Light sleep versus slow wave sleep in memory consolidation

Citation for published version:

Digital Object Identifier (DOI):

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Peer reviewed version

Published In:
Trends in Neurosciences

Publisher Rights Statement:
Author's final peer reviewed manuscript as accepted for publication

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
Light sleep vs. slow wave sleep in memory consolidation: A question of global vs. local processes?

Genzel L\textsuperscript{1}, Kroes MCW\textsuperscript{2}, Dresler M\textsuperscript{3}, Battaglia FP\textsuperscript{2,4}

\textsuperscript{1}CCNS, University of Edinburgh, Edinburgh/United Kingdom
\textsuperscript{2}Donders Institute for Brain, Cognition, and Behaviour, Radboud University Nijmegen, Nijmegen, The Netherlands
\textsuperscript{3}Max Planck Institute of Psychiatry, Munich/Germany
\textsuperscript{4}NERF/VIB Leuven Belgium

\textbf{Keywords:} memory consolidation, sleep, sleep stages, replay, downscaling, SWR, spindles

Correspondence should be addressed to L.G. (lgenzel@ed.ac.uk) and F.P.B. (F.Battaglia@science.ru.nl)

\textbf{Abstract}
Sleep is strongly involved in memory consolidation, but its exact role remains unclear. ‘Sleep replay’, the active potentiation of relevant synaptic connections via reactivation of patterns of network activity that occurred during previous experience, has received considerable attention. Alternatively, sleep has been suggested to homeostatically and nonspecifically regulate synaptic weights, thereby improving the signal-to-noise ratio of memory traces. Here, we attempt to reconcile these theories by highlighting the distinction between light and deep non-rapid eye movement (NREM) sleep. Specifically, we propose that studies in humans and animals suggest a link between light NREM and active potentiation, and between deep NREM and homeostatic regulation. Finally, we suggest this framework could serve as a key to interpreting the physiology of sleep stages.

\textbf{Sleep and memory: the need for an updated picture}
In roughly 100 years of research sleep has been shown to be beneficial for memory in a wide variety of tasks and species, e.g. insects, birds, rodents, and humans \textsuperscript{1-3}. In humans, positive effects of sleep have been reported for declarative memory \textsuperscript{4}, motor memory \textsuperscript{5,6}, visual discrimination \textsuperscript{7}, and many other tasks \textsuperscript{1,8}. However, there is still an intense debate about the mechanisms of this involvement. Contrasting hypotheses see sleep either as a moment of reprocessing, i.e. reinforcing and reorganisation of specific information and memory traces or the ‘active’ role of sleep \textsuperscript{1,9}, or as conducive to non-specific homeostatic processes, solving
neurobiological imbalances accumulated during prolonged waking periods, and restoring a suitable 'working point' for brain networks to operate (the 'downscaling' hypothesis\textsuperscript{10}).

Heterogeneous data, from human subjects as well as invasive neurophysiology experiments in animals, contributed to a detailed picture of sleep processes and their role in memory consolidation. However, different techniques provide views of brain processes at very different spatial and temporal scales, in different experimental contexts. A number of theoretical efforts have attempted to reconcile these variegated data. This is however a difficult task, partly because of confusing use of terminology across the field. As a key example, sleep is usually dissected in sequential 'stages' based on surface electroencephalography (EEG) recordings \textsuperscript{11} (Figure 1). Sleep stages are often taken as a proxy for the underlying physiological processes, and ascribed different functions in the memory consolidation process. The problem with this view is twofold. First, surface EEG provides a very partial picture of the underlying physiological phenomena, and the same dynamical phenomena appear, albeit with different frequency, during different sleep stages. Second, researchers investigating human and animal sleep use different terminology; for example, while slow wave sleep (SWS) constitutes solely deep sleep (stages 3+4) in humans, in animals it is customary to use this term to describe all Non-REM sleep. This has caused fundamental misunderstandings and makes conclusions from stage-based analysis of sleep and from more temporally precise invasive physiological experiments difficult to compare.

Here, we offer a proposal for how to recast the "sleep and memory consolidation" problem in terms of the underlying physiological processes. The main hypothesis is that during sleep, different mechanisms may synchronize brain activity at a more local or a more global level, and the local-global gradient is the crucial axis determining the functional role of synchronization events for memory. As we will explain below, 'active' systems consolidation, i.e. the consolidation of memory via information exchange between the hippocampus and the neocortex\textsuperscript{33}, should require global synchronization and communication. In contrast, local activations may subserve other roles, e.g. homeostasis that may contribute to some aspects of memory consolidation. We will review existing literature in this light, formulating hypotheses that may help harmonize competing theoretical views of the sleep/memory link. Finally, we will argue under which aspects currently available evidence is lacking, and which experiments may be necessary to complete the picture.

\textit{Network physiology of sleep and possible significance for memory consolidation}\par
On the coarsest level, a night's sleep is subdivided into Rapid Eye Movement (REM – see Glossary) and Non-REM (NREM) sleep based on surface EEG, electro-oculogram (EOG) and electro-myographic (EMG) recordings. NREM is further subdivided into light sleep (LS),
and deep or SWS; (Fig. 1). REM, LS and SWS all differ in terms of their neural dynamics\textsuperscript{12,13}. REM exhibits mostly desynchronized dynamics, more similar to that observed in the waking state. Critically, NREM sleep is considered the key stage for memory consolidation\textsuperscript{1}.

