Opinion

Cell PRESS

How schema and novelty augment memory formation

Marlieke T.R. van Kesteren[1,2], Dirk J. Ruiter[2], Guillén Fernández[1,3*] and Richard N. Henson[4*]

[1] Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, P.O. Box 9101, 6500 HB, Nijmegen, The Netherlands
[2] Department of Anatomy, Radboud University Nijmegen Medical Centre, P.O. Box 9101, 6500 HB, Nijmegen, The Netherlands
[3] Department of Cognitive Neuroscience, Radboud University Nijmegen Medical Centre, P.O. Box 9101, 6500 HB, Nijmegen, The Netherlands
[4] MRC Cognition & Brain Sciences Unit, Cambridge, CB2 7EF, UK

Information that is congruent with existing knowledge (a schema) is usually better remembered than less congruent information. Only recently, however, has the role of schemas in memory been studied from a systems neuroscience perspective. Moreover, incongruent (novel) information is also sometimes better remembered. Here, we review lesion and neuroimaging findings in animals and humans that relate to this apparent paradoxical relationship between schema and novelty. In addition, we sketch a framework relating key brain regions in medial temporal lobe (MTL) and medial prefrontal cortex (mPFC) during encoding, consolidation and retrieval of information as a function of its congruency with existing information represented in neocortex. An important aspect of this framework is the efficiency of learning enabled by congruency-dependent MTL–mPFC interactions.

Introduction

The existence of prior knowledge to which new information can be related generally improves memory for that information. Although the role of such schemas in learning has long been studied in psychology (Box 1), this role has only recently been studied in neuroscience [1,2]. In particular, whereas structures within the MTL, such as the hippocampus, have long been implicated in the learning of declarative information [3], recent neuroscientific data have implicated an additional, time-dependent involvement of the mPFC [4,5], particularly when new information is congruent with a schema [6–8].

A second line of research has studied how the novelty of information can also improve its retention (Box 2). This raises the question of when information conforming to a schema (congruent information) is remembered better or worse than information that does not (unrelated, or incongruent, information) [1,9], a question that has important implications for optimising learning in educational settings [10]. Below, we review recent neuroscientific research addressing this question before presenting a new framework that tries to explain the complex relationship between schema, novelty and memory.

Review of schema in systems neuroscience of memory

Several theories exist about how new information becomes consolidated into memory [1,11–13]. The so-called standard systems-level theory of consolidation [14] proposes

Glossary

Declarative memory: memories that can be declared, that is, have a propositional truth value (events or facts), normally associated with conscious recall as distinct from procedural (non-declarative) memories such as skill-learning, which cannot be verbalised and are often expressed unconsciously.
Episodic or instance memory: declarative memory for a specific event in space and time, which normally includes other contextual information present at that time (e.g. internal thoughts and states). We use instance to refer to a specific pattern of neocortical activity that is bound to an index in the MTL according to our SLIMM framework; we use episodic more generally to refer to memories with contextual information, which is often incidental (i.e. non-recurring, not part of an existing schema).
mPFC (medial prefrontal cortex): medial aspect in the prefrontal cortex, encompassing Brodmann areas (BA) 10, 11 and 32 in humans, and prelimbic, infralimbic and anterior cingulate cortex in rodents.
MTL (medial temporal lobe): part of the brain comprising hippocampus, perirhinal and enthorhinal cortices and parahippocampal gyrus.
Neocortex: association cortex that stores elements of a memory trace (visual, spatial, auditory, somatosensory, emotional, etc.). Note that the mPFC is part of the neocortex anatomically, but not considered to represent memory elements in the present framework.
Novelty: response to information that is not expected or predicted in a given context on the basis of prior experience. Note that we distinguish here between two types of novelty (Box 2): unrelated information that does not strongly match any schema, and incongruent information that is inconsistent with a dominant schema. Within the present SLIMM framework, only the latter improves memory, and note that this type of novelty cannot exist without a schema (i.e. the two concepts are intimately related).
Reactivation: reinstatement of a memory trace, either by online re-encountering of similar information or by replaying the memory trace during offline periods.
Resonance: neural state of co-activity of multiple mental representations (possibly across multiple brain regions), most probably bound via coherent (synchronous) activity.
Schema: network of neocortical representations that are strongly interconnected and that can affect online and offline information processing.
Semantic or schematic memory: general, factual declarative memory that captures regularities extracted from multiple encounters (instances) over time, and divorced from accompanying episodic details. We use schematic to refer to a (resonating) pattern of activity produced by strong connections within neocortex (i.e. an activated schema) within our SLIMM framework; we use semantic more generally to refer to acontextual knowledge that people possess.
Systems consolidation: time-dependent and offline process by which connections between elements of a memory trace in the neocortex are strengthened so they are retained over the long term, independently of MTL structures such as the hippocampus.

