D., Diekelmann, S., Hinse-Stern, P., Born, J., & Rauss, K. (2017). Sleep Supports the Slow Abstraction of Gist from Visual Perceptual Memories. *Scientific Reports*, *7*, 42950. https://doi.org/10.1038/srep42950

McKeon, S., Pace-Schott, E. F., & Spencer, R. M. C. (2012). Interaction of Sleep and Emotional Content on the Production of False Memories. *PLoS ONE*, *7*(11). https://doi.org/10.1371/journal.pone.0049353

Olewuezi, N. P. (2011). Note on the Comparison of Some Outlier Labeling Techniques. *Journal of Mathematics and Statistics 7(4)*, 353-355.

Pardilla-Delgado, E., & Payne, J. D. (2017). The impact of sleep on true and false memory across long delays. *Neurobiology of Learning and Memory*, *137*, 123–133. https://doi.org/10.1016/j.nlm.2016.11.016

Payne, J. D., Schacter, D. L., Propper, R. E., Huang, L. W., Wamsley, E. J., Tucker, M. A., … Stickgold, R. (2009). The role of sleep in false memory formation. *Neurobiology of Learning and Memory*, *92*(3), 327–334. https://doi.org/10.1016/j.nlm.2009.03.007

Rasch, B., & Born, J. (2013). About sleep's role in memory. *Physiological Reviews*, *93*(2), 681–766. https://doi.org/10.1152/physrev.00032.2012

Stickgold, R. (2005). Sleep-dependent memory consolidation. *Nature*, *437*(7063), 1272–1278. https://doi.org/10.1038/nature04286

Tamminen, J., Payne, J. D., Stickgold, R., Wamsley, E. J., & Gaskell, M. G. (2010). Sleep Spindle Activity is Associated with the Integration of New Memories and Existing Knowledge. *Journal of Neuroscience*, *30*(43), 14356–14360. https://doi.org/10.1523/JNEUROSCI.3028-10.2010

Chapter 3

Sleep-dependent false recognition of facial images is associated with slow wave sleep

Hikaru Tsujimura, University of Manchester, UK

Sonja Kotz, University of Manchester, UK

Karen Lander, University of Manchester, UK

Alessandra Tafuro, University of Manchester, UK

Penelope A. Lewis, University of Manchester, UK

3.1 Abstract

A previous study showed initial evidence that sleep can lead to an increase in false memories in areas of memory space where there is little competition and a decrease in false memories in areas of memory space where there is more competition. Extensive work has shown that slow wave sleep is important for declarative memory consolidation, however, as sleep was not recorded in the previous study, it is unclear whether the observed phenomena is associated with this sleep stage. The present study set out to replicate this finding with a paradigm that had been improved through a better balancing of face locations within the 3D face space, and also to extend it by recording sleep and search for an association between time spent in slow wave sleep and the observed memory effects. This work not only supported the prior study by again showing an across-sleep increase in the extent to which new faces in the non-competitive area of face space were endorsed as old relative faces in competitive area of face space ($p=0.03$), but also revealed a strong correlation between the proportion of time spent in slow wave sleep and the extent of this advantage for faces in the non-competitive areas, suggesting that consolidation in slow wave sleep does play a role in the observed false memory phenomena ($p=0.03$).

Because both our prior studies (Experiments 1 & 2) and Experiment 3 used upright face images, which may be consolidated differently from other visual stimuli as the to-be-remembered stimuli, we were interested to know if the observed pattern of consolidation extends to inverted faces. We thus performed an additional study (Experiment 5) using the same paradigm, but this time with inverted faces as our

stimuli. This confirmed that the pattern of false memory increases by sleep is not specific to upright faces but applies to faces in general.

3.2. Introduction

A growing body of evidence has shown that sleep facilitates gist-based false memory, falsely remembering items that have never been experienced but that are matched with generalised ideas or the gist of previously studied items (Pardilla-Delgado & Payne, 2017; Diekelmann, Born, and Wagner 2010; Lutz et al., 2017). The information overlap to abstract (iOtA) model proposed by Lewis and Durrant (2011) suggests that sleep promotes extraction of gist from overlapped memories in studied item. The model does not explicitly suggest that this process could also promote the formation of false memories where the gist of a set of learned items overlaps but was never actually learned, however this is a distinct possibility under the framework of the model.

Recently, a study by Tsujimura and Lewis (2018) provided a new insight about the iOtA hypothesis by showing that sleep's ability to promote gist-based false memory was attenuated by the presence of competition between overlapped studied items. Specifically, Tsujimura and Lewis (2018) built a novel experiment paradigm in which participants were trained on a series of computer generated facial stimuli and then given a subsequent old/new face recognition memory test. As participants were instructed to memorise each facial stimulus as a distinct item, competition existed between trained facial images while appearance of each image was perceptually overlapped. The results of this study showed that post-learning sleep promoted

generation of gist-based false recognition memory (i.e. participant's 'old' responses on 'new' items) only when away from this memory competition compared to post-learning wakefulness.

However, this previous work did not examine which specific sleep stage is important for the effects observed. Based on the active systems consolidation model of sleep-dependent memory consolidation (Born & Wilhelm 2012), as well as the iOtA model (Lewis & Durrant 2011), the amount of SWS obtained in a night was expected to predict the sleep-dependent generation of gist-based false recognition memory. The present work set out to test this prediction by determining whether SWS is important for the promotion of gist-based false recognition memories. To test this hypothesis, we ran our main experiment (Experiment 3), which employed a revised version of the study paradigm originally introduced by Tsujimura and Lewis (2018), while monitoring sleep activity during post-learning sleep and correlated overnight changes in performance with SWS. Importantly, we reduced the number of face stimuli compared to that used in our prior study in order to focus on the sleep-dependent changes in false recognition memory performance.

Face stimuli are often processed in a unique manner compared to other visual stimuli, but inverted faces are not processed in this manner (see a review Haxby, et al., 2000). Therefore, we also determined whether the impact of sleep observed in this present work with the face morphing paradigm are specific to upright face images or can be extended into face images in general. To examine this, we performed a second experiment (Experiment 4), this time using the exact same

behavioural paradigm as the first experiment (Experiment 3), but with inverted faces as stimuli.

In sum, the objectives of this present work were (1) to replicate the findings in the work by Tsujimura and Lewis (2018), (2) to extend these by determining if these overnight changes are specifically associated with SWS, and (3) to determine if the sleep-dependent production of gist-based false recognition also effects inverted images of faces.

3.3 Materials & Methods (Experiment 3)

Participants

In this current experiment (Experiment 3), 32 healthy Caucasian participants (11 males: age 26.64±6.25 (mean±SD), range 20-37 years old, and 21 females: age 23.29±2.70 years old, range 18-29 years old) were recruited from posters, e-mails and online advertisements. Participants were randomly assigned to either a sleep (n = 16, 5 males: age: 26.20±5.40 years old, range 21-35 years old and 11 females: age 23.73±3.10 years old, range 18-29 years old) or a wake group (n = 16, 6 males: age 27.00±7.38 years old, range 20-37 years old, 10 females: age 22.80±2.25 years old, range 20-26 years old). They were screened for a history of neurological and psychiatric diseases and sleep disorders. They were asked to abstain from caffeine and alcohol 24 hours prior to the experiment. This study was approved by the University of Manchester Research Ethics Committee. All participants were informed about the experiment details and signed consent forms.

Computationally generated face stimuli

To measure sleep-dependent changes in face recognition memory task, this experiment employed a collection of facial images computationally generated by FaceGen Modeller 3.5 (http://www.facegen.com). This face collection was originally invented by the work of Tsujimura and Lewis (2018). In detail, the original face set was created by gradually morphing a face of an actor in two facial features (i.e. age & gender) in 6 steps, assigning one feature morph as a horizontal axis and the other feature morph as a vertical axis, creating a 6 x 6 2D face morph scale. This was repeated for seven other actors, resulting in a collection of eight novel 2D face morph scales (labelling them as ID1 to 8). In the previous work, 18 face morphs (i.e., five old(items trained at encoding) and 13 new(items not trained)) were recruited from each invented 2D face morph scale, but this current experiment utilised only eight (i.e., four old, four new) of the morphs from each invented 2D face morph scale to balance a number of old and new items. For example, the four old items, defined as “O,” were the blue boxes in **Figure 3.1.**

Of the four new items, two (i.e. grey boxes in **Figure 3.1**) were allocated farther away from the old items than the other two new items to old items. Those farther new items were termed as “nfar.” The remaining two new items closer to one or two of old items were categorised as “n-” and ‘n+’. “N+” (red box in **Figure 3.1**) was the one that was surrounded by two old items (indicated by two red connecting lines in **Figure 3.1**) and “n-“ was the one that was surrounded by only one old item (indicated by one yellow connecting line in **Figure 3.1**).

The reduction of a number of face stimuli into eight items helped participants to focus on a smaller number of items and allowed them to make old/new judgements in an equal old/new distribution at retrieval. Moreover, the current experiment arranged “n+” and “n-“ items at neighbouring areas in the 2D scale so that images of “n+” and “n-” face morphs were perceptually highly resembled. This arrangement also made appearance of each old item distinct from “nfar” items, so that participants can discriminate between old and “nfar” items at an old/new recognition memory test.

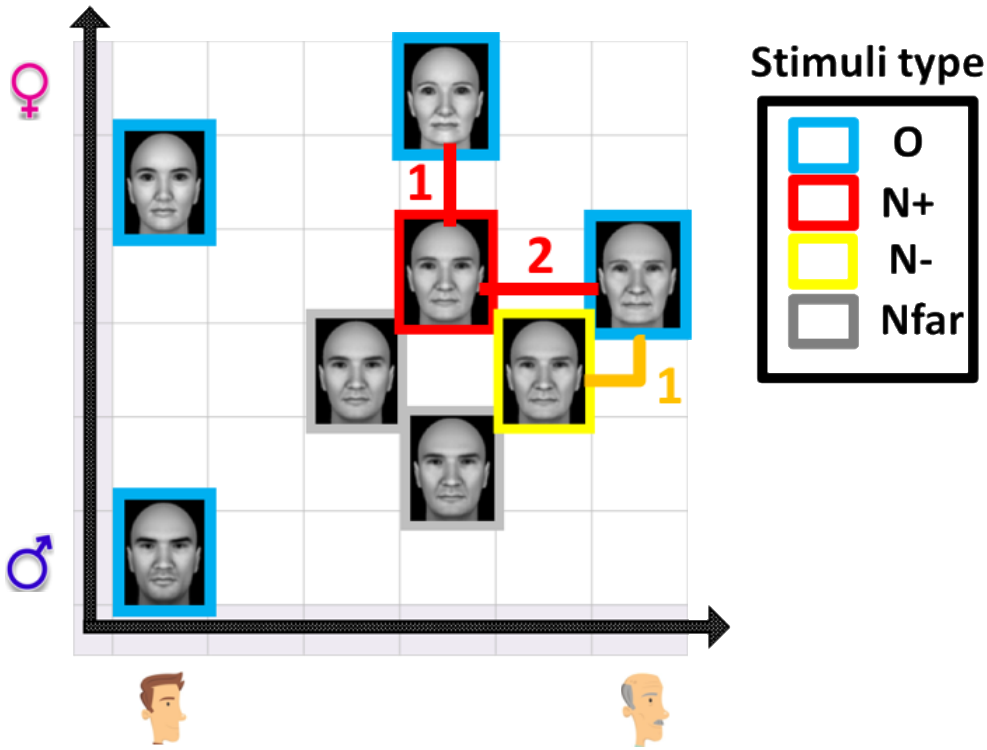

Figure 3.1. **An example of the invented 2D face morph scale**: In this example, a vertical axis represented face morphing in the feature of gender and a horizontal axis represented face morphing in the feature of age. Highlighted colour indicated a type of face stimuli. Blue meant that those items were classified as an “O”, which was trained in the initial encoding task. Red, yellow and grey items were classified as a “N+,” a “N-,” and a “Nfar” new items which were shown in the later retrieval task together with the “O” items. Of the new items, red colour indicated that those new items were surrounded by two numbers of nearby “O” items, yellow colour were

surrounded by one nearby “O” item and grey items were new stimuli surrounded by “O” farther than the other two new items.

Design

In this experiment, participants were assigned to wake or sleep groups randomly. Participants of both groups went through two sessions of the experiment. Each session was composed of an encoding and subsequent retrieval tasks with a delay between tasks. In one session, a delay lasted 10 minutes (short delay session) and a delay lasted 12 hours in the other session (long delay session). The order of the short and the long delay sessions was pseudo-randomly assigned to participants in each group. Thus, in both sleep and wake groups, half of the participants (i.e., n = 13 in sleep and n = 13 in wake) started the first session with a short delay and the other half started with the first session containing a long delay, and vice versa. The wake group started at 9:00. That meant that a half of wake group participants completed the first session and encoding of the second session in the morning and then were free to go back to their daytime activity. They were then asked to come back to lab again at 21:00 to complete retrieval task of the second session. The other half of wake group went through in an opposite way, completing encoding of the first session at 9:00 in the morning and completing a subsequent retrieval task of the first session and the second session in the evening. The sleep group followed same procedure as the wake group, but started at 21:00 and they were instructed to sleep over-night at the sleep lab of University of Manchester after completion of cognitive tasks. They were asked to complete their usual sleep rituals, such as brushing teeth, changing into pyjamas, etc. They were then set up with electrodes on their scalp so that we could collect standard polysomnography (PSG) recordings. After the PSG

setting, participants were allowed to go to sleep for eight hours. In the following morning, they were awoken by experimenters and allowed to take a shower before starting a retrieval task for this session. The wake group was asked not to take a nap during the long delay. Before the beginning of each session, participants completed questionnaires of Stanford Sleepiness Scale (SSS; Hoddes et al., 1973) and Karolinska Sleepiness Scale (Glenville, Broughton, Wing, & Wilkinson, 1978) to check their vigilance (represented as black diamond symbols in **Figure 3.2**). Please see **Figure 3.2** for detailed procedures.

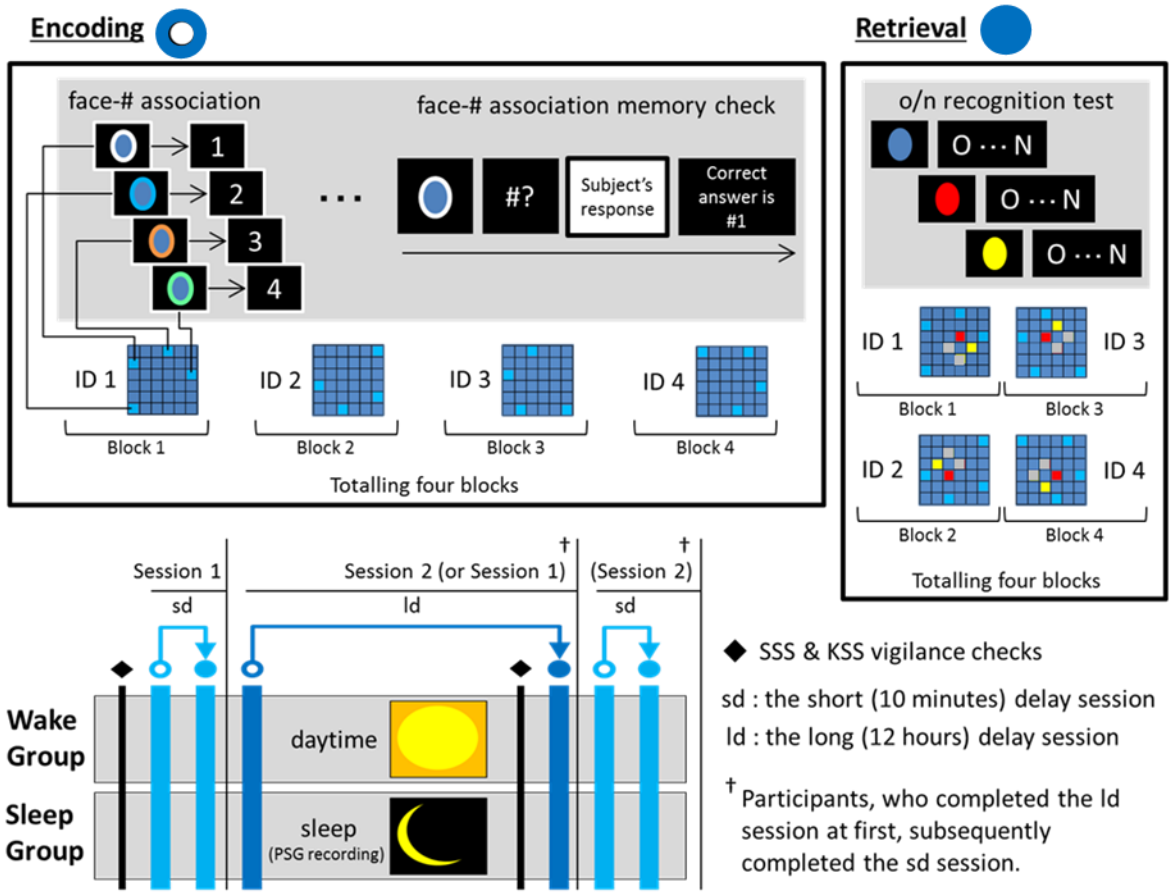

Figure 3.2. A schematic diagram of experiment design and procedures: Participants of both sleep and wake groups completed two sessions. Wake groups started in the morning and sleep groups started in the evening. Each session underwent same procedures, initially starting with the encoding task (memorised a collection of "O" facial stimuli) and then preceding the retrieval task (taken an old/new face recognition memory test). Between

tasks, there was either a short (10 minutes) or a long (12 hours) delay. Wake groups were freed to go back to their daytime activity during the long delay. Sleep groups slept at a bed at lab during the long delay while monitoring their sleeps by the PSG recording systems.

Cognitive Tasks

This current experiment employed two types of cognitive task, an encoding and a retrieval task. Both tasks contained four blocks; each block employed a face set of an actor (e.g., ID 1). In the encoding task, participants were instructed to memorise the association between four "O" facial stimuli and the numbers of 1 to 4 in each of the four blocks. This face-number pair was randomly assigned to each block. The facial stimuli used composed of four different versions of morphed faces of an actor within the 2D face morph map, as established in the section of **computationally generated face stimuli.** The exact face morphing pattern was highlighted in blue for the "O" items on the 2D face morph map in **Figure 3.1**. In each trial of the encoding task, each morphed facial image was shown on computer screen for 1.0 second followed by a number (number 1 - 4) shown for another 1.0 second. Participants were asked to associate the face with the number. The face-number presentations of the four morphed faces were presented twice. After that, participants were shown each of the four facial stimuli again for 1.5 seconds and then they were asked to respond with the associated number. There was no time limit for participants to decide their responses then a correct number was shown after their decision for 1.5 seconds. This repeated until participants reached 100% accuracy of the face-number pairing for all four face stimuli twice in a row. There were four blocks, each of which used four morphed faces of different actors (e.g., ID1-4). In total, participants had to memorize 4 (distinctly morphed facial images) x 4 (actors or IDs) facial images for

each session. The second session used a facial image set of four other actors (e.g. ID5-8) to develop face-number association memory of other 4x4 facial images. There was a short break between blocks. The aim of the encoding task was for participants to memorize each morphed stimulus perceptually distinctly so that the corresponding numbers were not required at retrieval tasks.

The other type of task was a face recognition memory test conducted at retrieval 10 minutes (the short delay session) or 12 hours (the long delay session) after an encoding task. In this retrieval task, participants were shown both old (encoded facial images) and new facial stimuli (never encoded images) of the 2D face morph map one at a time for 2.0 seconds. Participants were instructed to indicate if presented faces were old or new during the period. At the same time, participants were overtly informed that a half of presented images were old and the other half were new and instructed to give an equal amount of old and new responses in this task. Order of facial stimuli were quasi-randomly presented so that each of the eight facial stimuli (i.e. four old, one n+, one n- and two nfar items) was shown six times. This repeated across four blocks, each of which showed a face set of each facial ID as same as the encoding task. See **Figure 3.2** for detailed instructions.

Polysomnography (PSG) data acquisition and analysis

Polysomnography (PSG) data was acquired and digitized by employing Embla N7000 sleep monitoring system (Medcare-Embla®, Reykjavik, Iceland) at a sampling rate of 200Hz. EEG electrodes were placed on the scalp at F3, F4, Fz, C3, C4, Cz, O1, and O2, according to the international 10-20 system, referenced at left

and right mastoids (A1 and A2) and ground referenced at forehead. Eye movements were monitored on left and right eyes (E1 and E2), and chin movements were recorded at left, right and bottom of chin. Impedance of electrodes was verified until reaching under 5kΩ. PSG data was collected and scored offline by trained experimenters based on the AASM protocol (American Academy of Sleep Medicine, Westchester, IL), using RemLogic 1.1 software. Sleep stages were scored either wake, sleep stage 1(N1), stage 2(N2), slow wave sleep (SWS) or rapid eye movement (REM) sleep, then each percentage of N1, N2, SWS and REM(R) was calculated over a total sleep time.

Statistical analyses

In this experiment (Experiment 3), participants' old responses were collected and old response rates (a number of old responses over total events) of the four "o," two "nfar," two "n+," and two "n-" items were calculated per block (i.e. face ID) and each of the eight items was averaged across four face IDs for the short and the long delay sessions in wake and sleep groups. Veridical recognition memory performance was initially calculated by averaging old response rates on "o" items normalised by z-score in the short delay (sd) and long delay session (ld), respectively, and a ld-sd change in the wake and sleep groups were analysed. First, Experiment 3 employed one-sample *t*-tests in both groups to test if the ld-sd change of veridical recognition reduced lower than zero. As a wake group was expected to reduce after a long delay, one-tailed p-value was calculated. Second, an independent *t*-test (one-tailed) was conducted to determine whether veridical recognition in wake group was deteriorated by ld-sd change compared to that of sleep group. Then, this experiment

employed a 2 (group (between subject design): wake and sleep groups) x 2 (stimuli (within subject design): n+ and n- sessions) mixed analysis of variance (ANOVA) to test whether changes of veridical and false recognition memory across an overnight sleep differed from such changes across an equivalent amount of wakefulness delay. Besides this, a planned independent t-test was conducted to test if veridical recognition memory across an overnight sleep was maintained in sleep-dependent ld-sd change, compared to expected reduction of veridical recognition memory in wake-dependent ld-sd changes.

3.4 Results (Experiment 3)

Veridical Recognition Memory Performance

In this main experiment (Experiment 3), participants in both groups showed a decrease in veridical recognition (i.e., an old response rate on old items) over the course of the experiment (see **Figure 3.3A**). Importantly, however, this reduction was significantly lower in the sleep group (*m* = -0.13, *sd* = 0.08) than the wake group (*m* = -0.35, *sd* = 0.11), planned t-test *t*(30) = -1.71, *p* = .05. Furthermore, planned one sample t-tests revealed that the over-time change in the sleep group (m = -0.13, *sd* = 0.08) was not significantly different from zero *t*(15) = -1.67, *p* = .23, but the over-time reduction in the wake group (*m* = -0.35, *sd* = 0.11) was significant, *t* = -3.31, *p* = .01 (see **Figure 3.3A**). P-values were adjusted by Bonferroni correction. This result is in keeping with the literature in that it suggests memory is protected against decay over a period of sleep, but not wake.

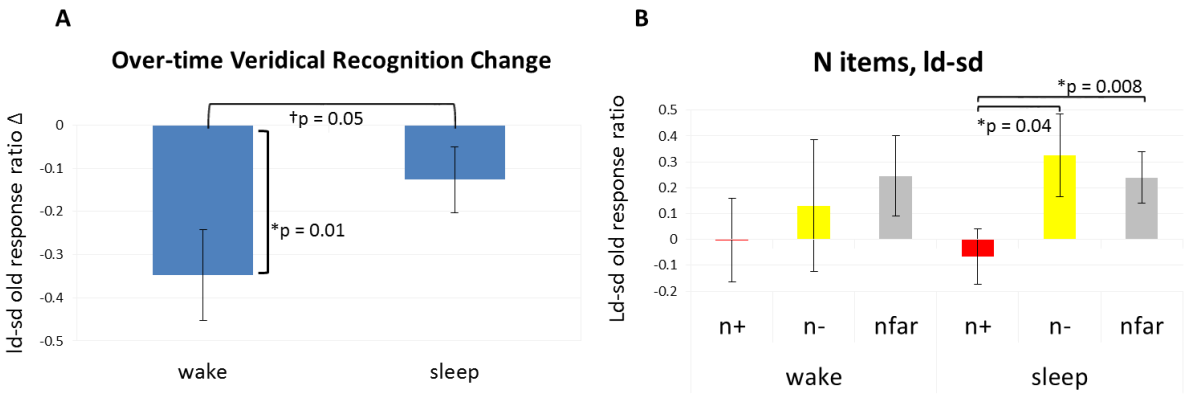

Figure 3.3. Changes over time (wake/sleep) in A) veridical recall and B) false recognition for upright face images. Y-axis of both graphs indicate an overtime change of old response rate across sleep or wake delay. On A), this data shows overtime old response changes on the studied items, so this shows an overtime change of veridical recall. In contrast, B) shows overtime old response changes on the unstudied new items, so this data indicates how false recognition is altered across a sleep or wake delay.