Invasive electrophysiology provides detailed insight into the neural dynamics have been associated with distinct sleep stages and whose surface EEG correlates are a major element for sleep scoring. Cortical slow oscillations (SO) are synchronous events affecting large cortical territories and predominantly associated with NREM. SO are alternations between states of generalised cortical excitation and depolarised membrane potentials (UP states) and states of relative neuronal silence (DOWN states). UP states are thought to be generated by excitatory feedback between cortical neurons, and to be terminated by activity-suppressing phenomena at the cellular synaptic scale, such as synaptic depression and potassium conductances. Their critical importance for sleep dynamics is underlined by the fact that SO orchestrate many other faster cortical phenomena such as delta waves, spindles, and gamma oscillations\textsuperscript{14}. Equally important, SO influence activity in other brain structures such as the striatum, locus coeruleus, and the hippocampus, possibly coordinating interactions between the neocortex and these brain regions.

The dominant activity pattern in the hippocampus (which is critical for memory consolidation\textsuperscript{15}) during NREM is the sharp wave-ripples (SWR) complex, a large burst activating up to 30% of all hippocampal cells in a 50-150 ms time window. SWRs are temporally linked to cortical SO; the latter were found to modulate membrane potential, firing rate and SWR probability (higher during UP state) in several hippocampal subfields\textsuperscript{16-18}. In turn, SWR are correlated with transient increases in cortical firing and tend to occur mostly at the transitions between DOWN and UP states\textsuperscript{19-21}, highlighting a likely bi-directional dialogue between cortex and hippocampus in the sleeping brain. Importantly, SWR have been associated with active memory replay (to be discussed below), suggestive of a role in active systems consolidation.

Critically, while all of NREM sleep is dominated by the UP/DOWN state bi-stability, there is increasing evidence for different types of slow events (i.e. SO, delta waves), some of them more confined to one brain area, and others more global, spanning large parts of the brain. The occurrences of local and global slow events are more or less common during different NREM stages. First, in cat and rodent models short DOWN states stochastically interspersed amongst on-going activity (sustained UP states), giving rise to isolated K-complexes\textsuperscript{14,22} and bouts of spindle oscillations, can be observed mainly during periods identified as LS. Second, more periodic alternation of UP and DOWN states, giving rise to EEG delta waves, is a hallmark of SWS. Although SO account for over 30% of periods classified as SWS and are sparser during LS, the synchronizations during LS tend to be more global than during SWS.
Human studies confirm this idea. Connectivity analyses of functional Magnetic Resonance Imaging (fMRI) data show that SWS is accompanied by a breakdown in cortico-cortical connectivity and presents with more local clustering\textsuperscript{12,13}. This “local” sleep during SWS has been confirmed in detailed analysis of intra-cranial EEG patterns\textsuperscript{17} and transcranial magnetic stimulation experiments\textsuperscript{23}. Conversely, LS actually shows increased cortico-cortical connectivity, with whole brain functional systems such as the default mode network still relatively intact, in comparison to waking\textsuperscript{13}. Furthermore, in humans, SWRs are more frequent during LS than SWS\textsuperscript{24}.

We propose here that global and local events play distinct roles in memory consolidation. Global events may help promote the reorganization of memory traces throughout the cortex and between cortex and hippocampus\textsuperscript{25}, via the ‘replay’ of neural patterns related to those traces resonating across brain areas. ‘Local’ synchronizations seem less suitable for this task and may subserve other purposes, such as memory consolidation within a brain area, and/or the homeostatic downregulation of synaptic strengths. Thus, we propose that the differences in memory consolidation correlates in LS and SWS that we review below may arise due to dissimilar global interaction strength during those stages. In the next section we will discuss the physiological sleep processes that support active systems consolidation.

*Figure 1 please near here*

**Sleep replay as an active mechanism of memory consolidation**

Here we will argue that sleep replay serves as an active mechanism of memory consolidation and occurs in concert with global SO (K-complexes) and SWR events associated with LS.

One success of high density neural recording in animals, allowing monitoring of the spike trains from tens or hundreds of single neurons, is the discovery of the so-called ‘replay’ phenomenon\textsuperscript{9,26} - the spontaneous reactivation, during sleep, of neural activity patterns that occurred during previous experience. Replay is currently the most enticing candidate mechanism supporting an active role of sleep in memory maintenance and consolidation\textsuperscript{27-29}, and has been extensively characterised in the hippocampus\textsuperscript{9,26}. Subsequently, it was also found in other structures such as the neocortex\textsuperscript{30,31} and the striatum\textsuperscript{32}. The link between hippocampal and cortical replay appears especially important for a possible memory function of sleep, as the standard model of systems memory consolidation assumes that memories are consolidated by exchange of information between the hippocampus (the initial site of memory acquisition) and the neocortex (the final memory store and already active during
memory acquisition). Global activity phenomena may play an important role as the ‘carrier’ of this information exchange. In the hippocampus, sharp waves-ripple bursts (SWR) (see glossary box for definitions) have been theoretically related to the retrieval of information stored in the auto-associative networks of the hippocampus, possibly as the information source for hippocampal based consolidation. In fact, hippocampal replay mostly occurs during sleep SWR and cortical replay is also strongest during hippocampal SWR. Thus, these bursts may indeed contribute an input to cortical memory reprocessing during sleep. Cortical replay is also thought to be shaped by cortical slow oscillations, which occur at least in part synchronously with SWR.