Corresponding author: van Kesteren, M.T.R. (marlieke.vankesteren@donders.ru.nl)
Keywords: schema; novelty; memory formation; medial temporal lobe; medial prefrontal cortex; prediction error.
* These authors made equal contributions.

0166-2236/$ – see front matter doi:10.1016/j.tins.2012.02.001

Box 1. History of schema research

The term schema was introduced from the philosophical work of Kant to developmental and cognitive psychology during the early 20th century by Piaget and Bartlett, respectively [83,84], and refers loosely to an abstract, structured mental representation. This concept led to a cascade of both empirical behavioural research [85,86] and theoretical developments in artificial intelligence and connectionist modelling [68,77,87]; it also influenced educational theory [88].

A primary focus of behavioural research was how schemas aid the retrieval of complex information by providing a scaffold for organising retrieval of that information. This reconstructive aspect of memory offered a natural explanation of biases and false memories that occur from an over-reliance on schema [83]. Importantly however, schemas may also affect the encoding and consolidation of memories ([68]; see the main text). For example, the superior recall of schema-congruent information cannot always be explained by facilitated retrieval [46] (e.g. by generation of schema-related information at test, followed by episodic recognition of information present at study [89]).

A primary focus of connectionist modelling was the extraction of regularities from exposure to new information (instances) during learning [90]. A core problem here is the stability–plasticity dilemma, which is the degree to which a new instance should alter existing knowledge about a class of instances (schema) without destabilizing such knowledge. One solution to this problem (adopted in adaptive resonance theory [69]) was a global parameter (vigilance) that determined whether or not a new instance needs to be represented separately as a function of its similarity to existing schemas. Another solution was to draw on different learning rules in complementary learning systems, in particular a fast-learning system (in MTL) that stores unique instances, which can then be replayed in an interleaved fashion to a slower-learning system (in neocortex) that extracts their commonalities [30,77].

Enthusiasm for schema research waned since the 1980s, partly because of the overextended definition of schema that arose from the explosion of interest and partly because of some apparently contradictory behavioural results, where novel information (that does not conform to a schema) can sometimes be remembered well (Box 2). Nonetheless, there has been a recent revival of interest in schema within the neuroscience community [1,6–8,29]. Here, the concept of a schema is simpler than in previous psychological research, operationalised, for example, as a familiar spatial layout such as the relationship between a number of locations within an arena in which a rat expects to find food [29], or as whether a word that must be associated with a novel visual stimulus is congruent with a simultaneously presented tactile stimulus [7]. Our present (neuroscientific) concept of a schema therefore refers simply to a network of neocortical representations that are strongly interconnected, activation of which affects processing of new information, as expanded in the main text.

that new (declarative) information is initially dependent on MTL structures such as the hippocampus, but over time (possibly through reactivation [15], e.g. during sleep [16,17]) this information becomes relatively more dependent on the neocortex. This proposal is based on evidence that MTL lesions not only impair the ability to form new memories (anterograde amnesia) but also impair the ability to retrieve memories formed within a period before the lesion (retrograde amnesia) [18]. After consolidation, long-term memories are believed to be represented by networks of interconnected neocortical brain regions representing the constituents of those memories, retrieval of which has become independent of the MTL (although see below and [19]). More recently, an additional role has been suggested for the mPFC in such consolidation [20–22], consistent with evidence of offline replay of learning-related brain activity in mPFC (as well as MTL) [23–25] and by its prominent anatomical location within memory-related brain networks [26–28].