False Recognition Performance

In contrast, false recognition (i.e., an old response rate on new items) increased over time between short delay and the long delay in both sleep and wake groups in all but the N+ condition in the sleep group.

A 3 (stimuli types) x 2 (groups) mixed ANOVA, comparing over-time change in false recognition on three types of new items (groups: N+, N- and Nfar) in the sleep and wake group, revealed that there was a main effect of stimuli type, $F(2, 60) = 3.49$, $p = .04$, but no main effect of group, $F(1, 30) = 2.37$, $p > .05$ and no interaction between stimuli type x group, $F(2, 60) = .66$, $p > .05$. Post hoc t-tests revealed that this was driven by a larger over-time change in Nfar false recognition ($m = 0.24$, $sd = 0.51$) than N+ ($m = -0.03$, $sd = 0.54$), $t(31) = 2.83$, $p = .008$ (Bonferroni corrected $p = .02$), and a larger over-time change in N- false recognition ($m = 0.23$, $sd = 0.84$)

than N+ (*m* = -0.03, *sd* = 0.54), *t*(31) = 2.17, *p* = .04 (Bonferroni corrected *p* > .05). Yet, there was no difference between over-time change in Nfar false recognition (*m* = 0.24, *sd* = 0.51) and N- false recognition (*m* = 0.23, *sd* = 0.84), *t*(31) = .11, *p* > .05. (**Figure 3.3B**).

A further planned one-tailed paired t-test (N- vs N+) was conducted on the sleep data because a relatively larger false recognition change on N- compared to N+ was expected based on the findings in the previous study (Tsujimura & Lewis 2018)). This comparison revealed that N- false recognition change (*m* = 0.32, *sd* = 0.64) was marginally larger than that of N+ (*m* = -0.07, *sd* = 0.43), *t*(15) = 2.09, *p* = .03 (**Figure 3.3B**).

Relationship between sleep and recognition memory

To determine whether there was a relationship between sleep stages and z-score for overnight changes in memory performance, I tested for correlations between these measures. Change in veridical recognition across a night of sleep did not correlate with SWS, *r*(14) = .28, *p* = .30. Importantly however, the difference between N- and N+ false recognition was positively correlated with SWS, *r*(14) = 0.74, *p* = .001 (Bonferroni corrected *p* = .002) (see **Figure 3.4A**). Looking at this more closely, z-score for change in false recognition of N- items was positively correlated with SWS, *r*(14) = 0.62, *p* = .01, (Bonferroni corrected *p* = .03), while overnight change in N+ and Nfar recognition showed no significant correlation with SWS, (N+: *r*(14) = -0.35,

p = .18, Nfar: *r*(14) = .28, *p* = .30)

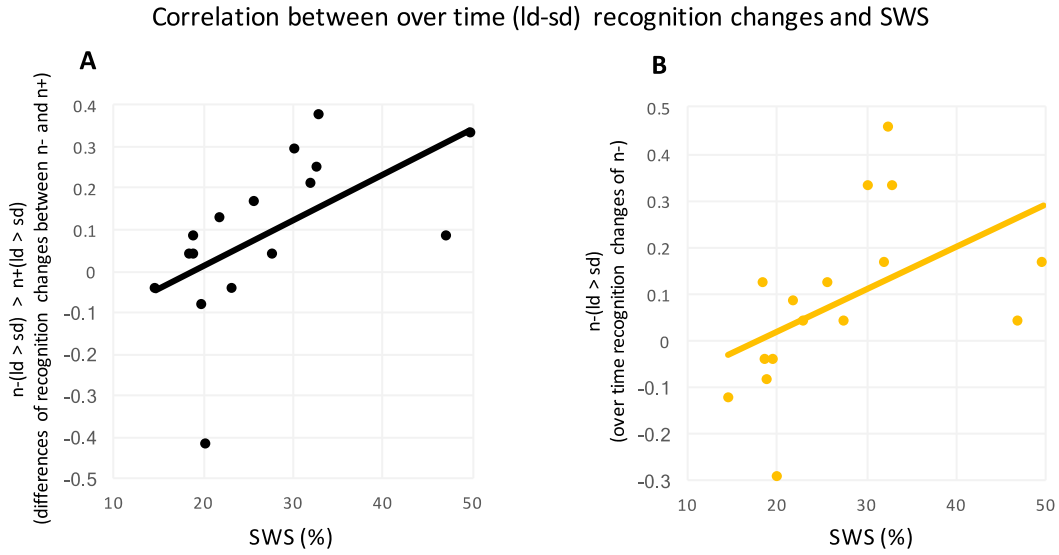

Figure 3.4. Correlation between SWS and overnight increases in z-score for false memory. A) for N- items only, B) for the interaction between N- and N+ items. Y-axis shows an amount of time spend in SWS in percentage. X-axis shows an overtime change of old response rate for B) n- and A) a difference of the over time change between n- and n+.

3.5. Control Experiment (Experiment 4)

The previous two experiments (experiments 2&3) examined sleep-dependent changes in face recognition memory, but we were interested to know if this change is limited to faces or can be observed in recognition memory of non-facial images as well. Compared to up-right faces, inverted (upside down) faces are regarded as non-facial images due to differences in their visual processing (see a review, Haxby, Hoffman & Gobbini, 2000). Thus, a control experiment was conducted to further examine if sleep-dependent changes in recognition memory were found with inverted facial images and whether these changes reflected those found with upright face stimuli.

3.5.1. Materials & Method (Experiment 4)

Participants and procedure

16 participants were recruited in a same way as the main experiment. All experiment procedures and design were same with two exceptions. Firstly, inverted facial morphs were used instead of upright ones and all participants were instructed to go back home to sleep in their own beds and to come back in the following morning to take a subsequent test.

3.5.2. Results (Experiment 4)

Over-time Veridical Recognition Change

In the control experiment (Experiment 4), veridical recognition did not change over-time. Thus, a one sample t-test revealed that the over-time change in veridical recognition (i.e., $m = 0.09$, $sd = 0.12$) was not significantly different from zero, $t(15) = 0.55$, $p > .05$ (see **Figure 3.5A**).

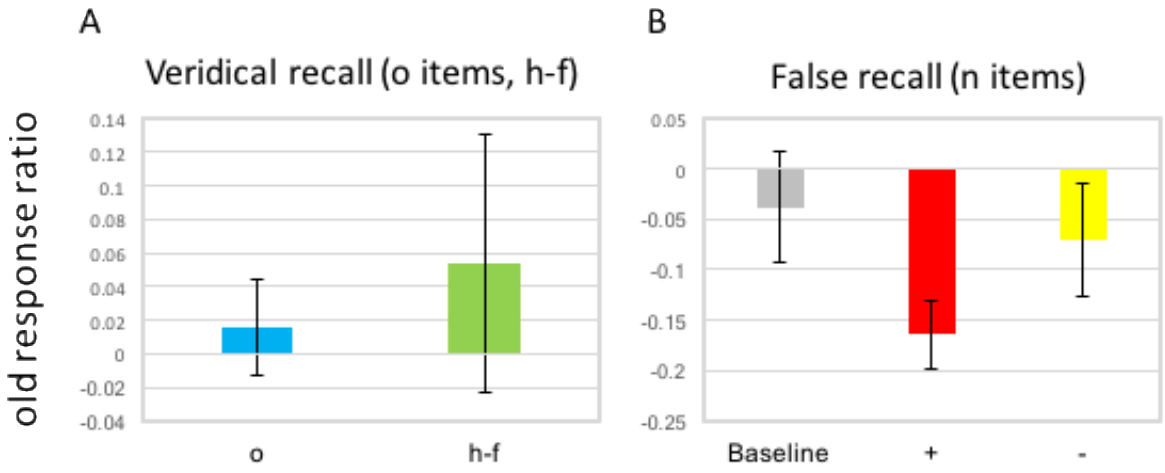

Figure 3.5. A) Veridical recall changes, B) False memory for the inverted face items. Here, two bar graphs of Experiment 4 (A and B) replicated results of Experiment 3. At first, as seen in Experiment 3, sleep maintained veridical memory (old response ratio on o items and even on overall memory accuracy (i.e. hits-false alarms)) in Experiment 4 (A) across sleep (ld-sd changes). Another result of Experiment 4 (B) shows that sleep also boosts false recognition of n-, compared to n+ across sleep (ld-sd changes).

Over-time False Recognition Change

Contrary to the results from the previous study (Tsujimura & Lewis, 2018), in the inverted face study (Experiment 4) false recognition decreased significantly over time. Then, a one way repeated measures ANOVA was run, by comparing over-time false recognition change on three types of new items (N+, N- and Nfar), and this revealed that there was no main effect of stimulus type ($F(2,30) = 2.3$, $p = .12$) (**Figure 3.5B**).

A planned one-tailed paired t-test (N- vs N+) was conducted to test for a larger false recognition change on N- compared to N+ which was expected, based on Experiment 3, and also on the results of the previous study (Tsujimura & Lewis, 2018). This comparison showed a marginal trend towards reduction in N- false

recognition being smaller (m = -.07, sd = .23) than reduction in N+ false recognition (m = -.16, sd = .14), $t(15) = 1.49$, $p = .08$ (**Figure 3.5B**).

3.5.3 General Discussion (Experiment 3 & 4)

In this present study, I replicated the findings of the previous study (Tsujimura & Lewis, 2018) that an over-night sleep supports increased false memories for items in a sparsely populated area of memory space (N- items), but not for items in a densely populated area of that space (N+ items). More importantly, I also found that sleep-dependent competition-free gist-based false recognition was significantly predicted by the time spent in SWS. Specifically, I found a positive correlation between the difference in false recognition z-score for items with several memory competitors (N+) and items with no memory competitors (N-). Interestingly, although I did find the expected overnight enhancement for veridical memory items, there was no correlation between the sleep-dependent veridical recognition change and SWS.

A strengthening and dual role of sleep

This study supported the suggestion that sleep, particularly SWS, plays a critical role in memory consolidation, and in the over-time transformation of recent memories. Yet, sleep is not regarded just as strengthening of memories that have actually been experienced before, but also reconstructing memories in two other ways. One way is to enhance the generalisation of recent memories to accept extending memory representations. Past work on gist-based false memories indicates that gist is already extracted at encoding either implicitly or explicitly (Howe, Garner,

Charlesworth & Knott, 2011). This process could be facilitated by interactions between overlapping neural replays during SWS (Lewis & Durrant, 2011). However, in a second form of restructuring, if the reactivated memories were competing rather than complimentary, for instance if a new memory conflicts with an existing schema – as in the case of newly learning about a penguin (who can swim but not fly) when you have a schema for birds that fly and don't swim. This process can lead to catastrophic interference (McClelland, McNaughton & O'Reilly, 1995) – a destruction of the existing schema. Under this logic, obtaining competing memories at encoding could potentially suppress the generation of gist-based false memories across sleep in densely populated memory spaces.

Similar to veridical memories that have actually been experienced, gist-based false memories undergo a process of memory consolidation over time. These memories are initially dependent on the hippocampus (e.g., Schacter, Norman & Koutstaal, 1998; Cabeza et al., 2001; Ciaramelli, Ghetti, Frattarelli & Làdavas, 2006) but gradually redistribute their neural underpinning to neocortical networks (e.g. Moscovitch & Melo,1997; Darsaud et al. 2010). In fact, gist-based false memories are more likely to be found after memory consolidation than veridical memories (McDermott, 1996; Seamon et al. 2002). This implies that gist extraction from memory consolidation may be resemble the pattern seen in the DRM paradigm (Payne et al., 2009; Diekelmann, Born & Wagner, 2010), where participants are trained on a list of semantically related words, and when later asked to retrieve them they tend to remember a 'gist' word that was not actually presented. Following this line of thought, even if the gist-based false memories are forgotten it may be possible to extract a similar gist again through a process of memory consolidation.

One approach to determining the mechanisms of time-dependent changes of the gist-based false memories is to test whether they are influenced by an interval of sleep. However, evidence of sleep's role in gist-based false memories is inconsistent. Some previous work showed that periods of sleep enhanced memory for both veridical memories and gist-based false memories created by training a list of words in the DRM paradigm (McKeon, Pace-Schott, & Spencer, 2012; Pardilla-Delgado & Payne, 2017; Payne et al. 2009). However, others did not show such enhancing effects of sleep and even showed a suppressing role of sleep on gist-based false memories (see Fenn et al., 2009; Lo, Sim & Chee, 2014). Interestingly, the free recall results are consistent with the hypothesis of gist memory promotion by sleep-dependent memory consolidation (Lewis et al., 2011; Nere et al., 2013), while the recognition results could support a suppressing role of sleep on the gist-based false memories under certain circumstances. This difference could suggest that memory consolidation in sleep works very differently for these two forms of memory.

Overall, my results support the idea that gist based false memory is sleep dependent. Because many sleep dependent memories are also hippocampal dependent, my results provide tentative support for the idea that these memories are also hippocampal dependent.

Are the effects specific to upright faces?

My control experiment (Experiment 4) supports the idea that the overnight impacts on false memory observed in the main experiment are not specific to face stimuli,

although effects with inverted face stimuli are weaker than those in upright faces. Nevertheless, there are three considerations to be discussed. First, only a sleep group was recruited without a wake group. This may lead to misinterpretation if recruiting additional data of a wake group shows the same pattern, showing relative larger false recognition in N- than N+, as the sleep group data did. Second, we may wonder why there was a lesser sleep effect on non-facial images. Is this because inverted faces were not typically processed holistically? Or is it because of difficulty in remembering upside down faces? Further future studies are required to understand these mechanisms. Finally, there has been a debate about whether inverted faces are actually not non-facial images (e.g. Richler, Mack, Palmeri & Gauthier, 2011), thus a claim here based on the control experiment result is not applicable. To challenge this problem, future sleep and recognition memory studies with typical non-facial images, such as cat, dog, car, should be examined.

References

Born, J., & Wilhelm, I. (2012). System consolidation of memory during sleep. *Psychological Research*. https://doi.org/10.1007/s00426-011-0335-6

Cabeza, R., Rao, S. M., Wagner, A. D., Mayer, A. R., & Schacter, D. L. (2001). Can medial temporal lobe regions distinguish true from false? An event-related functional MRI study of veridical and illusory recognition memory. *Proceedings of the National Academy of Sciences*, *98*(8), 4805–4810. https://doi.org/10.1073/pnas.081082698

Ciaramelli, E., Ghetti, S., Frattarelli, M., & Làdavas, E. (2006). When true memory availability promotes false memory: Evidence from confabulating patients. *Neuropsychologia*, *44*(10), 1866–1877. https://doi.org/10.1016/j.neuropsychologia.2006.02.008

Darsaud, A., Dehon, H., Lahl, O., Sterpenich, V., Boly, M., Dang-Vu, T., … Collette, F. (2011). Does Sleep Promote False Memories? *Journal of Cognitive Neuroscience*, *23*(1), 26–40. https://doi.org/10.1162/jocn.2010.21448

Diekelmann, S., Born, J., & Wagner, U. (2010). Sleep enhances false memories depending on general memory performance. *Behavioural Brain Research*, *208*(2), 425–429. https://doi.org/10.1016/j.bbr.2009.12.021

Ellenbogen, J. M., Hulbert, J. C., Jiang, Y., & Stickgold, R. (2009). The sleeping brain's influence on verbal memory: Boosting resistance to interference. *PLoS ONE*, *4*(1). https://doi.org/10.1371/journal.pone.0004117

Fenn, K. M., Gallo, D. a, Margoliash, D., Roediger, H. L., & Nusbaum, H. C. (2009). Reduced false memory after sleep. *Learning & Memory (Cold Spring Harbor, N.Y.)*, *16*(Gallo 2006), 509–513. https://doi.org/10.1101/lm.1500808

Glenville, M., Broughton, R., Wing, a M., & Wilkinson, R. T. (1978). Effects of sleep deprivation on short duration performance measures compared to the Wilkinson auditory vigilance task. *Sleep*, *1*(2), 169–76. https://doi.org/10.1093/sleep/1.2.169

Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences*. https://doi.org/10.1016/S1364-6613(00)01482-0

Hoddes, E., Zarcone, V., Smythe, H., Phillips, R., & Dement, W. C. (1973). Quantification of Sleepiness: A New Approach. *Psychophysiology*, *10*(4), 431–436. https://doi.org/10.1111/j.1469-8986.1973.tb00801.x

Howe, M. L., Garner, S. R., Charlesworth, M., & Knott, L. (2011). A brighter side to memory illusions: False memories prime children's and adults' insight-based problem solving. *Journal of Experimental Child Psychology*, *108*(2), 383–393. https://doi.org/10.1016/j.jecp.2010.08.012

Lewis, P. A., & Durrant, S. J. (2011). Overlapping memory replay during sleep builds cognitive schemata. *Trends in Cognitive Sciences*. https://doi.org/10.1016/j.tics.2011.06.004

Lo, J. C., Sim, S. K. Y., & Chee, M. W. L. (2014). Sleep Reduces False Memory in Healthy Older Adults. *Sleep*, *37*(4), 665–671. https://doi.org/10.5665/sleep.3564

Lutz, N. D., Diekelmann, S., Hinse-Stern, P., Born, J., & Rauss, K. (2017). Sleep Supports the Slow Abstraction of Gist from Visual Perceptual Memories. *Scientific Reports*, *7*, 42950. https://doi.org/10.1038/srep42950

McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neortex: Insights from the successes and failures of connectionist models of learning and memory.

Psychological Review, *102*(3), 419–57. https://doi.org/10.1037/0033-295X.102.3.419

McDermott, K. B. (1996). The Persistence of False Memories in List Recall. *Journal of Memory and Language*, *35*(2), 212–230. https://doi.org/10.1006/jmla.1996.0012

McKeon, S., Pace-Schott, E. F., & Spencer, R. M. C. (2012). Interaction of Sleep and Emotional Content on the Production of False Memories. *PLoS ONE*, *7*(11). https://doi.org/10.1371/journal.pone.0049353

Moscovitch, M., & Melo, B. (1997). Strategic retrieval and the frontal lobes: Evidence from confabulation and amnesia. *Neuropsychologia*, *35*(7), 1017–1034. https://doi.org/10.1016/S0028-3932(97)00028-6

Nere, A., Hashmi, A., Cirelli, C., & Tononi, G. (2013). Sleep-dependent synaptic down-selection (I): Modeling the benefits of sleep on memory consolidation and integration. *Frontiers in Neurology*, *4 SEP*. https://doi.org/10.3389/fneur.2013.00143

Olewuezi, N. P. (2011). Note on the Comparison of Some Outlier Labeling Techniques. *Journal of Mathematics and Statistics 7(4)*, 353-355.

Pardilla-Delgado, E., & Payne, J. D. (2017). The impact of sleep on true and false memory across long delays. *Neurobiology of Learning and Memory*, *137*, 123–133. https://doi.org/10.1016/j.nlm.2016.11.016

Payne, J. D., Schacter, D. L., Propper, R. E., Huang, L. W., Wamsley, E. J., Tucker, M. A., … Stickgold, R. (2009). The role of sleep in false memory formation.

Neurobiology of Learning and Memory, *92*(3), 327–334. https://doi.org/10.1016/j.nlm.2009.03.007

Richler, J. J., Mack, M. L., Palmeri, T. J., & Gauthier, I. (2011). Inverted faces are (eventually) processed holistically. *Vision Research*, *51*(3), 333–342. https://doi.org/10.1016/j.visres.2010.11.014

Schacter, D. L., Norman, K. A., & Koutstaal, W. (1998). The cognitive neuroscience of constructive memory. *Annu Rev Psychol*, *49*, 289–318. https://doi.org/10.1146/annurev.psych.49.1.289

Seamon, J. G., Luo, C. R., Kopecky, J. J., Price, C. A., Rothschild, L., Fung, N. S., & Schwartz, M. A. (2002). Are false memories more difficult to forget than accurate memories? The effect of retention interval on recall and recognition. *Memory & Cognition*, *30*(7), 1054–1064. https://doi.org/10.3758/BF03194323

Tsujimura, H., Lewis, P. A., (2018). Investigating the role of sleep on the gist-based face recognition memory.

Chapter 4

The role of sleep in curve-fitted face recognition memory

Hikaru Tsujimura, University of Manchester, UK

Sonja Kotz, University of Manchester, UK

Karen Lander, University of Manchester, UK

Timothy Stamper, University of Manchester, UK

Ellen M. Garlick, University of Manchester, UK

Penelope A. Lewis, University of Manchester, UK

4.1 Abstract

My prior work shows that sleep enhances false recognition of items that are free from competition, but not for items that compete with others. In this current study, I examined this phenomenon in detail using a finer grained 2D face space in which morphs were prepared on a much finer scale. This allowed me to examine how false recognition shifts away from competition and towards areas without competition in a more precise manner. Face morphs were only computed along the diagonal axis of the 2D space, and this allowed me to closely examine the shape of the memory distribution associated with each learned 'old' face. I was then able to fit a polynomial curve to each distribution and check for shifts in both the sharpness of the curve and how the curve shifted on the x axis (e.g. towards or away from the learned face). This analysis revealed that a night of sleep is associated with a marked sharpening ($p=0.04$) of the distribution associated with old faces, as well as a concomitant shift of this distribution away from areas of the face-space associated with competition and towards areas that are free of competition), $p=0.001$.

4.2. Introduction

Sleep is important for memory formation, particularly in the processes of generalising and abstracting gist of memories (Lewis & Durrant, 2011; Stickgold, & Walker, 2013), but there has been little evidence demonstrating the full mechanisms. To visualise the generalising processes, two previous studies (Tsujimura & Lewis, 2018; Tsujimura, Kotz, Lander, Tafuro and Lewis, 2018) developed a novel two-dimensional space by morphing the face of an actor along two features (age and gender) and assigning those morphing scales as axes in a 2D graph. Participants were trained until they learned to distinguish between individual face morph images selected from this 2D morph space. They were then shown a mixture of 'old' (trained) and 'new' (untrained) face morph stimuli from this space and asked to indicate if they were 'old (previously seen)' or 'new' (not previously seen). During the subsequent old/new recognition memory test, participants gave a higher proportion of old responses on old items (hits) after sleep compared to an equivalent amount of wakefulness, supporting the idea that sleep benefits veridical memory. Interestingly, however, the proportion of false alarms (old responses to new items) also changed across sleep, and was lower when the new items in question were located between old items in the constructed 2D morph space. Conversely, post-sleep false alarms to new items that were not near any older items were higher after sleep.

These findings suggest that overnight sleep enhances the over-generalisation of face recognition memories into competition-free face perceptual spaces, but not into perceptual spaces where existing face memories compete. These studies also showed that these effects are not specific to upright faces (see a work by Tsujimura,

Kotz, Lander, Tafuro and Lewis, 2018), and may thus be a more general property of learned images. Finally, they showed that these effects are predicted by the amount of time spent in slow wave sleep (SWS).

Although these results undoubtedly reveal something about the way memories are processed over sleep, they are nevertheless difficult to interpret. This is partially because the paradigms used in Tsujimura & Lewis, (2018) and Tsujimura, Kotz, Lander, Tafuro and Lewis (2018) differed slightly, e.g. in terms of amount and placement of images in the 2D face morph space (see Tsujimura & Lewis, 2018 and Tsujimura, Kotz, Lander, Tafuro and Lewis, 2018 for more details). More importantly, an ambiguity exists regarding how an overall representation of face memories alters across sleep. The works of Tsujimura and Lewis, (2018) and Tsujimura, Kotz, Lander, Tafuro and Lewis (2018) consistently show that more items in the non-competitive perceptual space are endorsed and less items in the competitive space are endorsed as seen after sleep. However, does this really mean that sleep shifts the overall representation of face memory across perceptual space?

The above questions stem from a structurally unsolved difficulty in the paradigm of the previous studies (see Tsujimura & Lewis, 2018 and Tsujimura, Kotz, Lander, Tafuro & Lewis, 2018). Specifically, these studies were unable to show how an "overall" representation of face memory is altered by a period of sleep. Thus, either four or five face stimuli were selected from the entire 2D face morph space for initial training. This paradigm compared the sleep effects in densely and sparsely populated face morph space on the two discrete 'conditions,' 1) post-sleep false alarms on new items locating at the highly competitive areas and 2) post-sleep false

alarms on new items locating outside of the competitive areas, but was not able to determine how memory representations actually shift in this space. In order to assess this explicitly, it would be better to represent the recognition memory performance in a unifying measure.