Replay is likely to take place in human subjects as well. Recent studies using sophisticated analyses of positron-emission tomography (PET) and fMRI data have been able to identify signs of replay in sleeping humans. Further, an increase in performance has been observed following the presentation during sleep of sound or odour cues previously coupled to content learned during an encoding experiment. This could perhaps be due to neural replay triggered by the cue presentation during sleep. This method also revealed that replay triggered while subjects were awake seems to have the opposite effect. Instead of leading to consolidation and strengthening of memories, memories cued while subjects were awake were destabilized and prone to inference effects. Similarly, in rodents, SWR and thus replay disruption during waking has been associated with disruption of working memory.

However, an important difference between human and non-human animal studies must be highlighted. Sleep studies in rodents quite consistently discuss NREM sleep only as SWS, subdivision into LS and SWS is seldom made. Yet, based on what can be judged from published examples, most of these rodent studies are likely to mostly contain LS, in part because of the typically short duration of the sleep sessions (but see ). Indeed, isolated K-complexes, which are mostly found during LS, are linked to cortical replay in these studies. Furthermore, hippocampal replay tends to fade after the first ~30 minutes of sleep in rodents, an epoch which is unlikely to contain SWS. Evidence for replay during confirmed SWS in rodents is therefore scarce, and with it the possibility for a full comparison of extant sleep data across NREM stages between rodents and humans.

According to ‘active systems consolidation’ theory, replay would be well suited for driving whole-brain reorganization of memory traces, carried by long range activity correlations, which could be the necessary underpinning for the hippocampal/cortical information exchange and for the formation of cortex-wide cell assemblies (thought to be the final outcome of the systems consolidation process). In support of this hypothesis, firing patterns with a brisk alternation of depolarisations and hyperpolarisations such as those observed during K-complexes are found to be optimal to induce plasticity in cortical slices.
K-complexes also represent a privileged state for cortico-hippocampal communication, with a strong concentration of SWRs around them. In addition, in a study in which replay time course has been precisely characterized, cortical replay peaks in the hundreds of milliseconds around K-complexes. This effect has been measured in prefrontal cortex, a cortical area that receives a direct afferent from the hippocampus, and whose dynamical link with the hippocampus has been extensively demonstrated. Such K-complexes occur predominantly but not exclusively during LS rather than SWS. Thus, although replay most likely occurs throughout all NREM, it is probably most dominant during LS.

In conclusion, we suggest that replay could be the substrate for cortico-hippocampal information exchange and related plasticity phenomena, reflecting an active mechanism of sleep-dependent systems memory consolidation. Next, we will discuss the role of sleep spindles in memory consolidation.

On the role of sleep spindles in active memory consolidation

Many different patterns – slow oscillations, delta waves, K-complexes, and sleep spindles – may be observed from EEG sleep records, each the signature of a different network mechanism. Sleep spindles are oscillations that appear in waxing and waning bouts usually lasting up to a few seconds throughout the entirety of NREM sleep. Sleep spindles have been proposed as the key cortical process for active memory consolidation, as their occurrence increases after learning sessions in comparison to baseline and they are related to hippocampal SWRs. However, mere correlations between sleep spindles and memory performance are not enough to support the possible mechanistic role of sleep spindles in active memory consolidation. This requires analysis of the much more detailed data coming from invasive physiological experiments, mostly on non-human animals.

Spindles are oscillations generated in the thalamocortical network, often in response to the transition to a cortical UP state (notable as part of a K-complex). The pace maker for this sleep oscillation resides in the thalamus. Attributing a crucial role to spindles in active memory consolidation would therefore imply that the thalamocortical loop plays a dominant role in mnemonic processing. Perhaps an even greater role than that played by intrinsic cortical networks that express themselves in SO. The phase locking between spindles and SWRs recorded in the rat hippocampus and in human parahippocampal cortices supports this hypothesis. However, there are a number of experimental findings contrasting with this view. First, cortical cells (in particular pyramidal neurons) respond to sleep spindles with only modest firing increases and sometimes moderate phase locking. Second, in prefrontal cortex, inhibitory interneurons show the highest degree of recruitment during spindles, hinting at enhanced feed-forward inhibition. Consequently, spindles reduce prefrontal cortex responses to hippocampal SWRs. Third, memory replay (as opposed to...
population activity) peaks not during spindles but a few hundreds of ms earlier \(^{30}\). Corresponding to the strong recruitment of cortical inhibition found in rats, in humans, spindles have been linked to noise resistance during sleep \(^{60}\), and have been proposed to be important for sleep maintenance via suppression of noise processing \(^{61,62}\). Although increased hippocampal-prefrontal connectivity has been observed during sleep spindles in humans \(^{63}\), these effects may be due to the concomitant slow phenomena e.g. SO. Based on the available data, we therefore propose an alternative hypothesis: sleep spindles do not drive active memory consolidation, instead we hypothesize that this process is led by cortical slow oscillations triggering long-range interactions in the brain, SWR, replay, and, hence, active systems consolidation. Still, spindles are related to SO and are a useful index reflecting SO density in memory consolidation experiments.