It has been suggested that the presence of a schema, in terms of a pre-existing network of interconnected neocortical representations (see Glossary), accelerates consolidation [2]. For example, a lesion study in rodents showed that memories congruent with a pre-learned spatial schema (Figure 1a) became hippocampally independent after only 48 h (Figure 1b) [29], whereas memories that lacked a prior schema were still hippocampally dependent. In addition, functional imaging in humans during a period of rest shortly after encoding revealed decreased hippocampal–mPFC functional coupling for more versus less congruent information (Figure 1d) [6], whereas successful retrieval of congruent information was associated with increased functional coupling between mPFC and a neocortical region coding that information (Figure 1e) [7]. A schema thus appears to act as a catalyst for consolidation, affecting interactions between mPFC, MTL and other neocortical regions, and possibly increasing the likelihood or effectiveness of replay of congruent information [1,30].

A schema can also influence processes occurring during initial acquisition. For example, functional imaging showed increased activity in mPFC for more versus less congruent information immediately after encoding in rodents (Figure 1c) and increased MTL–mPFC coupling in humans for less congruent information during encoding, related to the strength of the schema (Figure 1d) [6]. These results are consistent with a large body of evidence that MTL–mPFC interactions, along with activity in other brain regions [31–33], are important for successful encoding and retrieval [34–39]. They are also consistent with more general claims that the mPFC is important for making online predictions (e.g. during perception [40,41]) enabled by schemas, whereas the MTL is important for detecting the type of novelty [42–44] associated with an incongruent schema (Box 2).

Although there is much debate about whether patients with MTL damage can form new memories [45], and in many situations they appear to be unable to do so (anterograde amnesia), they can still show a congruency benefit [46] and there are certain situations in which they appear to be able to learn new information [47–52]; these situations are possibly related to the existence of schemas. In particular, recent data have suggested that such patients can learn some information as well as controls can [53–55] – so-called fast (cortical) mapping [54] – which may relate to schemas (see below). Damage to the mPFC, by contrast, has been associated with a reduced ability to filter and integrate incoming information, resulting in confabulation [56], a lack of a congruency benefit [46] and more errors during retrieval [57], which may reflect an inability to utilise schema (Box 1; although see also [58]). These observations, along with lesion data in rodents [59], suggest that memories mediated by MTL and mPFC might be different in nature, ranging from more detailed, episodic memories (instances) supported by MTL, to more general, semantic (schematic) memories integrated by mPFC, as expanded below.

Box 2. Novelty and prediction error

Novelty has long been suspected as an important factor in learning [91]. For example, people are often better able to remember an item that deviated from its prevailing context [92]. Conversely, there would seem little (e.g. metabolic) sense in the brain encoding information that is already fully predicted. For example, there is no need to encode the presence of your BathToy each time you enter your Bathroom (Figure 2a), assuming you always find it there. This is consistent with so-called predictive coding models of memory [81,93], in which the key factor that drives learning is the amount of prediction error (PE). Clearly, schemas still play an essential role, in that predictions are based on such knowledge. This perspective seems to entail greater learning for incongruent than for congruent information, however, in opposition to schema theories (Box 1). However, the precise predictions depend on the nature of the learned information and how it is subsequently retrieved, as expanded below.

From a Bayesian perspective, PE can be viewed as the divergence between prior and likelihood probability distributions. Thus, a familiar location would establish prior probabilities over the objects one expects to encounter there, whereas the (noisy) sensory input would provide the likelihood that certain objects are in fact present (Figure I). If one encounters a novel object in a novel location, such that both the likelihood and prior distributions are imprecise (flat), PE will be low (Figure Ia), at least relative to a familiar object in a novel location (Figure Ib). Thus, maximal overall novelty does not necessarily entail maximal learning; indeed, novel stimuli are often less well associated with unpredictive contexts than are familiar stimuli [82].