Accordingly, this new study set out to more explicitly test this question with a more systematically controlled set of face stimuli in the 2D space which can draw an overall representation of face recognition memories as one unifying representation. To do so, the new work devised a third version of the old/new face recognition memory paradigm with face stimuli selected from the invented 2D face morph space. In this version, only two 'old' items were trained, and these were placed diagonally across the age and gender face morphing space, while 'new' items were placed all along the remaining spaces of the diagonal line (see **Figures 4.1** and **4.2**). This new pattern of old item training along one diagonal line in the 2D face morph space helped participants to develop memory representations of the two old (trained) items and engendered strong competition between these old items in the space between these old items, while there was much less competition in the space on the outside of each old item (e.g. near one old item but not the other). This way, one unifying curve could be drawn to capture the distribution of memory strength along the diagonal line in our 2D face-space. Furthermore, unlike the studies of Tsujimura & Lewis, (2018) and Tsujimura, Kotz, Lander, Tafuro and Lewis (2018), the 5 fine-grained morphs between each 'perceptual' face representation were all considered. In this manner, the new task set up allowed us to visualise even minor shifts in memory, and also to examine the shape of the overall memory distribution both before and after sleep.

4.3. Materials & Method (Experiment 5)

Participants

42 healthy Caucasian participants (14 males: age 22.64±5.97 (mean±SD), range 18-40 years old, 28 females: age 21.36±3.77 (mean±SD), range 18-39 years old) were recruited from posters, e-mail and online advertisement. Participants were randomly assigned into either a sleep (n = 21, 6 males : age 21.33±4.80 (mean±SD), range 18-31 years old, 15 females: age 21.07±1.79 (mean±SD), range 18-26 years old) or a wake group (n = 21, 8 males : age 23.63±6.86 (mean±SD), range 18-40 years old, 13 females: age 21.69±5.30 (mean±SD), range 19-39 years old). One participant was excluded from the sleep group due to them reporting sleep deprivation on the night before the experiment, at the end of experiment. One participant was excluded from the wake group due to failing to follow experiment procedures. In total, 20 participants in the sleep group (5 males : age 19.40±0.89 (mean±SD), range 18-20 years old, 15 females: age 21.07±1.79 (mean±SD), range 18-26 years old) and 20 participants in the wake group (7 males : age 24.14±7.24 (mean±SD), range 18-40 years old, 13 females: age 21.69±5.30 (mean±SD), range 19-39 years old) were remained for data analysis. They were screened for a history of neurological and psychiatric diseases and sleep disorders. They were asked to abstain from caffeine and alcohol 24 hours prior to experiment. This study was approved by the University of Manchester Research Ethics Committee. All participants were instructed about the experiment in detail and signed a consent form.

Stimuli selection in the finer grained two-dimensional face morph scale

In this experiment, a modified version of the 2D face morph scale initially developed in the studies of Tsujimura and Lewis, (2018) and Tsujimura, Kotz, Lander, Tafuro and Lewis (2018) was employed. Briefly, the previous studies of Tsujimura and Lewis, (2018) and Tsujimura, Kotz, Lander, Tafuro and Lewis (2018) morphed an actor's face using the features of age and gender in 6 stages (ie a starting point as 0^{th} and then 10^{th}, 20^{th}, 30^{th}, 40^{th} and 50^{th} morphing point in FaceGen Modeller 3.5, http://www.facegen.com). One feature morphing scale was used as the vertical axis and the other feature morphing scale as the horizontal axis, which, as a result, created a 6x6 2D face morph scale. This process was repeated for the seven other actors' faces to create a total of eight 6x6 2D face morph scales (one per actor).

The current experiment morphed each face of eight actors in 25 stages (i.e. a starting point as 1^{st} and then 3^{rd}, 5^{th}, 7^{th}, …, 45^{th}, 47^{th} and 49^{th} morphing point in the FaceGen software) in each dimension, to create eight 25x25 finer grained 2D face morph scales (see an example of a finer grained 2D face morph scale in **Figure 4.1A**). In this new (modified) 25x25 face morph scale, a unit of face morphing (e.g. from 1^{st} to 3^{rd} FaceGen morphing point) was five times more subtle in facial manipulation than that of the past 6x6 face morph scale (e.g. from 0^{th} to 10^{th} FaceGen morphing point), where the 10 FaceGen face morphing points were the minimal degree of facial manipulation participants can detect a change of face identity perceptually in the previous studies (Tsujimura & Lewis, 2018; Tsujimura, Kotz, Lander, Tafuro & Lewis 2018). From each of the five times finer grained 2D face morph scale, 25 items were selected as face stimuli in this present experiment.

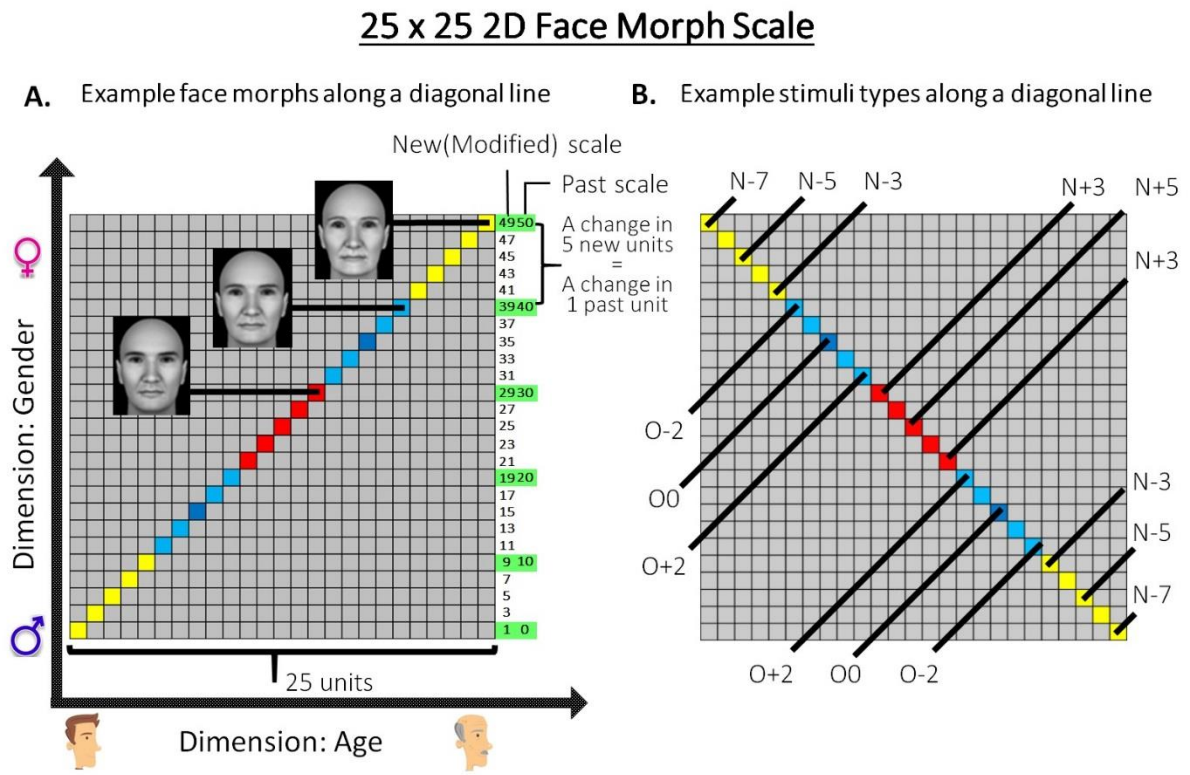

Figure 4.1. **An example of finer grained (25x25) 2D face morph scales.** (A) The left diagram showed a finer grained 25x25 2D face morph scale with example face morphs along a diagonal line. Green highlighted boxes indicated corresponding FaceGen morphing point for each face morph in the case of the new(modified) 25x25 face morph scale and in the case of the past 6x6 face morph scale. (B) The right diagram showed five major types (Yellow: N-, Light Blue: O- and O+, Dark Blue: O0, Red: N+) of face stimuli in the new 2D morph scale with a number next to it. The number indicated how far stimuli were away from the O0 stimuli.

Of the eight invented finer grained 2D face morph scales, a random selection of four 2D face morph scales were assigned to each participant. Those selected 2D face morph scales of four actors or identities were termed as ID1 (identity 1), ID2, ID3 and ID4. Two randomly selected IDs (e.g. ID1 and ID4) from the four IDs were used in one of two experimental sessions and the other two IDs (e.g. ID2 and ID3) were used in the other session (see the **design** section for detailed experimental procedures). In this study, a set of 25 face morphs along a right-top-left-bottom diagonal line were recruited from one ID (e.g. ID1) and the other set of 25 face morphs along a left-top-right-bottom diagonal line were recruited from the other ID

(e.g. ID4) in the same session (e.g. the 25 red, blue and yellow dots along the two diagonal lines in **Figure 4.1**). This was critical to control confounds relating to the position of the stimuli in the 2D face morph scales.

Stimuli types

In each set of the 25 face morph selection, there were five major types of face stimuli: (1) the O0 types (the dark blue dots) that were actually trained at initial encoding task, (2) the N- types (the yellow dots) that were not trained at encoding and spatially neighbouring to only one of the two O0s, (3) the N+ types (the red dots) that were not trained at encoding and sandwiched spatially between the two O0s, (4) the O- types (a half of the light blue dots closer to N-) that were not shown at encoding and spatially sandwiched between only an O0 and a N- and (5) the O+ types (the other half of the light blue dots closer to N+) that were not shown at encoding and sandwiched spatially between an O0 and a N+. In addition to this, a number next to the stimulus type indicated how far those stimuli were distanced from the O0 types. Please see **Figure 4.1B** for schematic illustration.

Design

In this experiment, participants were randomly assigned into either a wake or a sleep group. Participants in both groups were asked to attend two sessions in the experiment. Each session was mainly composed of encoding, subsequent retrieval and a delay between the two memory tasks. In one session, the delay lasted 10 minutes (the short delay session), and in the other session the delay lasted 12 hours

(the long delay session). The order of the short and the long delay sessions in this experiment was pseudo-randomly assigned in each group. Thus, both sleep and wake groups had half participants (i.e. n = 10 in sleep and n = 10 in wake) starting with the short delay session and then the long delay session, and vice versa. The wake group started at 9:00 at a quiet testing room at the sleep lab. This meant that, half of the wake group completed the first session and encoding of the second session in the morning. They were then free to go back to their daytime activity and asked to come back to sleep lab again at 21:00 to complete subsequent retrieval in the second session. The other half of the wake group went through in the opposite way, completing encoding of the first session at 9:00 in the morning. They then came back to the sleep lab at 21:00 to complete the subsequent retrieval of the first session and the second session in the evening. The sleep group followed same procedures as the wake group did, but started at 21:00 and then they were instructed to sleep overnight in a bed at the sleep lab after completion of encoding in the long delay session. Their sleep was monitored by polysomnography (PSG). Participants were instructed to complete their sleep rituals, such as brushing teeth, changing into pyjamas, etc, prior to setting up the PSG system. After the PSG setting, they lay down on a bed and the room light was turned off to sleep. The wake group was asked not to take a nap during the long delay. At the beginning of the first session or after the long delay, participants completed questionnaires of Stanford Sleepiness Scale (SSS; Hoddes et al., 1973) and Karolinska Sleepiness Scale (KSS; Glenville, Broughton, Wing, & Wilkinson, 1978) to check their vigilance. See **Figure 4.2A** for detailed instructions in this experiment.

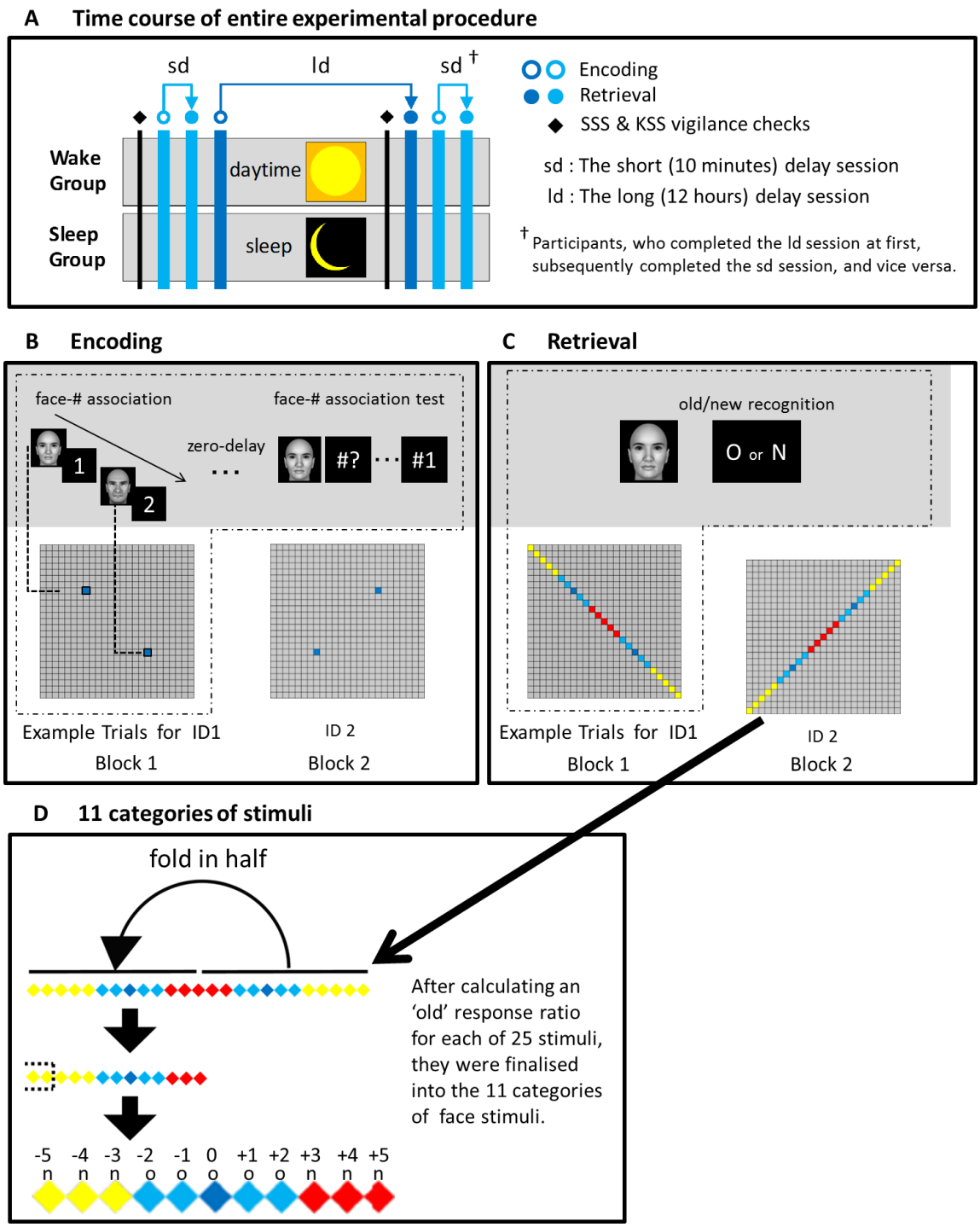

Figure 4.2. A Schematic Illustration of Experimental Procedures of Experiment 4. Here drew an illustration of (A) a time course of the entire experiment procedure of Experiment 4. The experimental procedure was mainly composed from (B) an encoding and (C) a retrieval tasks with a delay between them. During the encoding task (B), participants were trained to memorise and distinguish between two OLD stimuli (two blue boxes) of a same ID by taking a face-# association task and a zero-delayed test. During the retrieval task (C), participants took an old/new face recognition test and a # recognition test. Each task was run for two blocks; each block used an ID. See the **Design** and the **Tasks** section for more details. After collecting participant's old/new responses for the

25 stimuli, they were summed into 11 types (categories) of stimuli, which were analysed (D). N-6 and N-7 were removed for from the analysis in order to balance the number of N- and N+ stimuli.

Tasks

The current experiment slightly modified the tasks of encoding and retrieval from the ones used in the previous studies (Tsujimura & Lewis, 2018; Tsujimura, Kotz, Lander, Tafuro & Lewis 2018). Namely, in the previous studies, it was necessary to rotate the whole 2D face space by 90 degrees four times in order to control for confounds relating to the position of the stimuli in this space (e.g. right top, right bottom, left top and left bottom). Because of this, there were four blocks of trials in the previous studies, employing an ID per block in both the encoding and retrieval tasks. With this new improved design only two conditions (one diagonal and the other orthogonal to the first diagonal line) were necessary to be rotated. Thus, in the current experiment there were only two blocks of trials, employing one ID per block. The procedure of the encoding task was the same as that used in Tsujimura, Kotz, Lander, Tafuro and Lewis, (2018). Similarly, procedures of the retrieval task in the current experiment were same as that of Tsujimura, Kotz, Lander, Tafuro and Lewis, (2018). Participant's recognition memory of the number assigned to each face stimuli (either 1 or 2) was also tested. See **Figure 4.2B** and **4.2C** for schematic illustrations of those tasks.

Face recognition memory

In this current experiment, **'**old' responses for the 25 face stimuli were collected and an 'old' response ratio was calculated for each face stimulus ('old' responses vs total

'old'/'new' responses for that stimulus). Next, the old response ratios for the 25 face stimuli were averaged to collect the finalised old response ratios for the 11 categories of face stimuli (i.e. N-5, N-4, N-3, O-2, O-1, O0, O+1, O+2, N+3, N+4, N+5), which were used as parameters of face recognition memory (See a diagram of the 11 categories of stimuli in **Figure 4.2D**).

To ensure that the sleep and wake group showed similar results at baseline (i.e. sd session), I first calculated a 2 (sleep, wake) x 11 (category) two-way ANOVA for the (finalised) old response ratios at baseline. Next, I compared performance change across sleep and wake using a 3 way ANOVA with factors of delay (sd, ld) x sleep group (sleep, wake) x category (11 types). In both cases post-hoc t-tests were applied. In the analysis of the sd session only, post-hoc t-tests were applied to compare wake and sleep groups in all categories, and again between all pairs of categories within a group. In the analysis of change over sleep or wake these were applied to categories where I expected a difference based on findings of the previous studies (Tsujimura & Lewis, 2018; Tsujimura, Kotz, Lander, Tafuro & Lewis 2018). In other words, t-tests of the closest and farthest N- to the O0 (i.e. N-3 and N-5) were analysed in the sleep group to examine whether the sleep-dependent increases of the old response ratio on N-, which was found in the previous studies (Tsujimura & Lewis, 2018; Tsujimura, Kotz, Lander, Tafuro & Lewis 2018), was specific to either category. Similarly, t-tests of the exact O (i.e. O0) and farthest O (i.e. O-2 and O+2) were analysed in the wake group to examine whether the wake-dependent decreases of the old response ratio on O, which was found in the previous studies (Tsujimura & Lewis, 2018; Tsujimura, Kotz, Lander, Tafuro & Lewis 2018), was specific to either category. All t-tests were Bonferroni corrected for multiple

comparisons when the number of the comparisons were under 5; the Sidak correction was employed for more comparisons.

Data fitting into a second-order polynomial model (i.e. $y = ax^2+bx+c$, $a<0$)

Based on the results from the previous studies (Tsujimura & Lewis, 2018; Tsujimura, Kotz, Lander, Tafuro & Lewis 2018), I hypothesized that sleep plays a critical role in the time-dependent modulation of participant's overall mental representation of face recognition memory (which is represented as a grey inverted U-curve in **Figure 4.3A**). Specifically, I proposed that a role of sleep is to shift the participant's overall mental representation (e.g. movement of the whole curve away from memory competitors as shown in **Figure 4.3C**) and to sharpen the representation (e.g. moving towards a more precise curve as shown in **Figure 4.3B**).

To test this hypothesis, the collected old response ratios for the corresponding 11 categories (i.e. N-5, N-4, N-3, O-2, O-1, O0, O+1, O+2, N+3, N+4, and N+5) in each session (i.e. sd and ld) and group (wake and sleep) were first fit to an inverted U-shaped second-order polynomial model (i.e. $y = ax^2+bx+c$, $a<0$). In the inverted U-shaped parabolic model, I assigned values of x from -0.5 to 0.5 in steps of 0.1, to quantify how far each stimulus category was from the old item (O0) and from negative values in x-axis (e.g. N-/O-) and from positive values (e.g. N+/O+), e.g. O0 was 0, O+1 was .1, 0-1 was -.1, etc. Values of y were the old response ratios for each corresponding stimulus category, and the formula was used to calculate the three variables 'a', 'b', and 'c'. Namely, the value of "a" indicated a degree of sharpness of the inverted U curve. The larger the negative value of "a" was the more

its curve was sharper. To find the maximum, I calculated the derivative (i.e. y' = 2ax+b, a<0) of this inverted U-shaped second-order polynomial model for each session and group. Finally, I solved the formula in the case of 0 = 2ax+b or x = -b/2a and applied the obtained x value as well as the a, b and c values to the formula of y = ax^2+bx+c to get a maximal (or minimal) y point of this curve. See a schematic illustration of my proposed hypothesis in **Figure 4.3**.

Once all curves had been fitted, we tested that the curves had a similar shape after the sd by using a t-test to compare values of a in the sleep and wake groups, see **Figures 4.5B** and **4.5C**. Next, I used parameters from the curve fitting to determine how changes in memory distribution across the long delay differed between the wake and sleep groups. For this purpose, I used a 2 (sd, ld) x 2 (sleep, wake) way ANOVA, first to examine the curve shape 'a', see **Figure 4.7**. Next, to examine differences in how the maximum point changed across wake and sleep, I performed a 2 (wake, sleep) x 2 (x, y coordinate) way ANOVA, see **Figure 4.7**.

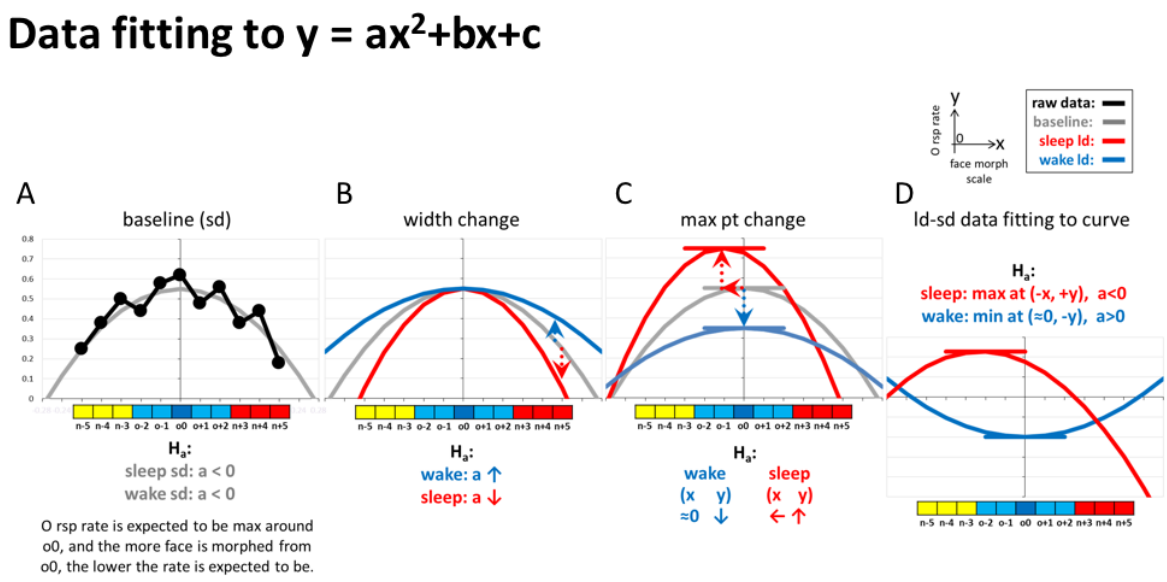

Figure 4.3. Data fitting into a parabolic curve of y = ax2+bx+c, (a<0). This diagram illustrated my hypothetic model of a sleep role in time-dependent modulation of participant's mental representation of face recognition memory. All drawn curves were estimated based on the results from previous studies (Tsujimura & Lewis, 2018; Tsujimura, Kotz, Lander, Tafuro & Lewis 2018). X-axis is a scale of distance of each stimulus category morphed away from the O0 and y-axis is a scale of old response ratios for corresponding category.