We believe that the current uncertainty about the role of spindles is the product of technical limitations. ‘Slow’ techniques such as fMRI and EEG density measures (with time constants of \(\sim 1\) s, thus comparable with SO frequency, but 10 times slower than the frequency of spindles) cannot disentangle the effects of fast rhythms such as spindles (or SWR \(^{64}\)) from the slower processes that embed them and are correlated with them.

Whatever their status is with respect to network interactions, spindles may play an important enabling role for synaptic plasticity \(^{65}\). Spindles massively modulate membrane potential in cortical neurons\(^{58}\), and in-vitro artificial spindle-like stimulation has been shown to induce short-term and long-term potentiation in neocortical pyramid cells through a massive \(\text{Ca}^{2+}\) influx\(^{52}\). Thus, spindles could trigger cellular processes that favor long-term potentiation of patterns that are reactivated immediately before spindles. Furthermore, by effectively ‘deafferenting’ the cortex from thalamic \(^{66,52}\) and hippocampal inputs \(^{20}\), spindles may enable local, undisturbed cortical reprocessing of previously replayed memories.

These considerations highlight the contribution of high spatio-temporal resolution data from invasive electrophysiology as a way to provide context for studies using coarser monitoring techniques \(^{64}\). Based on such data we suggest that spindles are important for memory consolidation, but are not global and are not the main driver of active memory consolidation.

**Sleep stages, neural dynamics and memory consolidation**

Our hypothesis, that SO (K-complexes)/SWR/replay support active memory consolidation and that spindles are important for local plasticity, is in agreement with human studies associating memory performance with sleep stages. Furthermore, this suggests an important role for LS.

Replay most likely occurs throughout all of NREM; however, LS dynamics seem to be more favourable for global information exchange. Indeed, SWR density is higher in LS than
SWS\textsuperscript{24} and SWR increase after learning does not correlate with delta waves, which suggests that it is not bound to SWS \textsuperscript{67}. In addition, sleep spindles and K-complexes/SO occur mostly during light sleep \textsuperscript{17,68-70}. Finally, a dominant role for LS is also compatible with analyses of correlations between spindles and memory maintenance \textsuperscript{71,72}. These are in fact strongest in LS, \textsuperscript{53,71-73} because spindles tend to co-occur with K-complexes in LS and are most likely triggered by them.

Initially, human sleep studies gave a preponderant weight to SWS and REM over LS for memory processes, despite LS comprising more than 50\% of human sleep. Due to the widespread usage of the half-night paradigm, for a long time procedural memory was linked to REM sleep, dominating the second half of the night, while declarative content was thought to be dependent on SWS, the most prevalent stage in the first half \textsuperscript{4,74,75}. However, studies using the half-night paradigm should be viewed with caution; while the first and second half of the night are dominated by SWS and REM sleep respectively, all sleep stages, and related dynamical processes, do occur in both night-halves. Further, the two night-halves dramatically differ in hormone levels (e.g. growth hormone and cortisol) as well as the actual testing time: learning in the evening vs. in the middle of the night and retest in the middle of the night vs. rested in the morning after a full night of sleep.

Based on studies using this half-night paradigm, theorists assumed that both replay and downscaling were tied to SWS\textsuperscript{1,10}. Presumably this persisted due to the fact that classically in animal sleep research all NREM is called SWS and both phenomena can only be directly observed with invasive techniques in animals. However, recent studies involving selective deprivation of either REM or SWS (during night sleep or daytime naps), with little or no effect on memory have contested this conceptualization \textsuperscript{53,71,76,77}, leading to the hypothesis that LS could be the more relevant stage. Similarly, significant memory consolidation can occur during a short nap even though only roughly half of the subjects normally achieve deep sleep in such short intervals \textsuperscript{53,77-79}. In one study, 6 minutes of LS appeared to be sufficient for consolidation \textsuperscript{80}. Additionally, fusiform-medial prefrontal connectivity during LS after learning correlated positively with overnight memory retention for arbitrary face-location associations. \textsuperscript{81}

Interestingly, GABA\textsubscript{A} agonists like benzodiazepines promote LS, partly at the cost of SWS \textsuperscript{82}. According to the prevalent SWS-focused model they should therefore impair memory consolidation. However, while GABA\textsubscript{A} agonists taken before the acquisition phase impair learning and result in anterograde memory deficits, they produce retrograde memory facilitation if taken after acquisition (for a review see \textsuperscript{83}). For example patients receiving GABA\textsubscript{A} agonists as sleep medication after learning experience significantly more S2 and better sleep-related memory consolidation compared to patients without GABA\textsubscript{A} agonists \textsuperscript{84}.  