Alternatively, when a familiar object (e.g. Cake) occurs unexpectedly in a familiar context (e.g. Bathroom) PE will be high (Figure Ic). This situation corresponds to a maximal match–mismatch [42], where an initial match (recognition of BathRoom) does not match other information (Cake). High PE results in substantial learning, that is, updating of the prior distribution to more closely match the posterior distribution. This can improve subsequent episodic recognition of the object by virtue of reactivating a distinctive context (Bathroom) when that object is repeated [81,82]. However, memory will not always be improved: if cued with the location instead, the overlap between the new predictions (updated prior) and the object representation still may not be sufficient for Cake to be recalled. This contrasts with finding a PlasticDuck in your BathRoom (Figure Id), for which PE will be low (assuming PlasticDucks and BathToys have similar representations), but the updated prior for your Bathroom will overlap with the PlasticDuck representation, allowing it to be recalled. Thus, although incongruent information may produce the greatest PE, the accuracy of subsequent retrieval of that information will depend on how it is cued. This may be one reason why an additional system (e.g. in the MTL) is needed to store incongruent instances, in case they recur and become important for extraction of new schemas (see the main text).

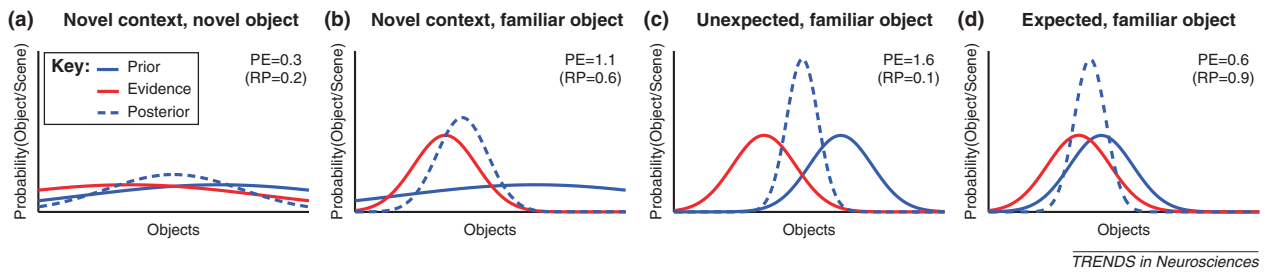

Figure I. Bayesian perspective on prediction-error-driven learningBayesian perspective on prediction-error-driven learningBayesian perspective on prediction-error-driven learning[81]. The curves represent probability distributions, for example over a dimension of objects (ordered by similarity). The red line represents the likelihood of an object being present, given (bottom-up, noisy) sensory evidence; the solid blue line represents the prior distribution, given (top-down) predictions from the current context (e.g. location in the environment); the dotted blue line represents the posterior probability of objects being present (and resembles the updated priors that would result from the learning experience). PE refers to the prediction error – the divergence between prior and likelihood distributions – whereas RP refers to recall prospect – proportional to the posterior probability (from updated priors) of retrieving the object when cueing with the previous context (both PE and RP have arbitrary units). **(a)** A novel object in a novel context, with flat (imprecise) prior and likelihood distributions (akin to the new sequence condition of [42], source memory for unfamiliar proverbs of [82], and the unrelated case in the main text). Although maximally novel overall, PE is relatively low and little can be learned. **(b)** A familiar object in a novel context (akin to the familiar proverbs of [82]), where PE is increased relative to **(a)**. **(c)** A familiar object that is not expected in a familiar context, giving highest PE (akin to the changed condition of [42] and incongruent case in the main text). Because of the residual divergence between posterior and likelihood distributions, however, RP is lower than in **(d)**, which corresponds to a familiar object that is expected in a familiar context (akin to the old sequence condition of [42], and the congruent condition in the main text). This has low PE, but high RP, given high overlap between posterior and likelihood distributions.