Polysomnographic (PSG) data

Polysomnographic (PSG) recordings were obtained and digitized at a sampling rate of 200Hz using the Embla N7000 system (Medcare-Embla®, Reykjavik, Iceland). The PSG electrodes were attached at F3, F4, Fz, C3, C4, Cz, O1, and O2, according to the international 10-20 system, as well as left and right mastoids (A1 and A2) as the EEG references, and the forehead as a ground, left and right eyes (E1 and E2) and left and right sides and bottom of chin. Electrode impedances were verified to be lower than 5kΩ. Collected PSG data was scored by trained experimenters according to the AASM manual (American Academy of Sleep Medicine, Westchester, IL), using RemLogic 1.1 software, to measure distribution of sleep stages for each participant. A summary of the PSG data was shown in **Table 4.1**.

4.4. Results (Experiment 5)

Sleep

	Sleep time (%)
S1	1.78 +/- 1.85
S2	66.27 +/- 9.72
SWS	19.14 +/- 5.03
REM	19.56 +/- 8.26

Table 4.1. Mean duration (in percentage) and standard deviation of sleep stages. SWS is made up from a combination of N3 & N4

Face recognition memory accuracy

First, to test the impact of sleep on memory performance, I analysed accuracy (ie hits – false alarms). 'Hits' were the old response ratio of O0 and the 'false alarms' were an average old response ratio across the remaining 10 categories. A 2 (group: wake or sleep) x 2 (session: sd or ld) mixed ANOVA on the face recognition memory accuracy was conducted and this revealed that there was a significant interaction effect of group x session, (F(1, 38) = 4.91, p = .033). Interestingly, face recognition improved across a night of sleep (ld-sd, m = .023, sd = .12) while it decayed across a day of wake (ld-sd, m = -.092, sd = .20), and a post-hoc t-test confirmed that this difference was significant t(30.54) = 2.22, p = .034 (**Figure 4.4**). There was no main effect of group or session, (group, F(1, 38) = .22, p = .65, session, F(1, 38) = 1.75, p = .19, respectively). These results supported the more general finding that sleep

confers an advantage in enhancing (face recognition) memory compared to an equivalent amount of wakefulness (e.g. Walker & Stickgold, 2004).

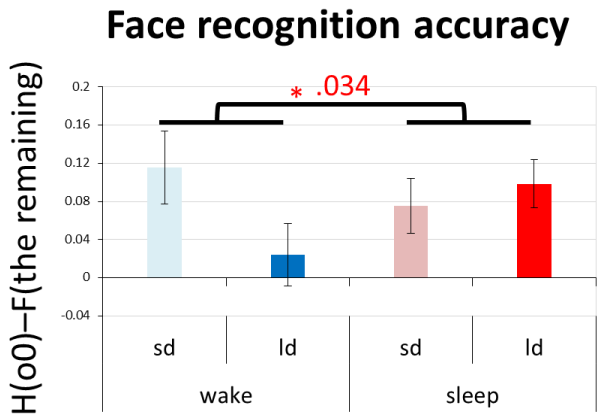

Figure 4.4. Face recognition memory accuracy. Participant's memory accuracy (hits - false alarms) of old/new face recognition memory test in sessions (sd vs ld) and groups (wake vs sleep). Y-axis showed values of hits (proportion of old response ratios for old stimuli 'O0') – false alarms (averaged old response ratios across all 10 categories of new stimuli).

Analysis of old response ratio

To confirm that both wake and sleep groups showed similar patterns after a short delay (sd session), I first performed a 2 (sleep, wake) x 11 (category) ANOVA on old response ratios in the sd session, see **Figure 4.5A**. This revealed that there was a main effect of category, $F(10, 380) = 3.77$, $p < .001$, but no main effect of group, $F(1, 38) = 0.0002$, $p = .99$, or interaction of group x category, $F(10, 380) = .55$, $p = .85$. As there were no main effect of group or interaction of group x category, this sanity test verified that both wake and sleep group followed a similar pattern after a short delay.

Next, I searched for differences in the way these response distributions changed across wake and sleep using a 2 (wake, sleep), x 2 (sd, ld), x 11 (category) ANOVA.

This revealed a main effect of category ($F(10, 380) = 5.65$, $p < .001$) and a three way interaction of group x category x delay ($F(10, 380) = 2.08$, $p = .025$), but there were no main effect of group, $F(1, 38) = .52$, $p = .48$, or delay, $F(1, 38) = .07$, $p = .79$, and no interaction of group x delay, $F(1, 38) = 1.09$, $p = .30$, group x category, $F(10, 380) = .23$, $p = .99$, or delay x category, $F(10, 380) = 1.21$, $p = .29$, see **Figures 4.5B** and **4.5C**.

According to prior works by Tsujimura and Lewis (2018) and Tsujimura, Kotz, Lander, Tafuro and Lewis, (2018), it was consistently found that veridical memory (i.e. correct old responses on the O items) deteriorated across a delay of wake. In addition, the same works showed that false memory (i.e. incorrect old responses on unstudied items) on N-, in which the unstudied items were away from competition of the O items, was enhanced across a night of sleep. Accordingly, I used planned t tests on O-2, O0 and O+2 in the wake group (**Figure 4.5B**) to examine where deterioration of veridical memory happened across a day of wake. Similarly, I focused on false memory of N-3, as the closest N- stimuli to O0, and N-5, as the farthest N- to O0, in the sleep group (**Figure 4.5C**) to examine which false memory was enhanced across a night of sleep. In the wake group, only old response ratios on O0 was decreased between sd and ld, (wake-O0, $t(19) = -2.29$, $p = .03$, wake-O-2, $t(19) = .58$, $p = .57$, wake-O+2, $t(19) = -1.41$, $p = .18$). In the sleep group, this revealed that only old response ratios on N-3 increased between sd and ld, $t(19) = 2.70$, $p = .01$. After Bonferroni correction, the decreasing of old response ratio O0 was no longer significant.

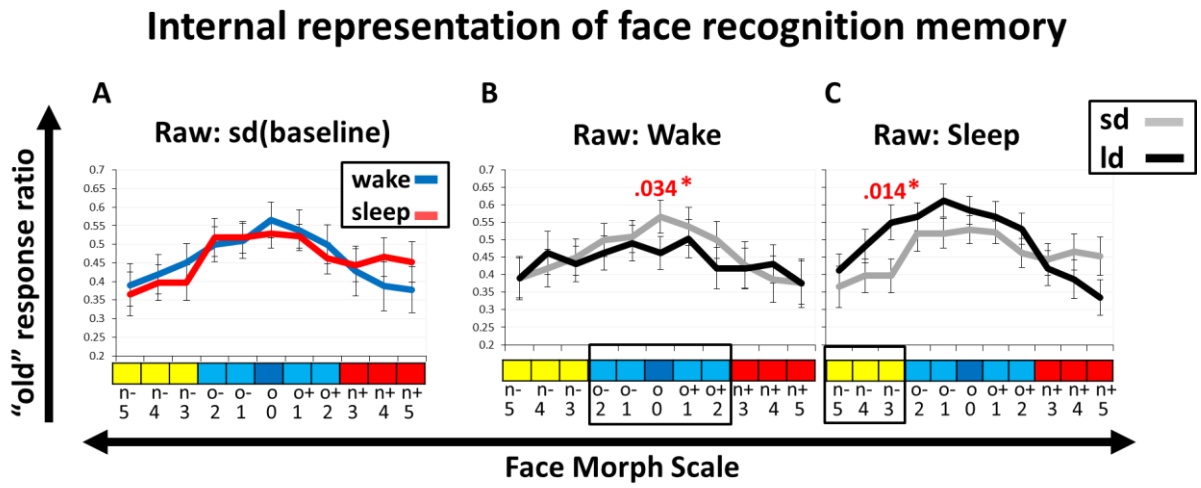

Figure 4.5. Analysis of old response ratio. A) Old response ration after a short delay for wake and sleep groups. B) Old response ration for short and long delays in the Wake group. C) Old response ratio for short and long delays in the sleep group. Thick highlighted black boxes are the expected areas where the previous studies (Tsujimura & Lewis, 2018; Tsujimura, Kotz, Lander, Tafuro & Lewis 2018) keep finding memory deterioration in wake group and over time increases of false recognition in sleep group. Thus, these areas were tested with post hoc tests after ANOVA, in order to specific which specific area is deteriorated in wake group and increasing false recognition in sleep group.

Curve fitting for a fine grained analysis at short and long delay

In the curve fitting analysis, a two-way ANOVA (2 group x 11 category) and a t-test were used to make sure that estimated old response ratios from the fitted curve and the curve parameter "a" did not differ in the sleep and wake groups in the sd session (**Figure 4.6A**). The two-way ANOVA revealed that there was a main effect of category, $F(10, 380) = 4.21$, $p < .001$, but no main effect of group, $F(1, 38) = .0002$, $p = .99$, or interaction of group x category, $F(10, 380) = .46$, $p = .92$, similar to the above analysis on raw data. This sanity test verified that both wake and sleep groups followed a similar pattern after a short delay. Similarly, a t-test analysis on the curve parameter "a" revealed that this parameter did not differ between wake ($m = -.16$, $sd = .20$) and sleep ($m = -.11$, $sd = .18$) groups a short delay, $t(38) = .79$, $p = .43$.

Next, I tested that the curves had a similar shape after the sd by using a t-test to compare values of “a” in the sleep and wake groups. This revealed no difference ($t(38) = .79$, $p > .05$) (**Figure 4.7A**). I also tested whether the curve shapes of both groups were inverted, indicated by parameter a being lower than zero, ($p<0.01$ in both sleep and wake groups). Inspection of the curves in **Figures 4.6B** and **4.6C** suggests that sleep leads to a sharpening of the curve, while wake leads to a flattening. To test this, I used a 2 (sd, ld) x 2 (sleep, wake) ANOVA, with “a” as the dependent variable, to examine how changes in memory distribution across the long delay differed between the wake and sleep groups. This revealed an interaction effect of delay x group, $F(1, 38) = 6.90$, $p = .01$, but no main effects of delay and group, (delay, $F(1, 38) = .14$, $p = .71$, group, $F(1, 38) = .73$, $p = .40$, respectively) (see **Figure 4.7A**). Matching with my hypothesis, “a,” the curve parameter, reduced across a night of sleep (m = -.11, sd = .20), indicating curve sharpening, and increased across a day of wake, indicating that curve became flattened over wake-time, (m = .08, sd = .25), $t(38) = 2.63$, $p = .012$. Further post hoc t tests revealed that this difference was driven by curve sharpening (reduction in ‘a’) between ld and sd session in the sleep group, $t(19) = -2.42$, $p = .03$ (Bonferroni correction, $p = .05$).

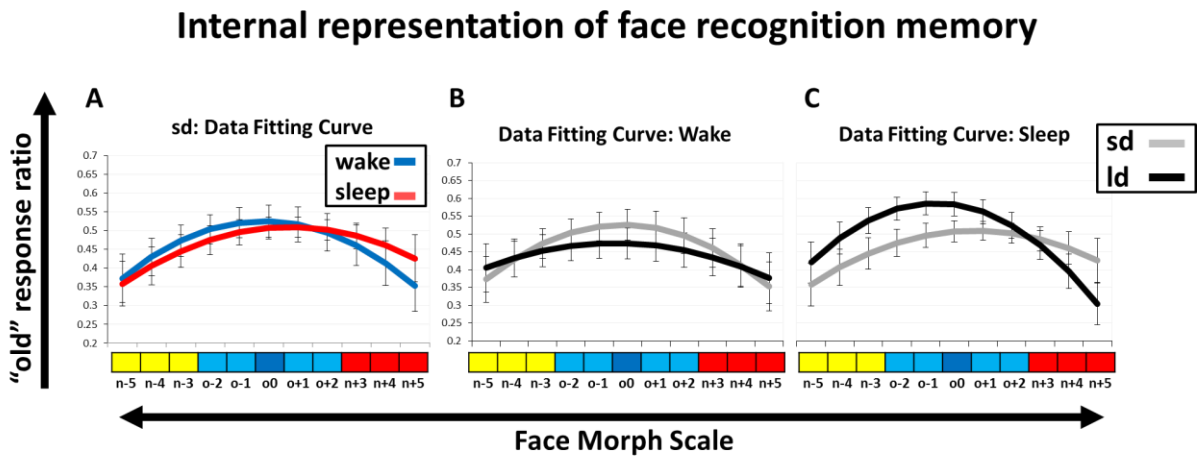

Figure 4.6. Analysis of polynomial curve fitting. A) Curves fit to data of short delay for wake and sleep groups. B) Curves fit to short and long delays in the Wake group. C) Old response ratio for short and long delays in the sleep group.

Inspection of the curves in **Figures 4.6B** and **4.6C** also suggests that sleep leads to a leftwards shift (away from the competitive N+ space and towards the non-competitive N-space), while wake leads to no shift in the position of the curve maximum on the y axis. To test this, I examined the differences in how the maximum point changed across wake and sleep, using a 2 (wake, sleep) x 2 (x, y coordinate) way ANOVA, see **Figure 4.7**. Several outliers (more than 2 standard deviations from the mean in x or y coordinate) were removed. The ANOVA then revealed a main effect of x,y coordinate, $F(1, 23) = 6.47$, $p = .018$, and an interaction between group and coordinate ($F(1, 31) = 8.53$, $p = .006$). Post-hoc t-tests showed that, in the sleep group, this was driven by a leftward (negative) shift in the X coordinate ($t(12) = -4.49$, $p = .001$) and an upward (positive) shift in the Y coordinate ($t(12) = 3.37$, $p = .006$), while there was no significant shift in the wake group.

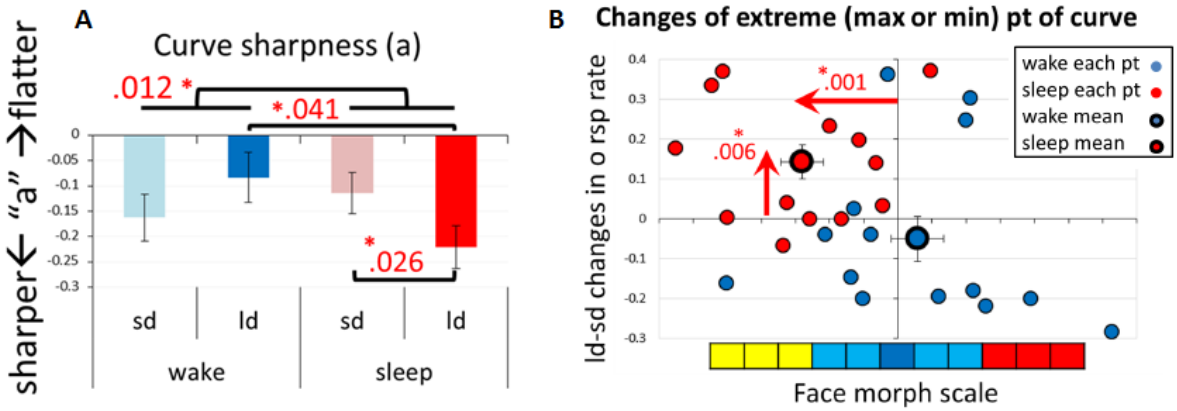

Figure 4.7. Statistical comparisons of curve fitting parameters. A) Results for the parameter 'a' which represents curve shape. B) Results for the maximum point of the curve, illustrating a leftward shift in the X coordinate and an upward shift in the Y coordinate.

4.5. Discussion (Experiment 5)

This experiment replicates the results of the previous studies (Tsujimura & Lewis, 2018; Tsujimura, Kotz, Lander, Tafuro & Lewis 2018) in that overnight sleep leads to increased false alarms for images falling within a comparatively empty portion of perceptual memory (N-), but not for images falling within an area of perceptual memory where there was competition (N+). Importantly, this study also extends my prior findings by illustrating that there is a shift in the whole distribution away from N+ and towards N- (e.g. the response curves shifted towards the left). It also shows a sharpening of the represented curve across sleep but not wake (where representations are flattened).

This fine-grained analysis provides additional insight into the process that is occurring during consolidation. It seems that, over a period of sleep, the representation of remembered information actively shifts away from the more

crowded (N+) memory space and into the less competitive (N-) space. Conversely, across a period of wake the memory simply seems to become more blurred.

These results support the notion of 'reconstructive' changes to memory over time, as proposed by Schacter, Guerin and Jacques (2011), and build on this by proposing that memory reconstruction across sleep involves shifting away from competitive memories and over-generalising memories/knowledge toward less competitive spaces. These results could also be in keeping with the CLS model if we think of the N+ items as competing memories, while N- items are non-competing memories. Thus, the competitive N+ memories (e.g. Penguin) require an extensive time to integrate with existing knowledge, whereas N- memories (e.g. cardinal) are, at least falsely, recognised as existing knowledge after sleep (McClelland, McNaughton & O'Reilly, 1995; McClelland, 2013).

Interestingly, this is contrary to a literature that increasingly talks about a role for sleep in integration (e.g. see Durrant, Taylor, Cairney & Lewis, 2011; Lewis & Durrant, 2011; Tamminen et al., 2010). My current result supports the idea that sleep also plays a critical role in distinguishing new learning from existing memory. Thus, learned representations are shifted away from competitor memories after sleep.

References

Durrant, S. J., Taylor, C., Cairney, S., & Lewis, P. A. (2011). Sleep-dependent consolidation of statistical learning. *Neuropsychologia*, *49*(5), 1322–1331. https://doi.org/10.1016/j.neuropsychologia.2011.02.015

Glenville, M., Broughton, R., Wing, a M., & Wilkinson, R. T. (1978). Effects of sleep deprivation on short duration performance measures compared to the Wilkinson auditory vigilance task. *Sleep*, *1*(2), 169–76. https://doi.org/10.1093/sleep/1.2.169

Hoddes, E., Zarcone, V., Smythe, H., Phillips, R., & Dement, W. C. (1973). Quantification of Sleepiness: A New Approach. *Psychophysiology*, *10*(4), 431–436. https://doi.org/10.1111/j.1469-8986.1973.tb00801.x

Lewis, P. A., & Durrant, S. J. (2011). Overlapping memory replay during sleep builds cognitive schemata. *Trends in Cognitive Sciences*. https://doi.org/10.1016/j.tics.2011.06.004

McClelland, J. L. (2013). Incorporating rapid neocortical learning of new schema-consistent information into complementary learning systems theory. *Journal of Experimental Psychology: General*, *142*(4), 1190–1210. https://doi.org/10.1037/a0033812

McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychological Review*, *102*(3), 419–57. https://doi.org/10.1037/0033-295X.102.3.419

Schacter, D. L., Guerin, S. A., & St. Jacques, P. L. (2011). Memory distortion: An adaptive perspective. *Trends in Cognitive Sciences*, *15*(10), 467–474. https://doi.org/10.1016/j.tics.2011.08.004

Stickgold, R., & Walker, M. P. (2013). Sleep-dependent memory triage: evolving generalization through selective processing. *Nature Neuroscience*, *16*(2), 139–145. https://doi.org/10.1038/nn.3303

Tamminen, J., Payne, J. D., Stickgold, R., Wamsley, E. J., & Gaskell, M. G. (2010). Sleep Spindle Activity is Associated with the Integration of New Memories and Existing Knowledge. *Journal of Neuroscience*, *30*(43), 14356–14360. https://doi.org/10.1523/JNEUROSCI.3028-10.2010

Tsujimura, H., Kotz S., Lander K., Tafuro A., Lewis P. A., (2018). Sleep-dependent false recognition of facial images is associated with slow wave sleep.

Tsujimura, H., Lewis, P. A., (2018). Investigating the role of sleep on the gist-based face recognition memory.

Walker, M. P., & Stickgold, R. (2004). Sleep-dependent learning and memory consolidation. *Neuron*. https://doi.org/10.1016/j.neuron.2004.08.031

Chapter 5:

Neural correlates of sleep role in curve-fitted face recognition memory

Hikaru Tsujimura, University of Manchester, UK

Sonja Kotz, University of Manchester, UK

Karen Lander, University of Manchester, UK

Penelope A. Lewis, University of Manchester, UK

5.1 Abstract

My prior work shows that sleep sharpens memory representations but also shifts memory representations away from areas of memory competition (with more than one learned item), leading to false recognition of items that were not previously learned and are only similar to one (not two) previously learned items. In a separate study, I showed that the overnight shift away from areas of competition is predicted by the proportion of time spent in slow wave sleep, which is heavily associated with memory replay and consolidation. In the current study, I set out to examine the neural correlates of these effects. To this end, I constructed a 2D face space by morphing faces along both age and gender, and then trained participants on just two 'old' faces which fell along the diagonal in this space. I then tested them for false recognition of all morphs along that diagonal, both those falling between the two old items (N+ morphs) and those falling to the outside of each old item (N- morphs). This testing was performed before and after sleep in order to detect the overnight shift in face representations. All testing was done in a functional magnetic resonance imaging scanner so I could assess brain responses. Although I was unable to replicate the prior behavioural results, I did show a numerical trend towards both curve sharpening and curve shifting away from N+ items across the night of sleep. In the neuroimaging data, the peak response of face selective areas mirrored the behavioural data by shifting away from N+ items and towards N- items, and the

distribution became sharper. By contrast, hippocampal responses shifted towards N+ items, and this shift was positively correlated with time spent in slow wave sleep ($p<0.05$). These findings suggest that the face selective areas may be coding for identity rather than the more superficial sensory characteristics of a face image. While the hippocampus may need to work harder to identify faces that fall within densely populated face-space, potentially due to the need for greater pattern separation.

5.2. Introduction

Sleep is important for consolidation of hippocampus-dependent memory (Marshall & Born, 2007), especially for processes of generalization (Friedrich, Wilhelm, Born, & Friederici, 2015), abstraction (Durrant, Cairney & Lewis, 2013), and gist extraction (Diekelmann, Born & Wagner, 2010: Lutz et al., 2017). It has been suggested that this process is as a result of restructuring its representation from original form (Hanert, Weber, Pedersen, Born, & Bartsch, 2017: Doxey, Hodges, Bodily, Muncy & Kirwan, 2018). Yet, the question of exactly how memories are processed during sleep remains to be fully understood.

In a prior study (Tsujimura & Lewis, 2018), we examined how facial memory representations change over sleep and found that, after a night of sleep, participants showed an increase in the tendency to mis-recognise face morphs to which they had not previously seen if these morphs were similar to just one pre-learned face, and a decrease in the tendency to mis-recognise morphs that were similar to more than one previously learned face. In a follow up study, we replicated this effect and found that these overnight performance changes were predicted by the time spent in slow wave sleep (SWS) (Tsujimura, Kotz, Lander, Tafuro & Lewis, 2018). Finally, we tested for the same effect using a more fine-grained analysis in which faces were morphed on a finer scale and we were able to model the memory representation of faces as a polynomial distribution, and show how this changed across sleep. This

showed a sharpening and a shift away from areas of memory competition over sleep, while wake was associated with flattening of the representation across perceptual space (Tsujimura, Kotz, Lander, Stamper, Garlick, & Lewis, 2018). However, none of our prior studies examined the neural correlates of these changes in behavioural performance. Thus, in the current study, aimed to determine how activity in areas underpinning face perception and memory is altered across sleep and how this relates to the behavioural shifts we have observed.

Rather than examining the whole brain, we chose *a priori* regions of interest in the hippocampus, face selective areas, and medial prefrontal cortex based on existing theoretical work. Thus, the hippocampus is essential for pattern separation, and may be important for dealing with competition between memories (Bakker, Kirwan, Miller, & Stark, 2008). We therefore predicted that hippocampal activation would be stronger in areas of greater competition after a night of sleep. Turning to face processing, faces are represented in a very special way (Kanwisher et al., 1997; Kanwisher and Yovel, 2006). There are three bilateral regions in the inferior occipital gyrus which respond more strongly to faces than to other stimuli, the fusiform face area (FFA), occipital face area (OFA) and the posterior superior temporal sulcus (pSTS). It has been proposed that these represent a core system for face processing which can be divided into two functions: one relates to detection of invariant features of face, such as identity and another relates to processing of variant features of face, such as face expressions (Haxby, Hoffman & Gobbini, 2000). Because our prior work shows that face representations are altered across sleep, we expect activation in these regions to be altered. Specifically, because our behavioural data shows an increase in false recognition for faces that are farther away from competition after

sleep, we expect face specific identity responses the brain to mirror this. Medial prefrontal cortex is associated with gist extraction and the representation of schematic information. Some authors (Takashima et al. 2009) have argued that as information becomes consolidated the MPFC takes on prior role of the hippocampus by acting as a hub that connects together the various cortical areas involved in that memory. Under this framework, we might expect an increase in MPFC activity after sleep.

In order to study these issues, we conducted an fMRI study using the same paradigm as in our prior work (Tsujimura, Kotz, Lander, Stamper, Garlick, & Lewis 2018), but with all test sessions conducted in the fMRI scanner. We specifically examined how activity in the hippocampus, MPFC, and empirically determined face selective regions changed across a night of sleep. Finally, given our prior observation that behavioural changes in this task are predicted by SWS, we tested for a relationship between time spent in SWS and overnight changes in the neuroimaging data.