In sum, we propose that active memory consolidation likely occurs in all stages of NREM, but dominates during LS when conditions for global brain interactions are most optimal. In the next section we will discuss how the distinction between local and global activations in NREM sleep suggests a hypothesis for the reconciliation of the “active role” and “homeostatic downscaling” theories of sleep and memory.

**Delta wave activity for local cortical homeostasis**

Having identified that SO (K-complexes), SWR, and replay are critical for active memory consolidation and dominate during LS, we further hypothesize that delta waves support homeostatic processes and dominate deep sleep. The synaptic homeostasis hypothesis states that synapses that have become potentiated during a wake period are downscaled during sleep to a baseline level that is energetically sustainable and allows efficient use of grey matter space. Downscaling would improve memory retention by curtailing weak synaptic connections while sparing stronger ones, thus improving the signal to noise ratio and freeing up capacity for new learning. Evidence has linked downscaling to cortical delta waves that arise from the potentiated state of the cortical network. The detection of delta waves in surface EEG would lead to a classification of SWS, linking deep sleep to downscaling. Delta waves create optimal conditions for synaptic depotentiation due to the low tone of neuromodulatory systems such as noradrenaline and BDNF that normally support synaptic potentiation. This contrasts with data showing synaptic potentiation arising from K-complexes during NREM sleep. Interestingly, sequences of UP-DOWN-UP states were found to trigger synaptic potentiation when presented in a random, irregular fashion as they appear in LS, but synaptic depression when expressed more periodically which is typical of SWS. Furthermore, delta waves seem to arise, at least partly, due to homeostatic pressure, in response to potentiation that accumulates during wake and reduces during sleep. Indeed, the slope of cortical evoked potentials that mark synaptic strength are steeper after a wake period than after sleep and this change is associated with the amount of delta waves, i.e. slow wave activity (SWA). Such pressure dissipates with sleep, which is reflected by decreased SWA in later phases of sleep. The final outcome of this downscaling process is reflected by extracellular glutamate levels and neural spike rates, which are increased after wakefulness but decrease during NREM sleep associated with SWA. Despite such supportive evidence for the synaptic downscaling hypothesis, a critical note is in place. First, homeostatic plasticity during sleep may not only involve synaptic weakening, but also synaptic strengthening. Second, neuromodulators have diverse effects on synaptic plasticity, thus low neuromodulator levels does not necessitate synaptic weakening. Finally, important changes in network connectivity accumulating during the day, in particular in the medial temporal lobe.
and the connections between this area and other structures involved in memory, are not accompanied by increased brain metabolism not even in the regions affected by connectivity changes\textsuperscript{93}. Regardless of these questions, synaptic homeostatic processes during sleep involve general changes in synaptic plasticity unlike synapse specific plasticity evoked by active memory consolidation due to replay.

While standard systems memory consolidation theory predicts that an ‘active’ role of sleep would be useful to build whole-brain synaptic representation, which requires global interactions between cortical areas, downscaling takes place at the level of single neurons and synapses, and so could be supported by more local phenomena. Concordantly, most delta waves are local\textsuperscript{17} and are locally regulated as a function of prior learning experiences and plasticity, which, depending on the nature of learning, may recruit different cortical areas\textsuperscript{5,94,95}. fMRI connectivity analysis shows a breakdown of cortico-cortical connectivity and a more local network with local clustering during SWS\textsuperscript{13}. Adding to this argument is the finding that after prolonged waking, neurons can go quiescent locally accompanied by SWA in local EEG during wakefulness. Hence, neurons can go offline in one region but not another while animals are active and scalp EEG suggests an awake state\textsuperscript{96}. Once again this example illustrates that surface EEG does not accurately reflect all underlying physiological processes.

These data led to the hypothesis that downscaling is beneficial for memory retention. For example, local increases in SWA over task-related brain regions correlate with memory improvement over sleep in humans\textsuperscript{5}. However, local SWA may only be a consequence of increased plasticity due to new learning and replay. In a complementary hypothesis, downscaling may affect memory by renormalizing synaptic weights, thus preparing a “clean slate” for new encoding. Compatible with this view, deep sleep suppression has been reported to affect subsequent encoding-related hippocampal activation and memory performance during encoding in humans\textsuperscript{97-99}. Additionally, the amount of SWS, reflecting a dominance of delta waves, in humans correlates positively with subsequent memory performance and negatively with hippocampus activation at recall\textsuperscript{100}. Further, SWS may be responsible for restoring prefrontal memory control functions\textsuperscript{101}. The prefrontal cortex is thought to exert executive control over memory retrieval, e.g. via the inferior temporal cortex to retrieve stored memories\textsuperscript{101,102}. At the same time cognitive impairments that arise from sleep deprivation are often found in executive control tasks as the Go/NoGO task, task switching, and inhibitory control tasks\textsuperscript{101}. Further, neuroimaging studies have provided evidence to suggest that sleep deprivation affects executive control dependent on the frontal lobes\textsuperscript{101}.

In conclusion, local downscaling can improve memory retention and free up capacity for new learning, further it is supported by local delta activity that occurs predominantly
during SWS. Our hypotheses imply specific functional roles for active memory consolidation and downscaling, a discussion of which we turn to below.