Memory for incongruent (novel) information can also be enhanced (Box 2). This novelty advantage has been associated with greater MTL activity during encoding [60–64]. Moreover, the precise type of novelty is likely to be important [42], for example whether information is novel because it is incongruent with an existing schema or because it is unrelated to any existing schemas. Here we focus on enhanced memory owing to the former type of novelty (or prediction error; Box 2), although the latter type of novelty (such as a completely new environment for a rodent [65]) might also improve memory through other means, such as arousal, reward and dopamine release [66,67]. Although the role of novelty has been acknowledged by some schema theorists (e.g. in terms of schemas being used to direct attention to novel aspects of an experience [68]), there is no clear consensus, at least within neuroscientific theories, about the precise conditions under which memory is superior for congruent or incongruent information. Below, we outline a framework termed SLIMM (schema-linked interactions between medial prefrontal and medial temporal regions) that draws on two complementary modes for learning new information, determined by MTL–mPFC interactions, to reconcile the facilitatory effects of schema and novelty on memory.

A new framework relating schema and novelty to memory

SLIMM extends standard consolidation theory, in terms of a time-dependent shift from MTL to neocortical representations, by adding a third component – the mPFC – that acts to accelerate direct neocortical learning independent of the MTL. Within SLIMM, the main function of the mPFC is to detect the congruency of new information with existing information in neocortex, which we term resonance (akin to adaptive resonance theory, ART [69]) in the sense that congruent information resonates with existing information.