5.3. Materials & Method (Experiment 6)

Participants

20 healthy Caucasian participants (13 females: age 24.77±5.07 (mean±SD), range 19-34 years old, 7 males: age 23.14±3.08 (mean±SD), range 18-27 years old) were recruited from posters, e-mail and online advertisement. One participant was excluded because of technical problems during sleep monitoring. Two participants

were excluded because of technical problems during brain image acquisition. In total, 17 participants (13 females: age 24.77±5.07 (mean±SD), range 19-34 years old, 4 males: age 23.50±3.87 (mean±SD), range 18-27 years old) were remained for data analysis. They were screened for a history of neurological and psychiatric diseases and sleep disorders. They were asked to abstain from caffeine and alcohol 24 hours prior to experiment. This study was approved by the University of Manchester Research Ethics Committee and the Cardiff University School of Psychology Ethics Committee. All participants were instructed about the experiment in detail and signed a consent form.

Facial stimuli in the two-dimensional face morph scale

In this study, a set of facial stimuli prepared in a same manner as the study of Tsujimura, Kotz, Lander, Stamper, Garlick, and Lewis (2018) was employed. Namely, the previous study morphed an actor's face using the features of age and gender in 25 stages (ie a starting point as 1st and then 3rd, 5th, 7th, …, 45th, 47th and 49th morphing point in the FaceGen software Modeller 3.5, http://www.facegen.com). One feature morphing scale was used as the vertical axis and the other feature morphing scale as the horizontal axis, which, as a result, created a 25x25 2D face morph scale. This process was repeated for the seven other actors' faces to create a total of eight 25x25 2D face morph scales (one per actor). See an example of a 25x25 2D face morph scale in **Figure 5.1A**. In this 25x25 face morph scale, a unit of face morphing (e.g. from 1st to 3rd FaceGen morphing point) was five times more subtle in facial manipulation than the minimal degree, which is a five-morphing-units (e.g. from 1st to 11th FaceGen morphing point), of facial manipulation participants can

detect a change of face identity perceptually in the previous studies (Tsujimura & Lewis, 2018; Tsujimura, Kotz, Lander, Tafuro & Lewis 2018). From each of the 25x25 2D face morph scale, 22 items were selected as face stimuli in this present experiment.

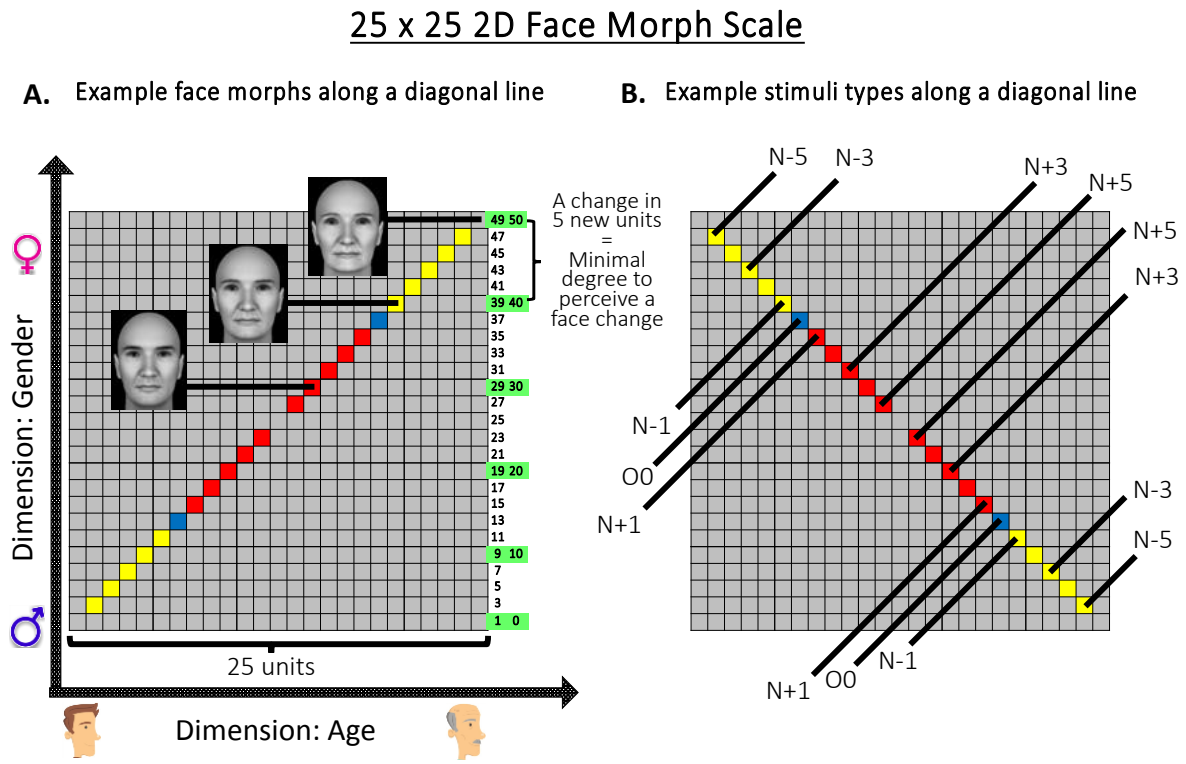

Figure 5.1. **An example of finer grained (25x25) 2D face morph scales.** (A) The left diagram showed a finer grained 25x25 2D face morph scale with example face morphs along a diagonal line. Green highlighted boxes indicated corresponding FaceGen morphing point for each face morph in 25x25 face morph scale. (B) The right diagram showed three major types (Yellow: N-, Dark Blue: O0, Red: N+) of face stimuli in the 2D morph scale with a number next to it. The number indicated how far stimuli were away from the O0 stimuli.

Of the eight invented finer grained 2D face morph scales, a random selection of four 2D face morph scales were assigned to each participant. Those selected 2D face morph scales of four actors or identities were termed as ID1 (identity 1), ID2, ID3 and ID4. Two randomly selected IDs (e.g. ID1 and ID4) from the four IDs were used in one of two experimental sessions and the other two IDs (e.g. ID2 and ID3) were

used in the other session (see the **design** section for detailed experimental procedures). In this study, a set of 22 face morphs along a right-top-left-bottom diagonal line were recruited from one ID (e.g. ID1) and the other set of 22 face morphs along a left-top-right-bottom diagonal line were recruited from the other ID (e.g. ID4) in the same session (e.g. the 22 red, blue and yellow dots along the two diagonal lines in **Figure 5.1**). This was critical to control confounds relating to the position of the stimuli in the 2D face morph scales.

Stimuli types

In each set of the 22 face morph selection, there were three major types of face stimuli: (1) the O0 types (the dark blue dots) that were actually trained at initial encoding task, (2) the N- types (the yellow dots) that were not trained at encoding and spatially neighbouring to only one of the two O0s, (3) the N+ types (the red dots) that were not trained at encoding and sandwiched spatially between the two O0s. In addition to this, a number next to the stimulus type indicated how far those stimuli were distanced from the O0 types. Please see **Figure 5.1B** for schematic illustration.

Design

In this experiment, all participants were assigned into the sleep group. They were asked to attend two sessions in the experiment. Each session was mainly composed of the encoding task outside of the fMRI scanner and the subsequent retrieval tasks inside of the scanner and a delay between the two memory tasks. In one session, the delay lasted 10 minutes (the short delay session) and in the other session the

delay lasted 12 hours across an overnight sleep (the long delay session). The order of the short and the long delay sessions in this experiment was pseudo-randomly assigned in each group. Thus, half participants (i.e. n=10) arrived at our sleep lab at 21:00 and completed the short delay session and the encoding task of the long delay session by 22:30 and then took the retrieval task of the long delay session after they spent an overnight of sleep at the sleep lab. One participant who started with the short delay session was excluded due to a technical problem with brain image acquisition, thus the total number of this condition was nine participants. The other half (n=10) of participants went through in the opposite way, completing the encoding task of the long delay session by 22:00 and then being instructed to sleep overnight in a bed at the sleep lab. Two participants who started with the long delay session were excluded due to technical problems of either sleep monitoring systems or brain image acquisition, thus a total number of this condition was eight participants. Their sleep was monitored by polysomnography (PSG). Participants were instructed to complete their sleep rituals, such as brushing teeth, changing into pyjamas, etc., prior to setting up the PSG system. After the PSG setting, all of the participants lay down in bed and the room light was turned off to sleep by 24:00. At the beginning of the first session or after the long delay, participants completed questionnaires of Stanford Sleepiness Scale (SSS; Hoddes et al., 1973) and Karolinska Sleepiness Scale (KSS; Glenville, Broughton, Wing, & Wilkinson, 1978) to check their vigilance. See **Figure 5.2A** for detailed instructions in this experiment.

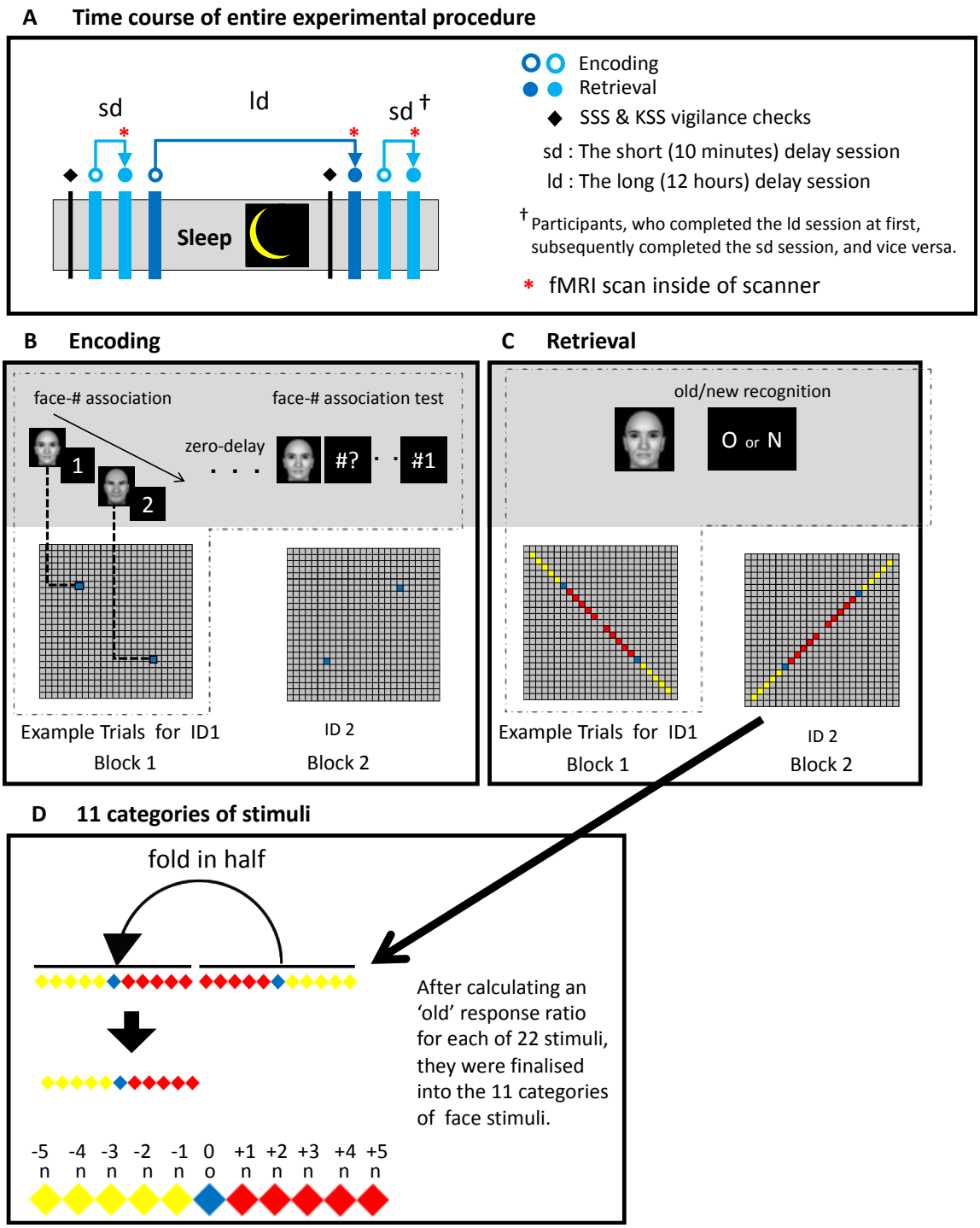

Figure 5.2. A Schematic Illustration of Experimental Procedures. Here drew an illustration of (A) a time course of the entire experiment procedure. The experimental procedure was mainly composed from (B) an encoding and (C) a retrieval tasks with a delay between them. During the encoding task (B), participants were trained to memorise and distinguish between two OLD stimuli (two blue boxes) of a same ID by taking a face-# association task and a zero-delayed test. During the retrieval task (C), participants took an old/new face recognition test while fMRI scan. Each task was run for two blocks; each block used an ID. See the **Design** and

the **Tasks** section for more details. After collecting participant's old/new responses for the 22 stimuli, they were summed into 11 types (categories) of stimuli, which were analysed (D).

Tasks

Procedures of the encoding and the retrieval tasks in the current experiment were same as that of Tsujimura, Kotz, Lander, Stamper, Garlick, and Lewis (2018), except that there was only old/new recognition task at retrieval and that participants took the encoding task on a PC monitor outside of fMRI scanner and then they took the retrieval task on the monitor inside of the fMRI scanner and that participants also took the face localiser task, in which they passively watched presentation of faces and scrambled faces, either before or after completing the old/new recognition task while participants stayed inside of the scanner. The face localiser task was a slight modification of the task initially developed by Eger, Schweinberger, Dolan, and Henson (2005). Previously, both encoding and retrieval tasks were taken on a same PC monitor in a same testing room. See **Figure 5.2B** and **5.2C** for schematic illustrations of those tasks and read the work of Tsujimura, Kotz, Lander, Stamper, Garlick, and Lewis (2018) for detailed procedures of this task.

Briefly, at encoding, participants were instructed to memorise two facial images taken from an invented 2D face morph space. Then at retrieval after a period of delay, participants were shown a collection of both 'old' (studied) and 'new' (unstudied) facial items, and asked to decide if they were either 'old' or 'new'. Each type of stimuli, in this case 22 items, was randomly shown for ten times. Retrieval was taken while brain activity was recorded inside of the fMRI scanner.

Polysomnographic (PSG) data

Polysomnographic (PSG) recordings were obtained and digitized at a sampling rate of 200Hz using the Embla N7000 system (Medcare-Embla®, Reykjavik, Iceland). The PSG electrodes were attached at F3, F4, Fz, C3, C4, Cz, O1, and O2, according to the international 10-20 system, as well as left and right mastoids (A1 and A2), forehead ground reference, left and right eyes (E1 and E2) and left and right sides and bottom of chin. Electrode impedance was verified to be lower than 5kΩ. Collected PSG data was scored by trained experimenters according to the AASM manual (American Academy of Sleep Medicine, Westchester, IL), using RemLogic 1.1 software, to measure distribution of sleep stages for each participant. A summary of the PSG data was shown in **Table 5.1**.

Brain imaging acquisition

In this study, a same protocol for brain imaging acquisition was used in an entire experiment. Specifically, fMRI images were acquired according to a standard echo-planar imaging (EPI) protocol in a 3 tesla Siemens scanner with repetition time (TR) = 2500 milliseconds, and echo time (TE) = 30 milliseconds. The main experiment of this present study employed an event-related design, where the images of face morphs were randomly presented for 1500 milliseconds, and inter trial intervals were jittered between 250 and 700, with a mean of 500 milliseconds. Data was acquired in 9 sessions of ~10 minutes each, with each session containing 48 trials. The face localiser task was run in a block design, where a half of participants were presented a collection of 16 faces for 1000 milliseconds one at a time, and then shown a

collection of 16 scrambled faces one at time while they passively saw those images. The other half of participants saw the two collection of images in an opposite order. This task lasted ~5 minutes. This face localiser task was run before or after the main old/new face recognition memory task. Please see resulting images responding to this face localizer task on **Figure 5.4**.

Data was analysed using a SPM12 (http://www.fil.ion.ucl.ac.uk). The retrieved functional EPI-BOLD (blood-oxygenation level-dependent) signals were realigned for motion correction. Then, images were co-registered into a standardized template of Montreal Neurological Institute (MNI) space, normalised and smoothed by spherical Gaussian kernel with a full width half-maximum of 8 mm. In an initial analyses, first-level design matrix used GLM to model the times of image presentations for the main old/new face recognition memory task and for the face localiser task, separately. Then at first, in the first level analysis, each participant's region of interests (ROI) was detected by brain activation responding to facial images contrasted to scrambled faces in the face localiser task. This was used as the first step of creating a functional mask of face localiser task. Then, brain activity responding to onsets of facial stimuli shown in the old/new face recognition memory task was contrasted with everything else to create a second mask in the main experiment. Then, conjunction of activated brain regions overlapped in both masks were defined as a mask of face responsive areas (Face). Functional masks of hippocampus (HPP), and medial prefrontal cortex (mPFC) were kindly given by colleagues. By employing these three masks as regions of interests (ROI), this present study examined how brain activity (i.e. BOLD signal) responds to the 22 types of stimuli described in **Figure 5.2** in the short and long delay session.

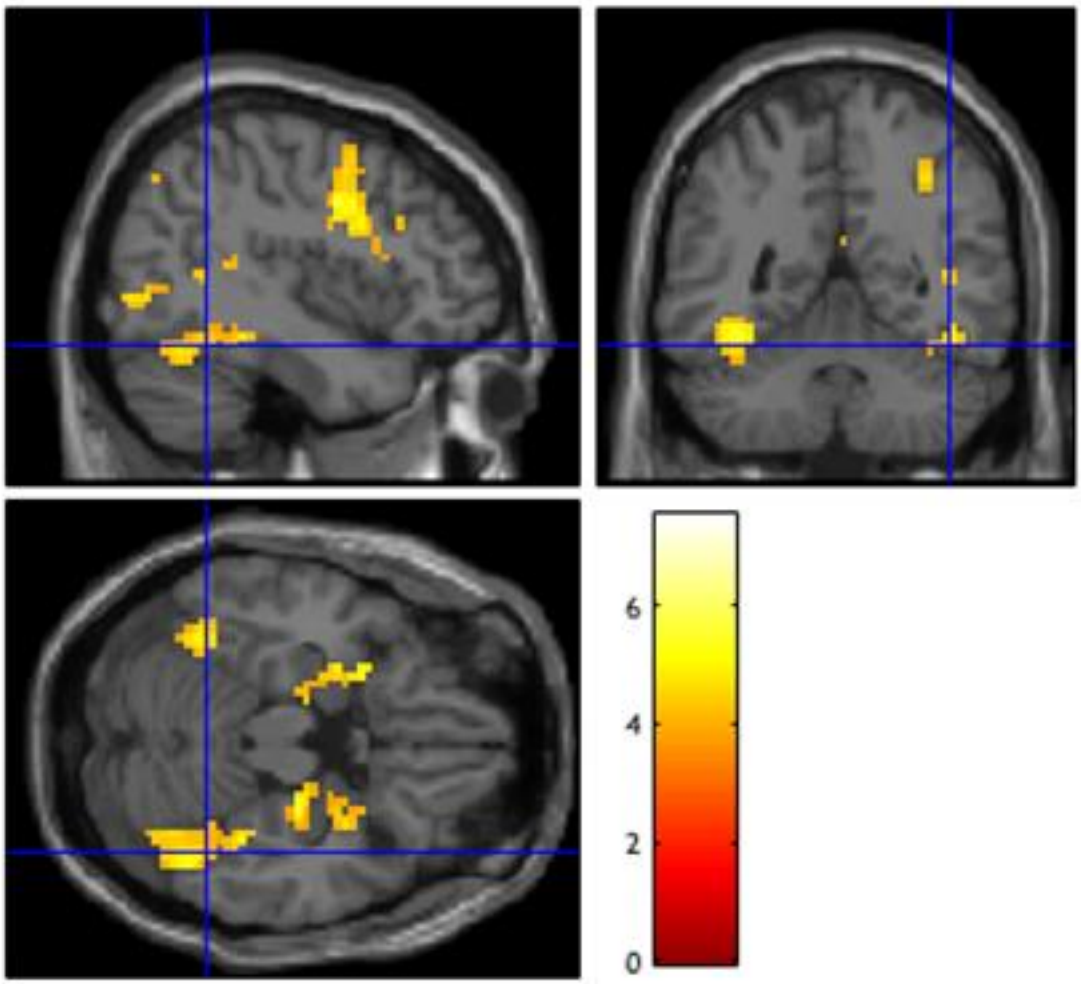

Figure 5.3. Images of functional responses in face localiser task. Face localiser task-related activation in the 2nd level analysis in the right face fusiform area (FFA) (42, -52, -19). The threshold is at uncorrected height threshold of $p < .001$, and an extent threshold of 0 voxels.

Statistical analysis

Behavioural measures

In this present study, ‘old’ responses for the 22 face stimuli were collected and an ‘old’ response ratio was calculated for each face stimulus (‘old’ responses vs total ‘old’/’new’ responses for that stimulus). Next, the old response ratios for the 22 face stimuli were averaged to collect the finalised old response ratios for the 11

categories of face stimuli (i.e. N-5, N-4, N-3, N-2, N-1, O0, N+1, N+2, N+3, N+4, N+5), which were used as parameters of face recognition memory (See a diagram of the 11 categories of stimuli in **Figure 5.2D**). Then, the collected old response ratios for the corresponding 11 categories (i.e. N-5, N-4, N-3, N-2, N-1, O0, N+1, N+2, N+3, N+4, N+5) in each session (i.e. sd and ld) were first fit to an inverted U-shaped second-order polynomial model (i.e. $y = ax^2+bx+c$, $a<0$). In the inverted U-shaped parabolic model, I assigned values of x from -0.5 to 0.5 in steps of 0.1, to quantify how far each stimulus category was from the old item (O0) and from negative values in x-axis (e.g. N-) and from positive values (e.g. N+), e.g. O0 was 0, N+1 was .1, N-1 was -.1, etc. Values of y were the old response ratios for each corresponding stimulus category, and the formula was used to calculate the three variables 'a', 'b', and 'c'. Namely, the value of a indicated a degree of sharpness of the inverted U curve. The larger the negative value of 'a' was the more its curve was sharper. To find the maximum, I calculated the derivative (i.e. $y' = 2ax+b$, $a<0$) of this inverted U-shaped second-order polynomial model for each session. Finally, I solved the formula in the case of $0=2ax+b$ or $x = -b/2a$ to obtain x value at peak of curve. See a schematic illustration of my proposed hypothesis in **Figure 5.3**.

Once curves of old response ratio in the short and long delay session had been fit, we calculated two types of curve parameters (i.e. curve sharpness, 'a' and x-coordinate at peak of curve) and tested whether these parameters changed across sleep compared to across a short delay (10 minutes). To test that, sample t-tests were used. Based on the previous results (Tsujimura, Kotz, Lander, Stamper, Garlick, & Lewis, 2018), I made an *a priori* hypothesis that both parameters would decrease after a sleep (i.e. ld-sd changes). Thus, one-tail was used.

Brain imaging data (Three ROIs: hippocampus, mPFC, and face responsive areas)

Same procedures were applied to BOLD signals responding to each of the 11 categories of stimuli fitting into the inverted u-shape curve, after BOLD signals were masking in the selected three ROIs (hippocampus (HPP), mPFC, and face responsive areas (Face)). That created two curves (the short delay (sd) and the long delay (ld) session) of BOLD signals for each ROI. In total, there were two (sd, ld) x three (HPP, mPFC, Face) curves for the brain imaging data. After calculating the ld-sd changes for the curve fitted brain imaging data, this time, one-way (ROIs: hippocampus, mPFC, face responsive areas) repeated ANOVA was used to assess if there was any differences between these three ROIs on each of the curve parameters. Again, based on the priori hypothesis proposed, the hippocampus curve was expected to shift toward right, indicated by positive values, while the other two were shifted left, indicated by negative values. Similarly, the hippocampus curve was expected to flatten, indicated by positive values of ‘a’, while the other two would sharpen, indicated by negative values. Therefore, again, one tail was used in this analysis.