*The functional role of replay and downscaling: Creating and updating memory networks*

Our hypotheses suggest a possible scenario that reconciles the two main theories for sleep and memory: global, transient fluctuations of activity, and the concomitant replay of memory traces, connecting hippocampus and neocortex, would be important for active systems consolidation and dominate during light NREM sleep. After these processes take place in the first part of each NREM/REM cycle, more local oscillations, at hyperpolarization levels favouring synaptic depression, engender downscaling and improve the signal to noise ratio of the surviving memories and this occurs mainly during SWS. Most likely, replay would continue during SWS, counteracting the downscaling process for the salient memory traces, which should not be erased.

It remains to be seen how replay and downscaling may act synergistically to shape memories during sleep (see Fig.2). Neocortical memory traces are thought to be organised in memory networks. Sleep has been proposed to play a role in memory network construction, possibly via hippocampal reactivation of episodic memories, propagated to cortex and there reprocessed to extract their statistical overlap. In turn, these memory networks may serve as schemas that come to guide consolidation of new memories. The push-pull action of replay (potentiating ‘important’ traces, for example those fitting the current cortical networks) and downscaling (weakening irrelevant traces) may support such network updating (for a model see Fig.2). Indeed, in humans the shift from HPC to neocortical retrieval networks occurs during sleep, and in rats schema updating takes place in a 3-48 hour time window, hinting at a role for sleep as well.

Determining the neural dynamics that may support such a complex process, fundamental for cognition, will require making full use of data from all sources, and to map behavioural and EEG data onto detailed physiology, a path that the research community has just begun to follow. New technologies may greatly facilitate progress in this area (see ‘Future experiments’ box), and so can careful study of psychiatric patients and their corresponding animal models, showing both disrupted sleep and memory consolidation (see box one). A better understanding of the distinct contribution of sleep phenomena will contribute to this progress.

*Figure 2 please near here*
Acknowledgement:
L. Genzel is funded by ERC-2010-Ad6-268800-Neuroschema grant.
Open questions box:
Future experiments

- **Determining the precise role of sleep oscillations in memory:** The approach of Girardeau et al, and Ego-Stengel et al.\textsuperscript{27,28} can be extended, making use e.g. of optogenetic stimulation, to selectively suppress or enhance the dynamical components of sleep (SWR, spindles, delta waves, slow oscillations, etc.) and test their effect on memory. Importantly, this should be done for recent vs. remote memories with and without previous knowledge/schema (which according to the systems memory consolidation hypothesis should be supported by different neural dynamics), as well as for memories with different content.

- **Long-term ensemble recordings** to understand whether replay takes place (and in which forms) during SWS, and to study the effects of synaptic downscaling at the neural ensemble level.

- **More high-resolution techniques in human sleep research:** Combining different techniques (high-resolution EEG (intra and extra-cranial), fMRI, magnetoencephalography (MEG)) to record memory encoding and replay during sleep, thus trying to replicate findings in humans with the high spatial and temporal resolution of animal research.

- **Investigating the role of REM:** Clarifying the emotional and fear component in animal memory tasks to better understand the role of REM sleep in memory consolidation, which is currently thought to be important for very emotional memory content.

- **Implement more sophisticated analysis** (e.g. regression or factor analysis, machine learning/classifiers) of human behavioural, fMRI and EEG data instead of focussing on correlations.

- Clarify **behavioural tasks** in animals and humans regarding their relevant processing requirements, brain areas, and previous knowledge (schema effect) to assess possible sleep related processes. Sleep may not always increase behaviour performance, but instead affect use of strategy/brain area (see \textsuperscript{110,114}).
Disturbed sleep is a key symptom of many psychiatric diseases\textsuperscript{115}. In major depression, a robust finding is a decrease in SWS and a disinhibition of REM sleep \textsuperscript{116,117}. The fact that many antidepressants suppress REM sleep seemingly without major detrimental cognitive effects has been proposed as a counterargument of claims that REM sleep is involved in memory consolidation in general \textsuperscript{118,119}. According to the classical two-process model of sleep-related memory consolidation, decreased SWS and suppressed REM sleep in medicated depression should impair declarative and procedural memory consolidation, respectively. Although medicated patients with depression experience a strong decrease in their procedural memory consolidation overnight \textsuperscript{120,121}, this consolidation impairment turned out to be unrelated to REM sleep suppression \textsuperscript{84,120,122}. Surprisingly, overnight consolidation of a declarative task was demonstrated to be unaffected in depression: despite experiencing 25\% less SWS than controls, depressed patients showed nominally even better declarative memory consolidation than controls\textsuperscript{84}. The relationship between sleep spindles and memory consolidation in depression is currently unclear. Some studies report reduced spindle activity in depressed patients \textsuperscript{123-125}, others did not find any difference\textsuperscript{126,127}, and a recent study even demonstrated increased spindle density in female patients with depression \textsuperscript{128}. In contrast, similar impairments of sleep-related procedural memory consolidation in schizophrenic patients \textsuperscript{121,129} have repeatedly been associated with deficits in several markers of spindle activity \textsuperscript{123,130,131}.