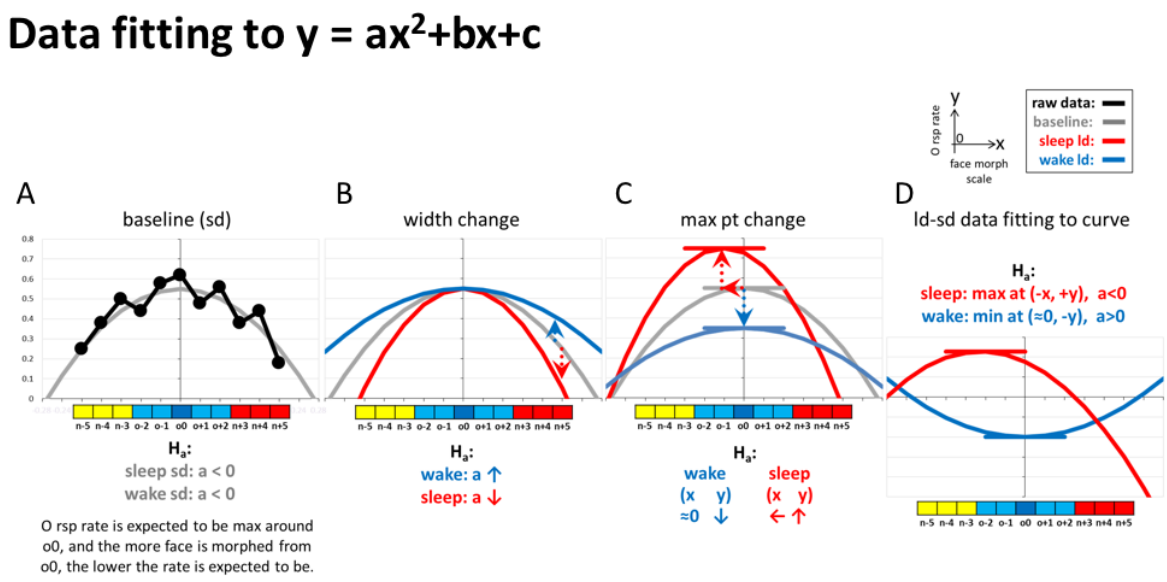

Figure 5.4. Data fitting into a parabolic curve of y = ax2+bx+c, (a<0). This diagram illustrated my hypothetic model of a sleep role in time-dependent modulation of participant's mental representation of face recognition memory. All drawn curves were estimated based on the results from previous studies (Tsujimura & Lewis, 2018; Tsujimura, Kotz, Lander, Tafuro & Lewis 2018; Tsujimura, Kotz, Lander, Stamper, Garlick & Lewis, 2018). X-axis is a scale of distance of each stimulus category morphed away from the O0 and y-axis is a scale of old response ratios for corresponding category.

5.4. Results (Experiment 6)

Sleep stage	Mean duration +/- SD
S1	2.82 +/- 3.60
S2	52.55 +/- 6.03
SWS	26.54 +/- 8.25
REM	18.09 +/- 6.11

Table 5.1. Mean duration (in percentage) and standard deviation of sleep stages. SWS is made up from a combination of N3 & N4.

5.4.1. Behavioural measures

5.4.1.1. Changes of curve fitted behavioural data across sleep

At first, this present study used one-tailed one-sample t-tests to assess if specific parameters (i.e. curve sharpness, “a”, and curve shifting, indicated as x coordinate change at peak of curve, see **Figure 5.1A** and **B**, respectively, for visual demonstration) of curve fitted behavioural data altered across an overnight sleep. This analysis revealed that neither curve sharpness (“a”), nor curve shifting (x coordinate change at peak of curve), significantly changed across sleep [curve sharpness, $t(16) = -.65$, $p = .53$, curve shifting, $t(16) = -.22$, $p = .83$], respectively (see **Figure 5.1A** and **B**). Yet, the results showed a numerical trend of curve sharpening and leftward curve shifting on the x(horizontal) axis across an overnight sleep, indicated by negative values of “a” (-.82+/- 5.39 (mean±SD)) and x coordinate change at peak of curve (-.02+/- .35 (mean±SD)) on the bar graphs of **Figure 5.1A** and **B**, respectively. Although the previous study with similar study paradigm (Tsujimura, Kotz, Lander, Stamper, Garlick, & Lewis, 2018) showed curve sharpening and leftward curve shifting across sleep, behavioural results of this present study failed to replicate the previous study, yet showed a trend that is matched with the previous study.

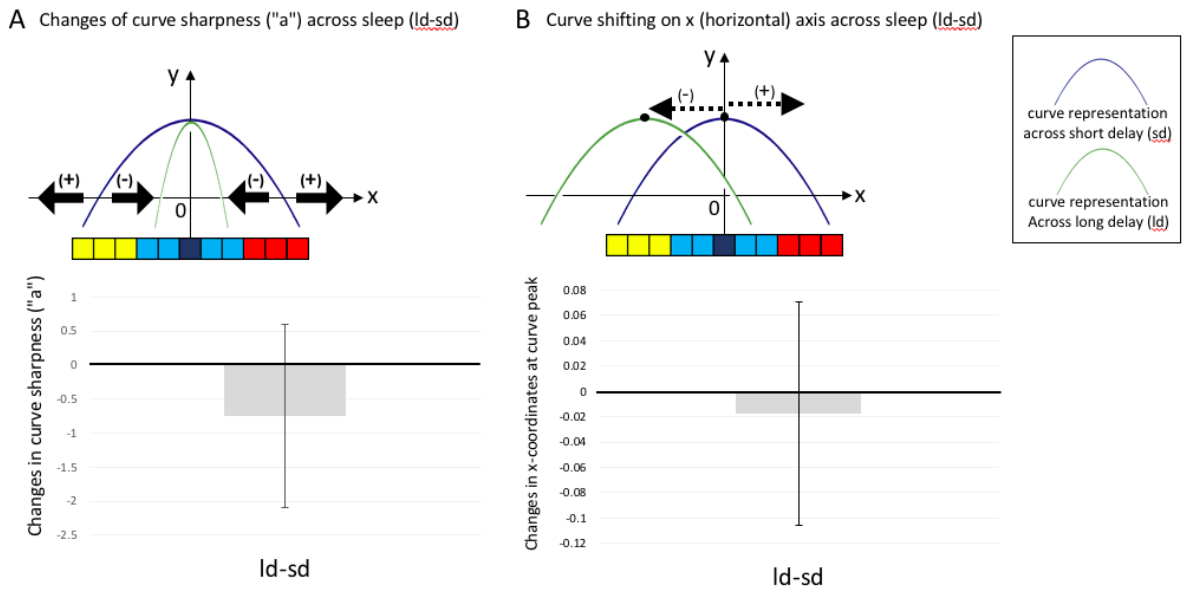

Figure 5.5. Parameter changes in curve fitted behavioural data across sleep. After behavioural data was fit to an inverted second-order polynomial curve (i.e. y = ax^2+bx+c, a<0), curve change across sleep (ld-sd) was calculated in the parameter of **A)** curve sharpness (“a”), and **B)** curve shifting on x (horizontal) axis at a peak of curve. The vertical axis of bar graph on **A** shows a change of curve sharpness (“a”) and curve became sharper when a bar went downward and flatter when a bar went upward across sleep. The vertical axis of bar graph on **B** shows a change of curve shifting on x (horizontal) axis at a peak of curve and curve shifted left when a bar went downward and curve shifted right when a bar went upward across sleep. Error bar represents a standard error.

5.4.2 Neuroimaging results

5.4.2.1. Changes of curve fitted brain imaging data across sleep

Similarly, in analysis of brain imaging data, BOLD signals of three selected regions (i.e. hippocampus (HPP), medial prefrontal cortex (mPFC), and Face selective area (Face)) of interests (ROIs) were fit into an inverted second-order polynomial curve individually, then two curve parameters (curve sharpening and curve shifting across sleep) of the curve fitted brain imaging data between the three ROIs were examined.

First, a repeated ANOVA with measure of curve sharpness (“a”) change across sleep revealed that this measure differed between the three ROIs, $F(2,30) = 3.71$, $p < .05$. Planned one-tailed repeated t-tests determined that a pattern of curve flattening, an opposite of curve sharpening indicated as a positive value of “a” on the bar graph of **Figure 5.2A**, of HPP (.52+/-4.75 (mean±SD)) significantly differed from a pattern of curve sharpening of Face (-2.19+/-5.07 (mean±SD)), $t(16) = 2.36$, $p < .05$. There was no significant difference between the other comparisons [HPP-mPFC comparison, $t(15) = 1.60$, $p = .13$; mPFC-Face comparison, $t(15) = -.39$, $p = .70$]. Planned one-tailed one-sample t-tests were also used and that revealed that such overnight curve sharpening of face selective area (Face) (-2.19+/-5.07 (mean±SD)) was significant, $t(16) = -1.79$, $p < .05$. However, the same analysis showed that overnight curve sharpening of mPFC (-.31+/-4.84 (mean±SD)) and curve flattening of HPP (.52+/-4.75 (mean±SD)) were not significant but each showed a numerical trend [mPFC, $t(15) = -.26$, $p = .40$; HPP, $t(16) = .45$, $p = .33$, respectively]. Here, all planned t-tests were not corrected for multiple comparisons for the purpose of exploratory analysis. See **Figure 5.2A** for summary.

Next, a repeated ANOVA with measure of curve shifting (changes of x-coordinate at peak of curve) across sleep revealed that this measure did not differ between the three ROIs, $F(2,30) = 2.20$, $p = .13$. Still, planned one-tailed repeated t-tests determined that a pattern of rightward curve shifting, indicated as a positive value of x coordinate change on the bar graph, see **Figure 5.2B**, for HPP (.22+/-.37 (mean±SD)) differed from a pattern of leftward curve shifting for Face (-.13+/-.36 (mean±SD)), $t(16) = 2.67$, $p < .05$. There was no difference between the other comparisons [HPP-mPFC comparison, $t(15) = 1.60$, $p = .07$; mPFC-Face

comparison, $t(15) = -.39$, $p = .35$]. Planned one-tailed one-sample t-tests were also used to assess if such curve shifting across sleep was significant. This analysis revealed that only rightward curve shifting of HPP (.22+/-.37 (mean±SD)) showed significance, $t(16) = 2.43$, $p < .05$, but leftward curve shifting of mPFC (-.22+/-.79 (mean±SD)) and Face (-.13+/-.36 (mean±SD)) showed only a numerical trend, [mPFC, $t(15) = -1.14$, $p = .14$; Face, $t(16) = -1.53$, $p = .07$, respectively. All planned t-tests were not corrected for multiple comparisons for the purpose of exploratory analysis. See **Figure 5.2B** for summary.

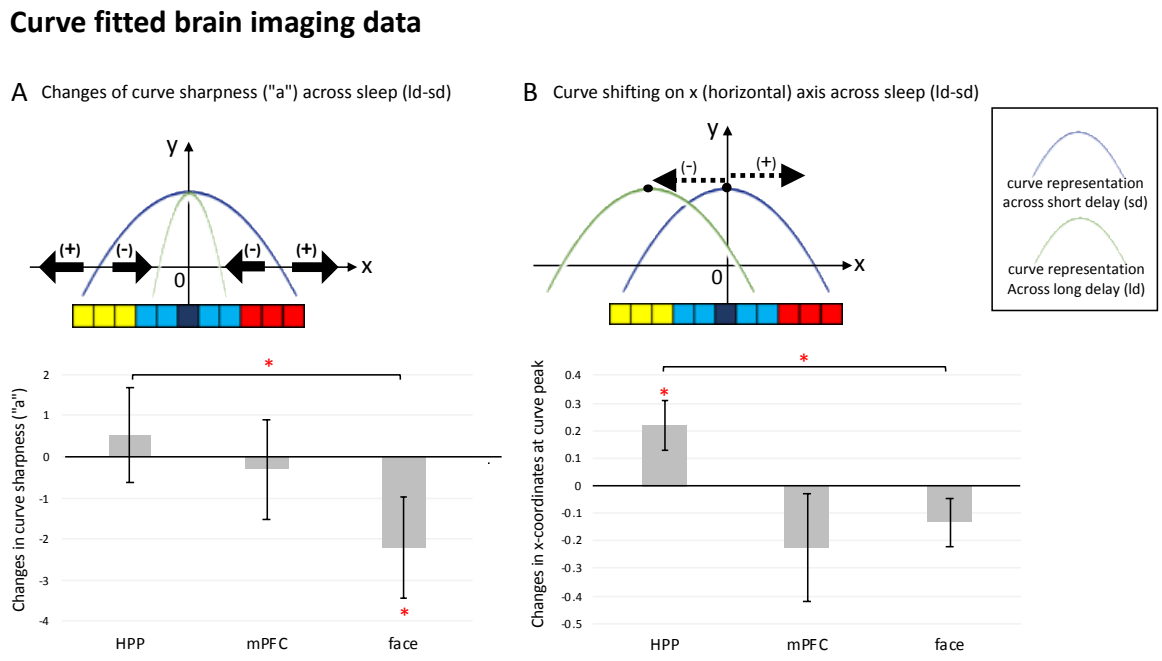

Figure 5.6. Parameter changes in curve fitted brain imaging data across sleep. After brain imaging data was fit to an inverted second-order polynomial curve (i.e. y = ax^2+bx+c, a<0), curve change across sleep (ld-sd) was calculated in the parameter of **A)** curve sharpness ("a"), and **B)** curve shifting on x (horizontal) axis at a peak of curve. The vertical axis of bar graph on **A** shows a change of curve sharpness ("a") and curve became sharper when a bar went downward and flatter when a bar went upward. The vertical axis on **B** shows a change of curve shifting on x (horizontal) axis at a peak of curve and curve shifted left when a bar went downward and curve shifted right when a bar went upward. Error bar is standard error.

In summary, these results including both behavioural and brain imaging data showed that an inverted u-shape curve of hippocampal (HPP) activity became flatter and shifted toward right on x-axis, which contrasted against a curve shifting of brain activity in Face selective area (Face) across sleep. That means that, a curve of HPP activity shifted toward the areas where memory competition existed between trained items, while a curve of brain activity in face selective area (Face) shifted away from the competition. Also, there was a mere numerical trend, but a curve of brain activity in medial prefrontal cortex (mPFC) and a curve of behavioural data followed the same pattern as a curve of brain activity in face selective area (Face). See **Figure 5.3** for a schematic summary illustrating overnight sleep changes of those curves.

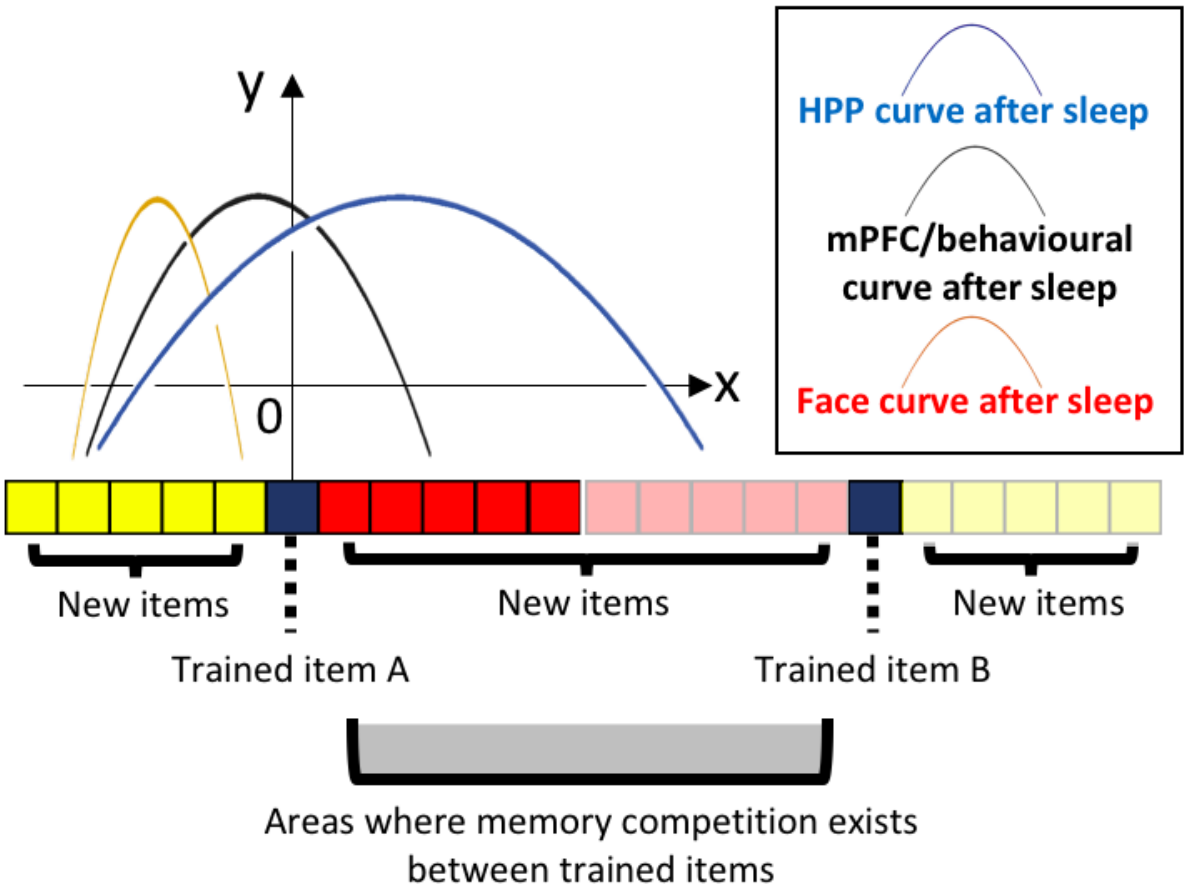

Figure 5.7. Schematic summary of over-night curve representations. This summary demonstrates over-night curve representations of brain activity in hippocampus (HPP), medial prefrontal cortex (mPFC) and Face

selective area (Face), as well as curve of behavioural performance in face recognition memory task. The horizontal axis represents the face morphing diagonal line described in **Figure 5.2D**. For curves of brain activity in those three ROIs, vertical axis represents amounts of BOLD signals responding to corresponding facial items on the horizontal axis. For a curve of behavioural performance, vertical axis represents a proportion of participant's old responses on corresponding facial items on the horizontal axis. Note that this is not actual data, but a schematic summary illustrating how each curve alters across an overnight sleep.

5.4.3. Correlation link between amounts of sleep stages (i.e. SWS) and curve fitted data of behavioural and/or brain imaging measure

Based on the prior results of Tsujimura, Kotz, Lander, Stamper, Garlick and Lewis (2018) showing that a time spent in slow wave sleep (SWS) predicted the extent to which participants falsely endorsed face images in non-competitive areas of the face-morph space when compared to face images in densely populated areas of the same face space, this present study formed an *a priori* hypothesis that overnight shifts in the curve fitting parameters, curve sharpness ('a') and curve shifting ('x coordinate change'), would also be predicted by time spent in SWS. To test this, a one-tailed Pearson correlation was used to assess if SWS was negatively correlated with two curve parameters of curve fitted behavioural data and curve fitted brain activity in medial prefrontal cortex (mPFC) and face responsive area (Face) and if there was an opposite pattern of correlation, which is a positive correlation, between SWS and hippocampal (HPP) activity. This analysis revealed that SWS was positively correlated with curve shifting of HPP activity [$r(15) = .62$, $p < .01$]. There was no other correlation. Yet, when removing an obvious outlier (> 3SD) of mPFC data, this analysis revealed that SWS was negatively correlated with curve shifting of brain activity in mPFC [$r(13) = -.46$, $p < .05$]. There was no correlation between SWS and curve shifting of curve fitted behavioural data, [$r(15) = .14$, $p > .05$] and between

SWS and curve fitted brain activity in face responsive areas (Face), [$r(16) = -.13$, $p > .05$]. There was no correlation between SWS and curve sharpness ('a') of any data, $p > .05$. This correlation test was not corrected for multiple comparisons for the purpose of exploratory analysis.

These results showed that, amounts of time spent in SWS predicted rightward curve shifting of brain activity in HPP matching with our *a priori* hypothesis, but did not show leftward curve shifting of brain activity in face selective area (Face). Yet, there was a trend that is congruent with our *a priori* hypothesis that leftward shifting of brain activity in face selective area (Face) was correlated with SWS. See **Figure 5.8** for summary**.**

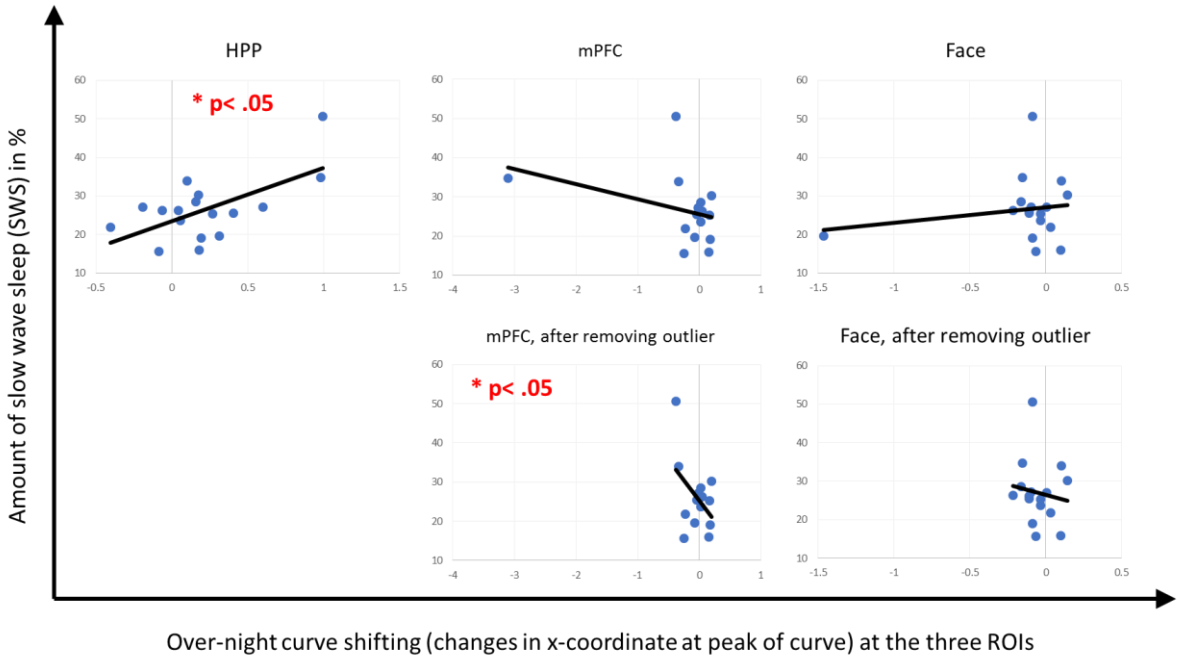

Figure 5.8. Correlation between SWS and over-night curve representation of brain activity. This illustrates overnight curve shifting of brain activity in the three ROIs (hippocampus (HPP), medial prefrontal cortex ((mPFC), and face selective area (Face)). Behavioural data was not included in this graph as there was no interesting finding in this present study. The upper three graphs showed full data and the lower two graphs showed data

after removing an obvious outlier (>3 SD). On the horizontal axis, negative values indicated that curve of brain activity shifted left across sleep, and vice versa. Black lines showed the trend lines of each graph.

5.5. Discussion (Experiment 6)

We have previously shown that face representations shift towards areas of non-competitive space across a night of sleep. Specifically, we showed that a curve fitted to the behavioural data such that it provided a representation of recognition memory performance became sharper overnight and shifted away from an area of perceptual space in which two older memories existed Tsujimura, Kotz, Lander, Stamper, Garlick, and Lewis (2018). While we have examined this type of phenomena in several other studies (Tsujimura & Lewis, 2018; Tsujimura, Kotz, Lander, Tafuro & Lewis, 2018) we have not previously examined the neural correlates of this shift. In the current study we built our own prior work, by repeating essentially the same behavioural design but now testing participants in the fMRI scanner so we could examine the neural correlates. While, we did not fully replicate the original behavioural findings, this present study did show a numerical trend towards the same pattern. This inconsistency between results of the present and the previous studies may have occurred because participants took the old/new face recognition memory task inside the fMRI scanner in this present study. Previously, both encoding and later recognition tasks were run on a same monitor in a same testing room. In contrast, in the case of this present study, initial training was run on a PC monitor at a testing room, but a later recognition task was run on a monitor inside of the scanner. This difference of monitors for initial encoding and later face recognition memory test might make difficult for participants to evaluate whether facial images

on the monitor inside of scanner were the ones studied at initial encoding. Therefore, a future study requires to improve this setting.

Turning to the neuroimaging data, we had two main findings in the present study. First, we showed that a polynomial curve, analogous to that fitted to the behavioural data in order to describe the memory distribution, is fitted to brain activity in face responsive areas shifted away from the area of memory competition across a night of sleep, and also became sharper (**Figure 5.7**). This finding matched our *a priori* hypothesis that a curve fitted cortical activity that represents acquired face memories would follow the same pattern as the curve fitted behavioural data shown in the previous study (Tsujimura, Kotz, Lander, Stamper, Garlick, & Lewis, 2018). In keeping with the idea that half of the core system is associated with detecting invariant aspects of a face (e.g. identity) (Haxby, Hoffman & Gobbini, 2000), these findings suggest that activity of the face responsive area is somewhat top-down, knowledge-based function to face perceptions, which become responsive only after acquisition of knowledge of trained face's identity.

Second, we showed that a parallel polynomial curve fitted to brain activity in the hippocampus shifted *towards* the area of memory competition across a night of sleep. This finding supported our *a priori* hypothesis that hippocampal activity was more responsive to the area where memory competes, and is in keeping with a role for hippocampus plays in pattern separation (Bakker, Kirwan, Miller, & Stark, 2008), which may be very important in this task in order to avoid confusion between the acquired memory representations.

Interestingly, the overnight rightward shifting in the hippocampal curve was predicted by with time spent in SWS. This is in keeping with our prior observation that SWS predicted the extent to which behavioural responses shifted across sleep (Tsujimura, Kotz, Lander, Stamper, Garlick, & Lewis, 2018). Our previous study also showed a reduction of false recognition on new items where memory competition existed and increases of false recognition outside of the competition. Thus, we propose that there is a demand to increase hippocampal activity after sleep in order to reject new items located in the area where trained face memories compete.

In sum, while this study did not fully replicate prior behavioural results it did show a trend towards the same shift in behaviour, with the memory representation curve shifting away from areas of completion. On the neural level, face specific responses shifted in parallel to this behavioural shift, suggesting that representations of *identity* itself may shift in this manner. Hippocampal responses, on the other hand, shifted towards areas of competition, and this shift was predicted by SWS, which could signify an increase in hippocampal engagement where pattern separation is needed.

References

Bakker, A., Kirwan, B. C., Miller, M., Stark, C. E. L., (2008). Pattern Separation in the Human Hippocampal CA3 and Dentate Gyrus. Science, 319(5870), 1640-42.

Diekelmann, S., Born, J., & Wagner, U. (2010). Sleep enhances false memories depending on general memory performance. *Behavioural Brain Research*, *208*(2), 425–429. https://doi.org/10.1016/j.bbr.2009.12.021

Doxey, C. R., Hodges, C. B., Bodily, T. A., Muncy, N. M., Kirwan, C. B. (2018). The effects of sleep on the neural correlates of pattern separation. Hippocampus, 28(2), 108-120.

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. https://doi.org/10.1093/cercor/bhs244

Eger, E., Schweinberger, S.R., Dolan, R.J., and Henson, R.N. (2005). Familiarity enhances invariance of face representations in human ventral visual cortex: fMRI evidence, Neuroimage, 26, 1128-39.

Friedrich, M., Wilhelm, I., Born, J., and Friederici, A. D. (2015). Generalization of word meanings during infant sleep, Nature communications, 6, 6004.

Glenville, M., Broughton, R., Wing, a M., & Wilkinson, R. T. (1978). Effects of sleep deprivation on short duration performance measures compared to the Wilkinson auditory vigilance task. *Sleep*, *1*(2), 169–76. https://doi.org/10.1093/sleep/1.2.169

Hanert, A., Weber, F. D., Pedersen, A., Born, J. and Bartsch, T. (2017). Sleep in humans stabilizes pattern separation performance. Journal of Neuroscience, 1189-17.

Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences*. https://doi.org/10.1016/S1364-6613(00)01482-0

Hoddes, E., Zarcone, V., Smythe, H., Phillips, R., & Dement, W. C. (1973). Quantification of Sleepiness: A New Approach. *Psychophysiology*, *10*(4), 431–436. https://doi.org/10.1111/j.1469-8986.1973.tb00801.x

Kanwisher, N., McDermott, J., and Chun, M. M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. Journal of Neuroscience, 17, 4302-11.

Kanwisher, N. and Yovel , G. (2006). The fusiform face area: a cortical region specialized for the perception of faces. Philosophical Transactions of the Royal Society B, 361, 2109-28.

Lutz, N. D., Diekelmann, S., Hinse-Stern, P., Born, J., & Rauss, K. (2017). Sleep Supports the Slow Abstraction of Gist from Visual Perceptual Memories. *Scientific Reports*, *7*, 42950. https://doi.org/10.1038/srep42950

Marshall, L., & Born, J. (2007). The contribution of sleep to hippocampus-dependent memory consolidation. *Trends in Cognitive Sciences*. https://doi.org/10.1016/j.tics.2007.09.001

Takashima, A., Nieuwenhuis, I. L. C., Jensen, O., Talamini, L. M., Rijpkema, M., & Fernandez, G. (2009). Shift from Hippocampal to Neocortical Centered Retrieval Network with Consolidation. *Journal of Neuroscience*, *29*(32), 10087–10093. https://doi.org/10.1523/JNEUROSCI.0799-09.2009

Chapter 6:

General Discussion

6.1. Synthesis

My first three experimental aims in this thesis were as follows:

(1) Determine how the density of learned stimuli represented in this 2D-space influences the shape of the resulting knowledge space within the 2D map.
(2) Determine how sleep alters the shape and size of such knowledge space.
(3) Determine which neural activity and sleep stage relates to constructed knowledge space and updating.

The studies described in Chapters 2-5 have addressed all of these goals. My findings are summarised as follows:

Firstly, people alter their (face) recognition memory differently across sleep and wake. Across sleep, overall memory accuracy is maintained, whereas competition free false memory is concurrently increased relative to false memory in areas where several veridical memories compete (Chapters 2-4). Further investigation shows that this over-time shift towards false memories in areas of little memory competition is specifically associated with the amount of time spent in SWS (Chapters 3-4). In contrast, across an equivalent amount of wakefulness, veridical recognition memory merely deteriorates, and false memory does not change (Chapters 2-4). Interestingly, these results do not appear to be specific to upright face images, but instead generalise at least to inverted faces (Chapters 3).

Secondly, my work demonstrates that an overall internal representation of recognition memory can be described using an inverted U-curve. By employing this simplified representation, I found evidence suggesting that the sleep-dependent recognition change may be due to a shift of this U-curve away from existing memory competitors, as well as a sharpening of its shape. In addition, memory deterioration across wakefulness is expressed as a flattening of the curve (Chapter 4). Furthermore, by examining brain activity during performance of this task with fMRI, I showed that this shift in internal representation is associated with a parallel shift in neural activity in the face selective areas, and a shift in the opposite direction in the hippocampus. Specifically, face responsive areas decrease activity in areas of competition and increase activity in areas free from competition, while hippocampus increases activity in areas of competition and reduce activity in areas where there is no competition (Chapter 5).

The shifting of the memory curve away from existing competitors suggests that the mental representation of each studied item may become perceptually more distinct and differentiated after sleep. Curve sharpening suggests that the recognition pattern becomes more precise, even though it is less accurate (as the represented curve shifted away from the original form). Sharpening of the response curve in face selective areas may therefore suggest that these regions respond to the perception of a learned identity, and are thus top-down mediated, rather than simply responding to perceptual input. In contrast, flattening of the brain activity curve in hippocampus suggests that the function of hippocampus becomes specialised into pattern separation of studied items to avoid confusion between internal representations.

Overall, it appears that, across a period of sleep, memory representations shift such that participants believe items in a new, sharply defined area of perceptual space, are actually 'old'. This 'false memory' area of perceptual space occurs just outside the area where several veridical memories are crowded together. Thus, in Chapter 4, for instance, it occurs just to the left (n-) side of where the old memory actually was, compared to the right (n+) side (see **figure 5.3**). Accordingly, I propose that sleep plays a critical role in creating a distinction between existing memory representations when they compete, in a manner that resembles pattern separation. This process promotes increases of certain false memory (i.e. n-) but discourages the other false memory (i.e. n+). Additionally, based on the finding in the neuroimaging study (Chapter 5), I suggest that the observed sleep-dependent changes in cognitive/perceptual processes is linked to engagement of two distinct brain regions (i.e. hippocampus and face responsive areas), where hippocampus plays a critical role in processes of pattern separation and face responsive areas play a role in top-down processing of promoting false memories.

6.2. Summing up

Although sleep has been shown to be important for integration (e.g. Durrant, Taylor, Cairney & Lewis, 2011; Lewis & Durrant, 2011; Tamminen, Payne, Stickgold, Wamsley, & Gaskell, 2010), recent work has suggested that sleep is also important for forgetting, and can help us to avoid confusion between competing (or similar) information (Wilhelm et al. 2011; Feld & Born, 2017; Poe, 2017). The work presented in this thesis provides key support for this idea by clearly demonstrating that sleep

can facilitate a shift away from the correct memory representation and towards a representation which, though false, facilitates discrimination between the different items (in this case faces) which were originally encoded.

References

Andrade, K. C., Spoormaker, V. I., Dresler, M., Wehrle, R., Holsboer, F., Sämann, P. G., & Czisch, M. (2011). Sleep spindles and hippocampal functional connectivity in human NREM sleep. *The Journal of Neuroscience*, *31*(28), 10331–10339. https://doi.org/10.1523/JNEUROSCI.5660-10.2011

Aserinsky E. & Kleitman N (1953). Regularly Occurring Periods of Eye Motility, and Concomitant Phenomena, during Sleep Science, 118(3062) 273-274.

Baddeley A. D. and Warrington E. K. (1970). Amnesia and the distinction between long- and short-term memory. Journal of Verbal Learning and Verbal Behavior 9(2), 176-189.

Bakker, A., Kirwan, B. C., Miller, M., Stark, C. E. L., (2008). Pattern Separation in the Human Hippocampal CA3 and Dentate Gyrus. Science, 319(5870), 1640-42.

Bartlett, F. C. (1932). Remembering : A Study in Experimental and Social Psychology. *Cambridge, Social Psychology*, 1–11. https://doi.org/10.1111/j.2044-8279.1933.tb02913.x

Bergmann, T. O., Mölle, M., Diedrichs, J., Born, J., & Siebner, H. R. (2012). Sleep spindle-related reactivation of category-specific cortical regions after learning

face-scene associations. *NeuroImage*, *59*(3), 2733–2742. https://doi.org/10.1016/j.neuroimage.2011.10.036

Bliss, T. V. P., & Cooke, S. F. (2011). Long-term potentiation and long-term depression: a clinical perspective. *Clinics (São Paulo, Brazil)*, *66 Suppl 1*(Suppl 1), 3–17. https://doi.org/10.1590/S1807-59322011001300002

Born, J., & Wilhelm, I. (2012). System consolidation of memory during sleep. *Psychological Research*. https://doi.org/10.1007/s00426-011-0335-6

Bruce, V., & Young, A. (1986). Understanding face recognition. *British Journal of Psychology*, *77*(3), 305–327. https://doi.org/10.1111/j.2044-8295.1986.tb02199.x

Buhl, D. L., & Buzsáki, G. (2005). Developmental emergence of hippocampal fast-field “ripple” oscillations in the behaving rat pups. *Neuroscience*, *134*(4), 1423–1430. https://doi.org/10.1016/j.neuroscience.2005.05.030

Bushey, D., Tononi, G., & Cirelli, C. (2011). Sleep and Synaptic Homeostasis: Structural Evidence in Drosophila. *Science*, *332*(6037), 1576–1581. https://doi.org/10.1126/science.1202839

Buzsáki, G. (1989). Two-stage model of memory trace formation: A role for “noisy” brain states. *Neuroscience*, *31*(3), 551–570. https://doi.org/10.1016/0306-4522(89)90423-5

Buzsáki, G. (1996). The hippocampo-neocortical dialogue. *Cerebral Cortex*. https://doi.org/10.1093/cercor/6.2.81

Buzsaki, G., Horvath, Z., Urioste, R., Hetke, J., & Wise, K. (1992). High-frequency network oscillation in the hippocampus. *Science*, *256*(5059), 1025–1027. https://doi.org/10.1126/science.1589772

Cabeza, R., Rao, S. M., Wagner, A. D., Mayer, A. R., & Schacter, D. L. (2001). Can medial temporal lobe regions distinguish true from false? An event-related functional MRI study of veridical and illusory recognition memory. *Proceedings of the National Academy of Sciences*, *98*(8), 4805–4810. https://doi.org/10.1073/pnas.081082698

Cain, K., Oakhill, J., & Lemmon, K. (2004). Individual Differences in the Inference of Word Meanings From Context: The Influence of Reading Comprehension, Vocabulary Knowledge, and Memory Capacity. *Journal of Educational Psychology*, *96*(4), 671–681. https://doi.org/10.1037/0022-0663.96.4.671

Ciaramelli, E., Ghetti, S., Frattarelli, M., & Làdavas, E. (2006). When true memory availability promotes false memory: Evidence from confabulating patients. *Neuropsychologia*, *44*(10), 1866–1877. https://doi.org/10.1016/j.neuropsychologia.2006.02.008

Clemens, Z., Mölle, M., Eross, L., Barsi, P., Halász, P., & Born, J. (2007). Temporal coupling of parahippocampal ripples, sleep spindles and slow oscillations in humans. *Brain*, *130*(11), 2868–2878. https://doi.org/10.1093/brain/awm146

Contreras, D., & Steriade, M. (1995). Cellular basis of EEG slow rhythms: a study of dynamic corticothalamic relationships. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *15*(1 Pt 2), 604–622.

Contreras, D., Timofeev, I., & Steriade, M. (1996). Mechanisms of long-lasting hyperpolarizations underlying slow sleep oscillations in cat corticothalamic networks. *The Journal of Physiology*, *494*(Pt 1), 251–264. https://doi.org/10.1113/jphysiol.1996.sp021488

Cousins, J. N., El-Deredy, W., Parkes, L. M., Hennies, N., & Lewis, P. A. (2014). Cued Memory Reactivation during Slow-Wave Sleep Promotes Explicit Knowledge of a Motor Sequence. *Journal of Neuroscience*, *34*(48), 15870–15876. https://doi.org/10.1523/JNEUROSCI.1011-14.2014

Darsaud, A., Dehon, H., Lahl, O., Sterpenich, V., Boly, M., Dang-Vu, T., … Collette, F. (2011). Does Sleep Promote False Memories? *Journal of Cognitive Neuroscience*, *23*(1), 26–40. https://doi.org/10.1162/jocn.2010.21448

De Gennaro, L., & Ferrara, M. (2003). Sleep spindles: An overview. *Sleep Medicine Reviews*. https://doi.org/10.1053/smrv.2002.0252

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

Destexhe, a, & Sejnowski, T. J. (2003). Interactions between membrane conductances underlying thalamocortical slow-wave oscillations. *Physiol Rev*, *83*(4), 1401–1453. https://doi.org/10.1152/physrev.00012.2003

Diekelmann, S., & Born, J. (2010). The memory function of sleep. *Nat Rev Neurosci*, *11*(2), 114–126. https://doi.org/nrn2762 [pii]\r10.1038/nrn2762

Diekelmann, S., Born, J., & Rasch, B. (2016). Increasing Explicit Sequence Knowledge by Odor Cueing during Sleep in Men but not Women. *Frontiers in Behavioral Neuroscience*, *10*(April), 74. https://doi.org/10.3389/fnbeh.2016.00074

Diekelmann, S., Born, J., & Wagner, U. (2010). Sleep enhances false memories depending on general memory performance. *Behavioural Brain Research*, *208*(2), 425–429. https://doi.org/10.1016/j.bbr.2009.12.021

Djonlagic, I., Rosenfeld A., Shohamy, D., Myers, C., Gluck, M., & Stickgold, R. (2009). Sleep enhances category learning. *Learning & Memory*, 16, 751-755.

Doxey, C. R., Hodges, C. B., Bodily, T. A., Muncy, N. M., Kirwan, C. B. (2018). The effects of sleep on the neural correlates of pattern separation. *Hippocampus*, 28(2), 108-120.

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. https://doi.org/10.1093/cercor/bhs244

Durrant, S. J., Cairney, S. A., & Lewis, P. A. (2016). Cross-modal transfer of statistical information benefits from sleep. *Cortex*, *78*, 85–99. https://doi.org/10.1016/j.cortex.2016.02.011

Durrant, S. J., & Lewis, P. A. (2009). Memory Consolidation: Tracking Transfer with Functional Connectivity. *Current Biology*. https://doi.org/10.1016/j.cub.2009.08.019

Durrant, S. J., Taylor, C., Cairney, S., & Lewis, P. A. (2011). Sleep-dependent consolidation of statistical learning. *Neuropsychologia*, *49*(5), 1322–1331. https://doi.org/10.1016/j.neuropsychologia.2011.02.015

Eger, E., Schweinberger, S.R., Dolan, R.J., and Henson, R.N. (2005). Familiarity enhances invariance of face representations in human ventral visual cortex: fMRI evidence, Neuroimage, 26, 1128-39.

Ego-Stengel, V., & Wilson, M. A. (2010). Disruption of ripple-associated hippocampal activity during rest impairs spatial learning in the rat. *Hippocampus*, *20*(1), 1–10. https://doi.org/10.1002/hipo.20707

Ellenbogen, J. M., Hu, P. T., Payne, J. D., Titone, D., & Walker, M. P. (2007). Human relational memory requires time and sleep. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(18), 7723–8. https://doi.org/10.1073/pnas.0700094104

Ellenbogen, J. M., Hulbert, J. C., Jiang, Y., & Stickgold, R. (2009). The sleeping brain's influence on verbal memory: Boosting resistance to interference. *PLoS ONE*, *4*(1). https://doi.org/10.1371/journal.pone.0004117

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*. https://doi.org/10.1016/j.conb.2006.10.006

Eschenko, O., Ramadan, W., Molle, M., Born, J., & Sara, S. J. (2008). Sustained increase in hippocampal sharp-wave ripple activity during slow-wave sleep after

learning. *Learning & Memory*, *15*(4), 222–228. https://doi.org/10.1101/lm.726008

Euston, D. R., Tatsuno, M., & McNaughton, B. L. (2007). Fast-Forward Playback of Recent Memory Sequences in Prefrontal Cortex During Sleep. *Science*, *318*(5853), 1147–1150. https://doi.org/10.1126/science.1148979

Feld, G. B., & Born, J. (2017). Sculpting memory during sleep: concurrent consolidation and forgetting. *Current Opinion in Neurobiology*, *44*, 20–27. https://doi.org/10.1016/j.conb.2017.02.012

Fenn, K. M., Gallo, D. a, Margoliash, D., Roediger, H. L., & Nusbaum, H. C. (2009). Reduced false memory after sleep. *Learning & Memory (Cold Spring Harbor, N.Y.)*, *16*(Gallo 2006), 509–513. https://doi.org/10.1101/lm.1500808

Frankland, P., & Bontempi, B. (2005). The organization of recent and remote memories. *Nature Reviews Neuroscience*, *6*, 119–130. https://doi.org/10.1038/nrn1607

Friedrich, M., Wilhelm, I., Born, J., and Friederici, A. D. (2015). Generalization of word meanings during infant sleep, Nature communications, 6, 6004.

Gais, S. & Born, J. (2004). Declarative memory consolidation: Mechanisms acting during human sleep. *Learning & Memory*, *11*(6), 679–685. https://doi.org/10.1101/lm.80504

Gais, S., Hüllemann, P., Hallschmid, M., & Born, J. (2006). Sleep-dependent surges in growth hormone do not contribute to sleep-dependent memory consolidation.

Psychoneuroendocrinology, *31*(6), 786–791. https://doi.org/10.1016/j.psyneuen.2006.02.009

Genzel, L., Rossato, J. I., Jacobse, J., Grieves, R. M., Spooner, P. A., Battaglia, F. P., … Morris, R. G. M. (2017). The Yin and Yang of Memory Consolidation: Hippocampal and Neocortical. *PLoS Biology*, *15*(1). https://doi.org/10.1371/journal.pbio.2000531

Genzel, L., Spoormaker, V. I., Konrad, B. N., & Dresler, M. (2015). The role of rapid eye movement sleep for amygdala-related memory processing. *Neurobiology of Learning and Memory*. https://doi.org/10.1016/j.nlm.2015.01.008

Ghosh, V. E., & Gilboa, A. (2014). What is a memory schema? A historical perspective on current neuroscience literature. *Neuropsychologia*. https://doi.org/10.1016/j.neuropsychologia.2013.11.010

Girardeau, G., Benchenane, K., Wiener, S. I., Buzsáki, G., & Zugaro, M. B. (2009). Selective suppression of hippocampal ripples impairs spatial memory. *Nature Neuroscience*, *12*(10), 1222–1223. https://doi.org/10.1038/nn.2384

Griessenberger, H., Heib, D. P. J., Lechinger, J., Luketina, N., Petzka, M., Moeckel, T., … Schabus, M. (2013). Susceptibility to Declarative Memory Interference is Pronounced in Primary Insomnia. *PLoS ONE*, *8*(2). https://doi.org/10.1371/journal.pone.0057394

Girardeau, G., & Zugaro, M. (2011). Hippocampal ripples and memory consolidation. *Current Opinion in Neurobiology*. https://doi.org/10.1016/j.conb.2011.02.005

Glenville, M., Broughton, R., Wing, a M., & Wilkinson, R. T. (1978). Effects of sleep deprivation on short duration performance measures compared to the Wilkinson auditory vigilance task. *Sleep*, *1*(2), 169–76. https://doi.org/10.1093/sleep/1.2.169

Groch, S., Schreiner, T., Rasch, B., Huber, R., & Wilhelm, I. (2017). Prior knowledge is essential for the beneficial effect of targeted memory reactivation during sleep. *Scientific Reports*, *7*, 39763. https://doi.org/10.1038/srep39763

Grosvenor, a, & Lack, L. C. (1984). The effect of sleep before or after learning on memory. *Sleep*, *7*(2), 155–167.

Hanert, A., Weber, F. D., Pedersen, A., Born, J. and Bartsch, T. (2017). Sleep in humans stabilizes pattern separation performance. Journal of Neuroscience, 1189-17.

Harrison, Y., & Espelid, E. (2004). Loss of negative priming following sleep deprivation. *Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology*, *57*(3), 437–446. https://doi.org/10.1080/02724980343000224

Hashmi, A., Nere, A., & Tononi, G. (2013). Sleep-dependent synaptic down-selection (II): Single-neuron level benefits for matching, selectivity, and specificity. *Frontiers in Neurology*, *4 OCT*. https://doi.org/10.3389/fneur.2013.00148

Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences*. https://doi.org/10.1016/S1364-6613(00)01482-0

Head, H., & Holmes, G. (1911). Sensory Disturbances From Cerebral Lesions. *Brain*, *34*(2–3), 102–254. https://doi.org/10.1093/brain/34.2-3.102

Hennies, N., Lambon Ralph, M. A., Kempkes, M., Cousins, J. N., & Lewis, P. A. (2016). Sleep Spindle Density Predicts the Effect of Prior Knowledge on Memory Consolidation. *The Journal of Neuroscience*, *36*(13), 3799–3810. https://doi.org/10.1523/JNEUROSCI.3162-15.2016

Hoddes, E., Zarcone, V., Smythe, H., Phillips, R., & Dement, W. C. (1973). Quantification of Sleepiness: A New Approach. *Psychophysiology*, *10*(4), 431–436. https://doi.org/10.1111/j.1469-8986.1973.tb00801.x

Howe, M. L., Garner, S. R., Charlesworth, M., & Knott, L. (2011). A brighter side to memory illusions: False memories prime children's and adults' insight-based problem solving. *Journal of Experimental Child Psychology*, *108*(2), 383–393. https://doi.org/10.1016/j.jecp.2010.08.012

Iber C, Ancoli-Israel S, Chesson AL, Quan SF. (2007). The AASM Manual for the Scoring of Sleep and Associated Events. Rules, terminology and technical specifications.. American Academy of Sleep Medicine. Westchester: IL.

Iidaka, T., Harada, T., Kawaguchi, J., & Sadato, N. (2012). Neuroanatomical substrates involved in true and false memories for face. *NeuroImage*, *62*(1), 167–176. https://doi.org/10.1016/j.neuroimage.2012.04.044

St. Jacques, P. L., Olm, C., & Schacter, D. L. (2013). Neural mechanisms of reactivation-induced updating that enhance and distort memory. *Proceedings of*

the National Academy of Sciences, *110*(49), 19671–19678. https://doi.org/10.1073/pnas.1319630110

Jadhav, S. P., Kemere, C., German, P. W., & Frank, L. M. (2012). Awake Hippocampal Sharp-Wave Ripples Support Spatial Memory. *Science*, *336*(6087), 1454–1458. https://doi.org/10.1126/science.1217230

Jenkins, J. G., & Dallenbach, K. M. (1924). Obliviscence during Sleep and Waking. *The American Journal of Psychology*, *35*(4), 605–612. https://doi.org/10.2307/1414040

Ji, D., & Wilson, M. A. (2007). Coordinated memory replay in the visual cortex and hippocampus during sleep. *Nature Neuroscience*, *10*(1), 100–107. https://doi.org/10.1038/nn1825

Kanwisher, N., McDermott, J., and Chun, M. M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. Journal of Neuroscience, 17, 4302-11.