Similarly, whether SWS is involved in consolidation impairments in schizophrenia is still debated. While several studies did not find SWS changes in schizophrenia \textsuperscript{131,132}, others reported decreased SWS to be related to impaired memory consolidation in schizophrenic patients \textsuperscript{129,133}. As a possible neural mechanism underlying impaired memory consolidation in schizophrenic patients, recent electrophysiological research using an animal model of schizophrenia suggests that fragmented NREM sleep and impaired slow-wave propagation culminates in deficient ripple-spindle coordination and disrupted spike timing due to deficits in hippocampal-cortical connectivity, which would indicate that spindles are more of an “index” for deficits in connectivity \textsuperscript{134}. In conclusion, whilst psychiatric disorders are often accompanied by sleep disturbances and cognitive deficits their relationship remains elusive.
Box Two: Some consequences for the analysis of human EEG

The important distinctions we made in the previous section between sleep stages have important consequences for the practitioners of human EEG. In particular, care must be taken in EEG analysis to avoid confusion between SO (fluctuations between UP and DOWN states at a frequency of 0.5-1 Hz) and delta waves. These two phenomena have different origins; SOs, as described, are generated within the cortical networks. An isolated UP/DOWN/UP cycle is known as a K-complex and is visible in the surface EEG with a frequency of <1Hz and amplitude of >140µV\(^{135}\). Delta waves, in contrast, originate from the interaction between pace-making neurons in the thalamus and the neocortex, and manifest themselves in surface EEG as 1-4Hz, 75-140µV\(^{135}\) amplitude oscillations. SOs and delta waves show a different homeostatic regulation; through the night delta waves decrease in strength and the periods during which delta waves dominate (SWS) also decrease while K-complexes and LS amounts remain stable\(^{14,17,22,136}\). Importantly for their possible function, delta waves tend to be more localized in restricted cortical regions, while K-complexes are more global – making them more suitable to orchestrate systems consolidation \(^{17}\). Moreover, K-complexes and not delta waves are most likely important for replay, as only the former are associated with SWR density implicated to be critical to the replay-complex \(^{67}\) and only K-complexes are positively correlated with hippocampal replay \(^{30,47}\).

In general, the surface EEG should be viewed in light of the dominating processes creating the electrical phenomena measured. All of NREM is a continuum with only artificial sharp boundaries. Replay occurs throughout all NREM whereas downscaling becomes increasingly dominant as sleep deepens. As long as replay dominates, LS will be seen in the EEG, when downscaling prevails we classify the sleep EEG as SWS in humans.
**Glossary Box:**

**REM:** Rapid Eye Movement sleep (REM) makes up roughly 25% of all human sleep and is characterized by a paradoxical and wake-like EEG and muscle atonia.

**NREM:** Roughly 75% of human sleep is non-REM sleep (NREM), which is further subdivided by sleep “depth” into S1-4.

**S1-4:** sleep stages 1-4 of NREM

**LS:** Light sleep (LS) consisting of sleep stages 1 and 2.

**SWS/deep sleep:** Slow wave sleep also known as deep sleep is made up from the sleep stages 3 and 4. Many animal sleep researchers use the term SWS to describe all NREM.

**SWR:** Sharp wave ripples. Transient excitatory burst originating in the hippocampus, and associated in hippocampal subfield CA1 to 200 Hz “ripple” oscillations

**Spindles:** Sleep spindles are waxing and waning oscillations at 12-16 Hz (humans)/9-15 Hz (rodents) seen in the surface EEG and characteristic for S2 LS.

**SO:** slow UP/DOWN state transition with a frequency of 0.5-1 Hz, seen throughout NREM sleep

**K-Complexes:** K-complexes are large waves (<1Hz and >140µV) visible in the surface EEG and defining for S2 LS.

**Delta Wave:** oscillations classically defined at 2-4 Hz (animals)/1-4 Hz (human), usually nested within SOs

**SWA:** Slow wave activity. Active-silent phases at 0.1-1 Hz; when periodic identical to delta waves.

**UP/DOWN states:** states of, respectively, elevated and reduced cortical activity, whose alternation gives rise to SO.

**Downscaling:** A theory concerning global down-regulation of synaptic strength during SWS via SWA proposed by Tononi and Cirelli (2006).

**Replay-complex:** The replay-complex consists of SO, SWR and spindle and has been linked to replay of memory traces in the hippocampus, striatum and cortex during sleep.
Fig 1:

**Figure 1: Sleep stages**

Mammalian and avian sleep consists of two very distinct types of sleep: REM and nonREM (NREM) sleep. In all species NREM and REM alternate in a cyclic fashion with NREM always preceding REM. As seen in panel b, in humans NREM generally is further subdivided into sleep stages 1-4 (S1-4) with S1 and S2 known as light sleep (LS), and S3 and S4 – characterized by large delta waves – as deep or slow wave sleep (SWS). On the left exemplary EEG traces are shown for each sleep stage. On the right a representative sleep hypnogram is presented with a typical sleep stage distribution over an 8h-nightime sleep episode. Early in the night the first NREM-REM cycles are dominated by SWS, while later cycles contain more REM sleep. Note, while SWS and REM sleep have classically drawn most of the attention in sleep research, human sleep is actually dominated by S2 light sleep, which comprises more than 50% of the sleep period. As seen in panel a on the right, rats and mice, nocturnal animals sleeping during the light hours of the day, show polyphasic sleep with many very short NREM-REM cycles interrupted by phases of wake. On the left exemplary EEG traces are depicted for the different stages. In animal sleep research it is customary, even though it is possible, to make no distinction between different stages of NREM sleep lumping them together under the SWS label, which has led to a fundamental mismatch between human and animal data. It would be extremely beneficial for this research field, if uniform and more precisely defined distinctions between sleep stages were adopted across species. Mouse hypnogramm courtesy of V. Jakubcakova and M. Kimura, Max Planck Institute of Psychiatry.
**Figure 2: The functional role of replay and downscaling:**

Top: During encoding different sets of episodic memory traces with information coded in different cortical modules (smell, vision etc.) are connected via the hippocampus (red lines).

**Creating and Updating Memory Networks:** (A) The replay of these episodic memory traces (see D: ‘spiking activity’) is initiated by the cortex via slow oscillations (SO, purple arrow), executed during SWR, and propagates from the hippocampus (blue arrows) and back with global plasticity. The hippocampus drives replay, plasticity, and active memory reorganization in the cortex. (B) During sleep spindles (seen in the surface EEG) the cortex is deafferented from the hippocampus (blue line) perhaps enabling local cortical processing via Ca\(^{2+}\) influx. Through repeated replay (over many nights after repeated learning) the overlapping cortical modules of the episodic traces form connections with each other creating a memory network (red lines between cortical modules in A and B). (C) During subsequent downscaling, local plasticity (blue arrows) cause weaker connections to disappear (dashed grey lines) leaving only the overlapping network (orange lines), and decreasing the noise level and preparing the hippocampus for future encoding. (D) The electrical signals linked to these phenomena (SO, K Complexes, SWR, spindles, SWA/delta waves) can be visualized in depth and surface EEG recordings. Which sleep stage is classified from the surface EEG.
depends on the dominant process at the time (Bottom two panels). When global events dominate, which we hypothesize to allow replay and active systems memory consolidation, S2 light sleep is defined. In contrast, when local network dynamics dominate, such as delta waves, which we hypothesize to accompany downscaling, deep sleep/SWS is defined. Note that replay occurs through all stages of NREM but predominates during LS, whilst downscaling dominates over replay during SWS. SO = slow oscillations, SWR = sharp wave ripple, SWA = slow wave activity, SWS = slow wave sleep
Reference List


7 Karni,A. et al. (1994) Dependence on REM sleep of overnight improvement of perceptual skill. *Science* 265, 679-682


17 Nir,Y. et al. (2011) Regional slow waves and spindles in human sleep. *Neuron* 70, 153-169


34 Kudrimoti,H.S. et al. (1999) Reactivation of hippocampal cell assemblies: effects of behavioral state, experience, and EEG dynamics. The Journal of Neuroscience 19, 4090-4101


39 Diekelmann,S. et al. (2012) Offline consolidation of memory varies with time in slow wave sleep and can be accelerated by cuing memory reactivations. Neurobiol. Learn. Mem. 98, 103-111

40 Rudoy,J. et al. (2009) Strengthening individual memories by reactivating them during sleep. Science 326, 1079


81 van Dongen, E.V. *et al.* (2011) Functional connectivity during light sleep is correlated with memory performance for face-location associations. *NeuroImage* 57, 262-270


88 Vyazovskiy, V.V. et al. (2008) Molecular and electrophysiological evidence for net synaptic potentiation in wake and depression in sleep. *Nat Neurosci* 11, 200-208

89 Vyazovskiy, V.V. et al. (2009) Trigerring Slow Waves During NREM Sleep in the Rat by Intracortical Electrical Stimulation: Effects of Sleep/Wake History and Background Activity. *Journal of Neurophysiology* 101, 1921-1931


91 Frank, M.G. (2012) Erasing synapses in sleep: is it time to be SHY? *Neural Plasticity* 264378


95 Huber, R. et al. (2006) Arm immobilization causes cortical plastic changes and locally decreases sleep slow wave activity. *Nat Neurosci* 9, 1169-1176


99 Van Der Werf, Y.D. et al. (2009) Sleep benefits subsequent hippocampal functioning. *Nat Neurosci* 12, 122-123


101 Wilckens, K. et al. (2012) Age-related decline in controlled retrieval: The role of the PFC and sleep. *Neural Plasticity* 2012


104 Tamminen, J. et al. (2010) Sleep spindle activity is associated with the integration of new memories and existing knowledge. J. Neurosci. 30, 14356-14360


126 Goetz, R. et al. (1980) Spindle characteristics in prepubertal major depressives during an episode and after sustained recovery: a controlled study. *Sleep* 6, 369-375


128 Plante, D.T. et al. Topographic and sex-related differences in sleep spindles in major depressive disorder: A high-density EEG investigation. *Journal of Affective Disorders*