Kanwisher, N. and Yovel , G. (2006). The fusiform face area: a cortical region specialized for the perception of faces. Philosophical Transactions of the Royal Society B, 361, 2109-28.

Keenan, J. P., Wheeler, M., Platek, S. M., Lardi, G., & Lassonde, M. (2003). Self-face processing in a callosotomy patient. *European Journal of Neuroscience*, *18*(8), 2391–2395. https://doi.org/10.1046/j.1460-9568.2003.02958.x

Kim, E. Y., Mahmoud, G. S., & Grover, L. M. (2005). REM sleep deprivation inhibits LTP in vivo in area CA1 of rat hippocampus. *Neuroscience Letters*, *388*(3), 163–167. https://doi.org/10.1016/j.neulet.2005.06.057

Kirchner, W. K. (1958). Age differences in short-term retention of rapidly changing information. *Journal of Experimental Psychology*, *55*(4), 352–358. https://doi.org/10.1037/h0043688

Koutstaal, W., & Schacter, D. L. (1997). Gist-Based False Recognition of Pictures in Older and Younger Adults. *Journal of Memory and Language*, *37*(4), 555–583. https://doi.org/10.1006/jmla.1997.2529

Kriegeskorte, N. (2008). Representational similarity analysis – connecting the branches of systems neuroscience. *Frontiers in Systems Neuroscience*. https://doi.org/10.3389/neuro.06.004.2008

Lansink, C. S., Goltstein, P. M., Lankelma, J. V., Joosten, R. N. J. M. A., McNaughton, B. L., & Pennartz, C. M. A. (2008). Preferential Reactivation of Motivationally Relevant Information in the Ventral Striatum. *Journal of Neuroscience*, *28*(25), 6372–6382. https://doi.org/10.1523/JNEUROSCI.1054-08.2008

Lau, H., Alger, S. E., & Fishbein, W. (2011). Relational memory: A daytime nap facilitates the abstraction of general concepts. *PLoS ONE*, *6*(11). https://doi.org/10.1371/journal.pone.0027139

Lau, H., Tucker, M. A., & Fishbein, W. (2010). Daytime napping: Effects on human direct associative and relational memory. *Neurobiology of Learning and Memory*, *93*(4), 554–560. https://doi.org/10.1016/j.nlm.2010.02.003

Lewis, P. A., & Durrant, S. J. (2011). Overlapping memory replay during sleep builds cognitive schemata. *Trends in Cognitive Sciences*. https://doi.org/10.1016/j.tics.2011.06.004

Lo, J. C., Sim, S. K. Y., & Chee, M. W. L. (2014). Sleep Reduces False Memory in Healthy Older Adults. *Sleep*, *37*(4), 665–671. https://doi.org/10.5665/sleep.3564

Loftus, E. F., & Pickrell, J. E. (1995). The formation of false memories. *Psychiatric Annals*. https://doi.org/10.1016/S0193-953X(05)70059-9

Lundqvist, D., Flykt, A., & Ohman, A. (1998). The Karolinska directed emotional faces (KDEF). *CD ROM from Department of Clinical Neuroscience, Psychology Section, Karolinska Institutet*. https://doi.org/10.1017/S0048577299971664

Lutz, N. D., Diekelmann, S., Hinse-Stern, P., Born, J., & Rauss, K. (2017). Sleep Supports the Slow Abstraction of Gist from Visual Perceptual Memories. *Scientific Reports*, *7*, 42950. https://doi.org/10.1038/srep42950

Mandler, J. M. (1984). Stories, Scripts, and Scenes: Aspects of Schema Theory. Hillsdale, NJ : Lawrence Erlbaum.

Marr, D. (1970). A theory for cerebral neocortex. *Proceedings of the Royal Society of London. Series B, Containing Papers of a Biological Character. Royal Society (Great Britain)*, *176*(1043), 161–234. https://doi.org/10.1098/rspb.1970.0040

Marr, D. (1971). Simple Memory: A Theory for Archicortex. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *262*(841), 23–81. https://doi.org/10.1098/rstb.1971.0078

Marshall, L., & Born, J. (2007). The contribution of sleep to hippocampus-dependent memory consolidation. *Trends in Cognitive Sciences*. https://doi.org/10.1016/j.tics.2007.09.001

McClelland, J. L. (2013). Incorporating rapid neocortical learning of new schema-consistent information into complementary learning systems theory. *Journal of Experimental Psychology: General*, *142*(4), 1190–1210. https://doi.org/10.1037/a0033812

McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychological Review*, *102*(3), 419–57. https://doi.org/10.1037/0033-295X.102.3.419

McDermott, K. B. (1996). The Persistence of False Memories in List Recall. *Journal of Memory and Language*, *35*(2), 212–230. https://doi.org/10.1006/jmla.1996.0012

McGaugh, J. L. (2000). Memory--a Century of Consolidation. *Science*, *287*(5451), 248–251. https://doi.org/10.1126/science.287.5451.248

McKeon, S., Pace-Schott, E. F., & Spencer, R. M. C. (2012). Interaction of Sleep and Emotional Content on the Production of False Memories. *PLoS ONE*, *7*(11). https://doi.org/10.1371/journal.pone.0049353

Melo, B., Winocur, G., & Moscovitch, M. (1999). False Recall and False Recognition: an Examination of the Effects of Selective and Combined Lesions To the Medial Temporal Lobe/Diencephalon and Frontal Lobe Structures. *Cognitive Neuropsychology*, *16*(3–5), 343–359. https://doi.org/10.1080/026432999380825

Melton, A. W. (1963). Implications ol Short-Term Memory for a General Theory of Memory. *Journalof Verbal Learning and Verbal Behavior*, *2*(1), 1–21. https://doi.org/10.1017/CBO9781107415324.004

Milner, B. (1962). Les troubles de la mémoire accompagnant des lésions hippocampiques bilatérales. In *Physiologie de l'hippocampe* (pp. 257–272).

Mölle, M., Marshall, L., Gais, S., & Born, J. (2002). Grouping of spindle activity during slow oscillations in human non-rapid eye movement sleep. *The Journal of Neuroscience*, *22*(24), 10941–10947. https://doi.org/22/24/10941 [pii]

Mölle, M., Yeshenko, O., Marshall, L., Sara, S. J., & Born, J. (2006). Hippocampal Sharp Wave-Ripples Linked to Slow Oscillations in Rat Slow-Wave Sleep. *Journal of Neurophysiology*, *96*(1), 62–70. https://doi.org/10.1152/jn.00014.2006

Moscovitch, M., & Melo, B. (1997). Strategic retrieval and the frontal lobes: Evidence from confabulation and amnesia. *Neuropsychologia*, *35*(7), 1017–1034. https://doi.org/10.1016/S0028-3932(97)00028-6

MÜlberger, A. (2017). MENTAL ASSOCIATION: TESTING INDIVIDUAL DIFFERENCES BEFORE BINET. *Journal of the History of the Behavioral Sciences*, *53*(2), 176–198. https://doi.org/10.1002/jhbs.21850

Nádasdy, Z., Hirase, H., Czurkó, A., Csicsvari, J., & Buzsáki, G. (1999). Replay and Time Compression of Recurring Spike Sequences in the Hippocampus. *The Journal of Neuroscience*, *19*(21), 9497–9507. https://doi.org/10.1126/science.1182395

Nadel, L., & Moscovitch, M. (1997). Memory consolidation, retrograde amnesia and the hippocampal complex. *Current Opinion in Neurobiology*. https://doi.org/10.1016/S0959-4388(97)80010-4

Nere, A., Hashmi, A., Cirelli, C., & Tononi, G. (2013). Sleep-dependent synaptic down-selection (I): Modeling the benefits of sleep on memory consolidation and integration. *Frontiers in Neurology*, *4 SEP*. https://doi.org/10.3389/fneur.2013.00143

Nevi, A., Cicali, F., & Caudek, C. (2016). The role of familiarity on viewpoint adaptation for self-face and other-face images. *Perception*, *45*(7), 823–843. https://doi.org/10.1177/0301006616643661

Nieuwenhuis, I. L. C., Folia, V., Forkstam, C., Jensen, O., & Petersson, K. M. (2013). Sleep Promotes the Extraction of Grammatical Rules. *PLoS ONE*, *8*(6). https://doi.org/10.1371/journal.pone.0065046

O'Neill, J., Pleydell-Bouverie, B., Dupret, D., & Csicsvari, J. (2010). Play it again: reactivation of waking experience and memory. *Trends in Neurosciences*. https://doi.org/10.1016/j.tins.2010.01.006

Olewuezi, N. P. (2011). Note on the Comparison of Some Outlier Labeling Techniques. *Journal of Mathematics and Statistics 7(4)*, 353-355.

Palchykova, S., Winsky-Sommerer, R., Meerlo, P., Dürr, R., & Tobler, I. (2006). Sleep deprivation impairs object recognition in mice. *Neurobiology of Learning and Memory*, *85*(3), 263–271. https://doi.org/10.1016/j.nlm.2005.11.005

Pardilla-Delgado, E., & Payne, J. D. (2017). The impact of sleep on true and false memory across long delays. *Neurobiology of Learning and Memory*, *137*, 123–133. https://doi.org/10.1016/j.nlm.2016.11.016

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. https://doi.org/10.1523/jneurosci.4342-04.2005

Payne, J. D., Schacter, D. L., Propper, R. E., Huang, L. W., Wamsley, E. J., Tucker, M. A., … Stickgold, R. (2009). The role of sleep in false memory formation. *Neurobiology of Learning and Memory*, *92*(3), 327–334. https://doi.org/10.1016/j.nlm.2009.03.007

Peigneux, P., Laureys, S., Fuchs, S., Collette, F., Perrin, F., Reggers, J., … Maquet, P. (2004). Are spatial memories strengthened in the human hippocampus during

slow wave sleep? *Neuron*, *44*(3), 535–545. https://doi.org/10.1016/j.neuron.2004.10.007

Piaget, J. (1952). The origins of intelligence in children. New York: International Universities Press

Piaget. J. (1953). The origins of intelligence in children. *Journal of Consulting Psychology*, *17*(6), 467–467. https://doi.org/10.1037/h0051916

Poe, G. R. (2017). Sleep Is for Forgetting. *The Journal of Neuroscience*, *37*(3), 464–473. https://doi.org/10.1523/JNEUROSCI.0820-16.2017

Rasch, B., & Born, J. (2008). Reactivation and consolidation of memory during sleep. *Current Directions in Psychological Science*. https://doi.org/10.1111/j.1467-8721.2008.00572.x

Rasch, B., & Born, J. (2013). About sleep's role in memory. *Physiological Reviews*, *93*(2), 681–766. https://doi.org/10.1152/physrev.00032.2012

Rasch, B., Buchel, C., Gais, S., & Born, J. (2007). Odor Cues During Slow-Wave Sleep Prompt Declarative Memory Consolidation. *Science*, *315*(5817), 1426–1429. https://doi.org/10.1126/science.1138581

Rechtschaffen, A., & Kales, A. (1968). A manual of standardized techniques and scoring system for sleep stages of human subjects. *Washington, D.C.: U.S. Government Printing Office*, *NIH Public*, 12. https://doi.org/10.1016/0013-4694(69)90021-2

Richler, J. J., Mack, M. L., Palmeri, T. J., & Gauthier, I. (2011). Inverted faces are (eventually) processed holistically. *Vision Research*, *51*(3), 333–342. https://doi.org/10.1016/j.visres.2010.11.014

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

Rumelhart, D. E. (1995). Brain style computation: Learning and generalization. In *An introduction to neural and electronic networks (2nd ed.). Neural networks: Foundations to applications* (p. 431–446 ST–Brain style computation: Learning an).

Sanchez-Vives, M. V, & McCormick, D. A. (2000). Cellular and network mechanisms of rhythmic recurrent activity in neocortex. *Nature Neuroscience*, *3*(10), 1027–1034. https://doi.org/10.1038/79848

Schabus, M., Gruber, G., Parapatics, S., Sauter, C., Klösch, G., Anderer, P., … Zeitlhofer, J. (2004). Sleep spindles and their significance for declarative memory consolidation. *Sleep*, *27*(8), 1479–1485. https://doi.org/10.1093/sleep/27.7.1479

Schacter, D. L. (1999). The seven sins of memory: Insights from psychology and cognitive neuroscience. *American Psychologist*, *54*(3), 182–203. https://doi.org/10.1037//0003-066X.54.3.182

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. https://doi.org/10.1098/rstb.2007.2087

Schacter, D. L., Guerin, S. A., & St. Jacques, P. L. (2011). Memory distortion: An adaptive perspective. *Trends in Cognitive Sciences*, *15*(10), 467–474. https://doi.org/10.1016/j.tics.2011.08.004

Schacter, D. L., Norman, K. A., & Koutstaal, W. (1998). The cognitive neuroscience of constructive memory. *Annu Rev Psychol*, *49*, 289–318. https://doi.org/10.1146/annurev.psych.49.1.289

Scholpp, S., Delogu, A., Gilthorpe, J., Peukert, D., Schindler, S., & Lumsden, A. (2009). Her6 regulates the neurogenetic gradient and neuronal identity in the thalamus. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(47), 19895–900. https://doi.org/10.1073/pnas.0910894106

Scoville, W. B., & Milner, B. (1957). Loss of recentmemory after bilateral hippocampal lesions. *Journal of Neurology, Neurosurgery & Psychiatry*, *20*(1), 11–21. https://doi.org/10.1136/jnnp.20.1.11

Seamon, J. G., Luo, C. R., Kopecky, J. J., Price, C. A., Rothschild, L., Fung, N. S., & Schwartz, M. A. (2002). Are false memories more difficult to forget than accurate memories? The effect of retention interval on recall and recognition. *Memory & Cognition*, *30*(7), 1054–1064. https://doi.org/10.3758/BF03194323

Sheth, B. R., Varghese, R., & Truong, T. (2012). Sleep Shelters Verbal Memory from Different Kinds of Interference. *SLEEP*. https://doi.org/10.5665/sleep.1966

Silvestri, A. J. (2005). REM sleep deprivation affects extinction of cued but not contextual fear conditioning. *Physiology and Behavior*, *84*(3), 343–349. https://doi.org/10.1016/j.physbeh.2004.11.011

Sirota, A., Csicsvari, J., Buhl, D., & Buzsaki, G. (2003). Communication between neocortex and hippocampus during sleep in rodents. *Proceedings of the National Academy of Sciences*, *100*(4), 2065–2069. https://doi.org/10.1073/pnas.0437938100

Slotnick, S. D., & Schacter, D. L. (2004). A sensory signature that distinguishes true from false memories. *Nature Neuroscience*, *7*(6), 664–672. https://doi.org/10.1038/nn1252

Smith, C. N., & Squire, L. R. (2008). Experience-Dependent Eye Movements Reflect Hippocampus-Dependent (Aware) Memory. *Journal of Neuroscience*, *28*(48), 12825–12833. https://doi.org/10.1523/JNEUROSCI.4542-08.2008

Spoormaker, V. I., Sturm, A., Andrade, K. C., Schröter, M. S., Goya-Maldonado, R., Holsboer, F., … Czisch, M. (2010). The neural correlates and temporal sequence of the relationship between shock exposure, disturbed sleep and impaired consolidation of fear extinction. *Journal of Psychiatric Research*, *44*(16), 1121–1128. https://doi.org/10.1016/j.jpsychires.2010.04.017

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

Squire, L. R., & Dede, A. J. O. (2015). Conscious and unconscious memory systems. *Cold Spring Harbor Perspectives in Biology*, *7*(3), a021667. https://doi.org/10.1101/cshperspect.a021667

Squire, L. R., & Zola, S. M. (1998). Episodic memory, semantic memory, and amnesia. *Hippocampus*. https://doi.org/10.1002/(SICI)1098-1063(1998)8:3<205::AID-HIPO3>3.0.CO;2-I

St George, M., Mannes, S., & Hoffman, J. E. (1997). Individual Differences in Inference Generation: An ERP Analysis. *Journal of Cognitive Neuroscience*, *9*(6), 776–787. https://doi.org/10.1162/jocn.1997.9.6.776

Steriade, M. (2005). Sleep, epilepsy and thalamic reticular inhibitory neurons. *Trends in Neurosciences*. https://doi.org/10.1016/j.tins.2005.03.007

Steriade, M., McCormick, D., & Sejnowski, T. (1993). Thalamocortical oscillations in the sleeping and aroused brain. *Science*, *262*(5134), 679–685. https://doi.org/10.1126/science.8235588

Sterpenich, V., Schmidt, C., Albouy, G., Matarazzo, L., Vanhaudenhuyse, A., Boveroux, P., … Maquet, P. (2014). Memory Reactivation during Rapid Eye Movement Sleep Promotes Its Generalization and Integration in Cortical Stores. *Sleep*, *37*(6), 1061–1075. https://doi.org/10.5665/sleep.3762

Stickgold, R. (2005). Sleep-dependent memory consolidation. *Nature*, *437*(7063), 1272–1278. https://doi.org/10.1038/nature04286

Stickgold, R., & Walker, M. P. (2013). Sleep-dependent memory triage: evolving generalization through selective processing. *Nature Neuroscience*, *16*(2), 139–145. https://doi.org/10.1038/nn.3303

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. https://doi.org/10.1162/089892900562075

Suzuki, W. A. (2006). Encoding New Episodes and Making Them Stick. *Neuron*. https://doi.org/10.1016/j.neuron.2006.03.029

Takashima, A., Nieuwenhuis, I. L. C., Jensen, O., Talamini, L. M., Rijpkema, M., & Fernandez, G. (2009). Shift from Hippocampal to Neocortical Centered Retrieval Network with Consolidation. *Journal of Neuroscience*, *29*(32), 10087–10093. https://doi.org/10.1523/JNEUROSCI.0799-09.2009

Tamminen, J., Payne, J. D., Stickgold, R., Wamsley, E. J., & Gaskell, M. G. (2010). Sleep Spindle Activity is Associated with the Integration of New Memories and Existing Knowledge. *Journal of Neuroscience*, *30*(43), 14356–14360. https://doi.org/10.1523/JNEUROSCI.3028-10.2010

Tamminen, J., Ralph, M. A. L., & Lewis, P. A. (2013). The role of sleep spindles and slow-wave activity in integrating new information in semantic memory. *The*

Journal of Neuroscience, *33*(39), 15376–15381. https://doi.org/10.1523/jneurosci.5093-12.2013

Taxidis, J., Mizuseki, K., Mason, R., & Owen, M. R. (2013). Influence of slow oscillation on hippocampal activity and ripples through cortico-hippocampal synaptic interactions, analyzed by a cortical-CA3-CA1 network model. *Frontiers in Computational Neuroscience*, *7*. https://doi.org/10.3389/fncom.2013.00003

Tomita, A., Yamamoto, S., Matsushita, S., & Morikawa, K. (2014). Resemblance to familiar faces is exaggerated in memory. *Japanese Psychological Research*, *56*(1), 24–32. https://doi.org/10.1111/jpr.12032

Tononi, G., & Cirelli, C. (2006). Sleep function and synaptic homeostasis. *Sleep Medicine Reviews*. https://doi.org/10.1016/j.smrv.2005.05.002

Tse, D., Langston, R. F., Kakeyama, M., Bethus, I., Spooner, P. A., Wood, E. R., … Morris, R. G. M. (2007). Schemas and Memory Consolidation. *Science*, *316*(5821), 76–82. https://doi.org/10.1126/science.1135935

Tse, D., Takeuchi, T., Kakeyama, M., Kajii, Y., Okuno, H., Tohyama, C., … Morris, R. G. M. (2011). Schema-Dependent Gene Activation and Memory Encoding in Neocortex. *Science*, *333*(6044), 891–895. https://doi.org/10.1126/science.1205274

Tsujimura, H., Kotz S., Lander K., Lewis P., (2018). Neural correlates of sleep role in curve-fitted face recognition memory.

Tsujimura, H., Kotz S., Lander K., Stamper T., Garlick E., Lewis P., (2018). The role of sleep in curve-fitted face recognition memory.

Tsujimura, H., Kotz S., Lander K., Tafuro A., Lewis P. A., (2018). Sleep-dependent false recognition of facial images is associated with slow wave sleep.

Tsujimura, H., Lewis, P. A., (2018). Investigating the role of sleep on the gist-based face recognition memory.

Tulving, E. (1972). Episodic and semantic memory. *Organization of Memory*. https://doi.org/10.1017/S0140525X00047257

Van Alphen, B., Yap, M. H. W., Kirszenblat, L., Kottler, B., & van Swinderen, B. (2013). A Dynamic Deep Sleep Stage in Drosophila. *Journal of Neuroscience*, *33*(16), 6917–6927. https://doi.org/10.1523/JNEUROSCI.0061-13.2013

Van der Helm, E., Gujar, N., Nishida, M., & Walker, M. P. (2011). Sleep-Dependent Facilitation of Episodic Memory Details. *PLoS ONE*, *6*(11). https://doi.org/10.1371/journal.pone.0027421

Van Kesteren, M. T. R., Beul, S. F., Takashima, A., Henson, R. N., Ruiter, D. J., & Fernández, G. (2013). Differential roles for medial prefrontal and medial temporal cortices in schema-dependent encoding: From congruent to incongruent. *Neuropsychologia*, *51*(12), 2352–2359. https://doi.org/10.1016/j.neuropsychologia.2013.05.027

Van Kesteren, M. T. R., Fernandez, G., Norris, D. G., & Hermans, E. J. (2010). Persistent schema-dependent hippocampal-neocortical connectivity during memory encoding and postencoding rest in humans. *Proceedings of the National Academy of Sciences*, *107*(16), 7550–7555. https://doi.org/10.1073/pnas.0914892107

Van Kesteren, M. T. R., Rijpkema, M., Ruiter, D. J., & Fernandez, G. (2010). Retrieval of Associative Information Congruent with Prior Knowledge Is Related to Increased Medial Prefrontal Activity and Connectivity. *Journal of Neuroscience*, *30*(47), 15888–15894. https://doi.org/10.1523/JNEUROSCI.2674-10.2010

Van Kesteren, M. T. R., Rijpkema, M., Ruiter, D. J., Morris, R. G. M., & Fernández, G. (2014). Building on Prior Knowledge: Schema-dependent Encoding Processes Relate to Academic Performance. *Journal of Cognitive Neuroscience*, *26*(10), 2250–2261. https://doi.org/10.1162/jocn

Wagner, U., Gais, S., Haider, H., Verleger, R., & Born, J. (2004). Sleep inspires insight. *Nature*, *427*(6972), 352–355. https://doi.org/10.1038/nature02223

Walker, M. P., & Stickgold, R. (2004). Sleep-dependent learning and memory consolidation. *Neuron*. https://doi.org/10.1016/j.neuron.2004.08.031

Warren, D. E., Jones, S. H., Duff, M. C., & Tranel, D. (2014). False Recall Is Reduced by Damage to the Ventromedial Prefrontal Cortex: Implications for Understanding the Neural Correlates of Schematic Memory. *Journal of Neuroscience*, *34*(22), 7677–7682. https://doi.org/10.1523/JNEUROSCI.0119-14.2014

Werchan, D. M., & Gómez, R. L. (2013). Generalizing memories over time: Sleep and reinforcement facilitate transitive inference. *Neurobiology of Learning and Memory*, *100*, 70–76. https://doi.org/10.1016/j.nlm.2012.12.006

Wierzynski, C. M., Lubenov, E. V., Gu, M., & Siapas, A. G. (2009). State-Dependent Spike-Timing Relationships between Hippocampal and Prefrontal Circuits during Sleep. *Neuron*, *61*(4), 587–596. https://doi.org/10.1016/j.neuron.2009.01.011

Wilford, M. M., & Wells, G. L. (2010). Does Facial Processing Prioritize Change Detection? Change Blindness Illustrates Costs and Benefits of Holistic Processing. *Psychological Science*, *21*(11), 1611–1615. https://doi.org/10.1177/0956797610385952

Wilhelm, I., Diekelmann, S., Molzow, I., Ayoub, A., Molle, M., & Born, J. (2011). Sleep Selectively Enhances Memory Expected to Be of Future Relevance. *Journal of Neuroscience*, *31*(5), 1563–1569. https://doi.org/10.1523/JNEUROSCI.3575-10.2011

Wilson, M., & McNaughton, B. (1994). Reactivation of hippocampal ensemble memories during sleep. *Science*, *265*(5172), 676–679. https://doi.org/10.1126/science.8036517

Xie, L., Kang, H., Xu, Q., Chen, M. J., Liao, Y., Thiyagarajan, M., … Nedergaard, M. (2013). Sleep Drives Metabolite Clearance from the Adult Brain. *Science*, *342*(6156), 373–377. https://doi.org/10.1126/science.1241224