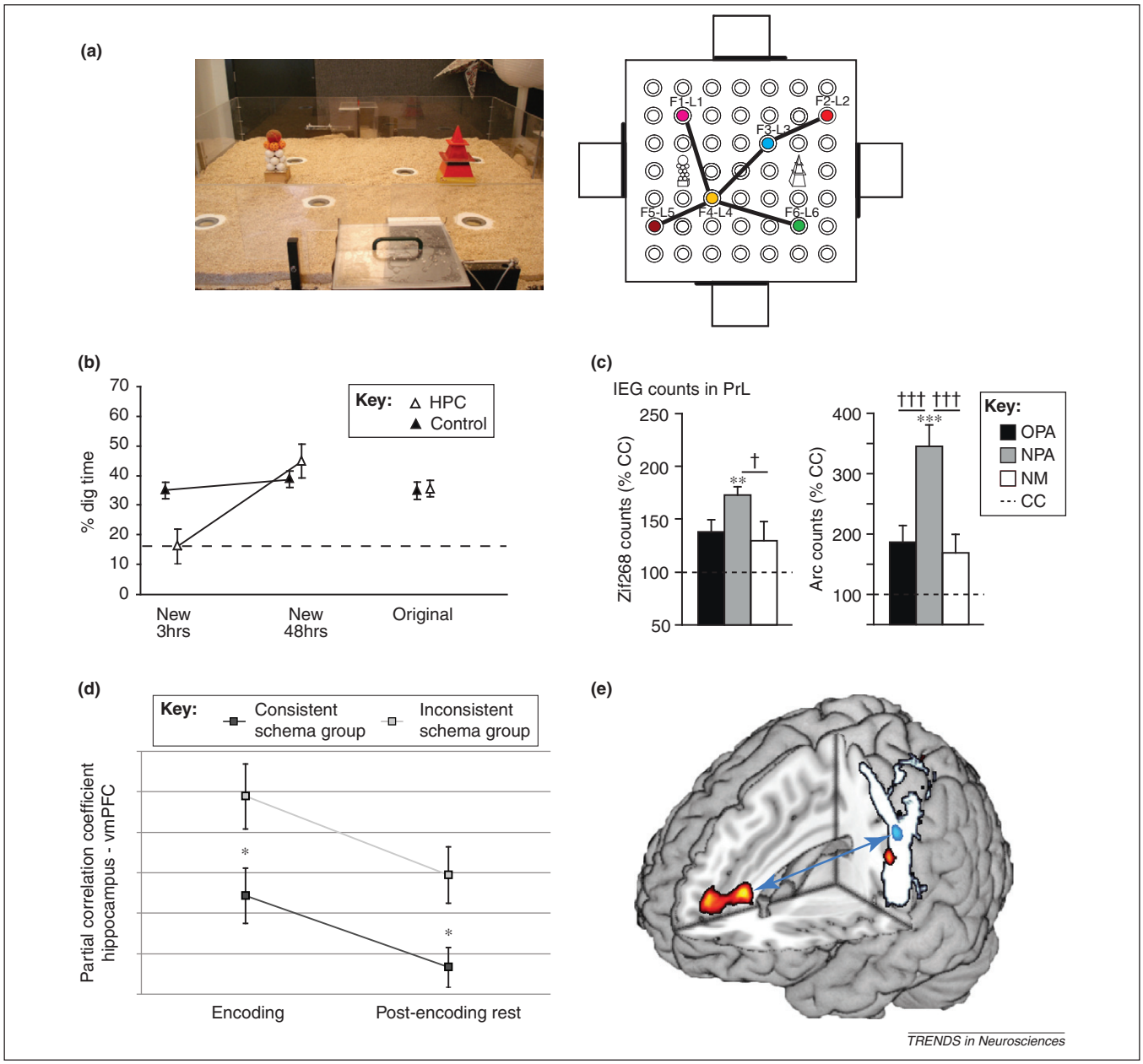

Figure 1. Overview of (a–c) rodent and (d–e) human data on schema and memory. **(a)** Rodent studies have used an event arena in which rodents initially learn a number of flavour–location associations [8,29]. A photo of the area and a schematic of the six locations (L1–6) of the wells and their association with six different flavours (F1–6) are shown. The different landmarks used to navigate in the arena are also shown. **(b)** After learning such a schema, rats showed rapid hippocampal independence (after 48 h, but not after 3 h) of new flavour–location associations within the same arena [29]. The graph shows data for hippocampally lesioned (HPC) versus control animals, represented as percentage dig time in the correct well. A separate group of rodents who had not learned the initial schema did not show such rapid consolidation of the new associations (data not shown). **(c)** In a later study [8], the expression of two immediate early genes (IEGs), zinc finger protein 225 (Zif268, left panel) and activity-regulated cytoskeletal protein (Arc, right panel), was higher immediately after encoding of the new associations (NPA, new paired associates), relative to retrieval of the original associations (OPA, old paired associates), learning of associations in a completely new area (NM, new map) and the performance of caged control (CC) animals, in prelimbic (PrL) structures (equivalent to human mPFC) and in the hippocampus (not shown). **(d)** In humans, mPFC–hippocampal connectivity was greater, both while participants watched a movie (i.e. during the encoding period) and during a resting period shortly afterwards (i.e. post-encoding rest period), the less congruent that movie was with the first part of the movie watched the previous day (i.e. the inconsistent schema group) [6]. **(e)** In a later study [7], mPFC activity and the connectivity between mPFC and a neocortical (somatosensory) region representing the schema-related information were higher during retrieval of information congruent with a schema than of information incongruent with a schema. For an explanation of these different effects of congruency on regional activity and inter-regional connectivity, see Figure 2. Reproduced, with permission, from [29] **(a,b)**, [8] **(c)**, [6] **(d)** and [7] **(e)**.

Greater resonance leads to greater mPFC activity, which in turn is assumed to potentiate direct connections between neocortical representations (e.g. through phase synchronisation [70]). Note that these are the same connections assumed to be more gradually strengthened in the absence of such mPFC input as in standard consolidation theory; the mPFC thus accelerates neocortical learning [54]. Importantly, mPFC is assumed to have a reciprocal relationship with MTL, such that mPFC activity inhibits MTL activity [71,72]. This relates to the assumption that MTL automatically captures new experiences [73] except, according to SLIMM, when inhibition from mPFC means that the new

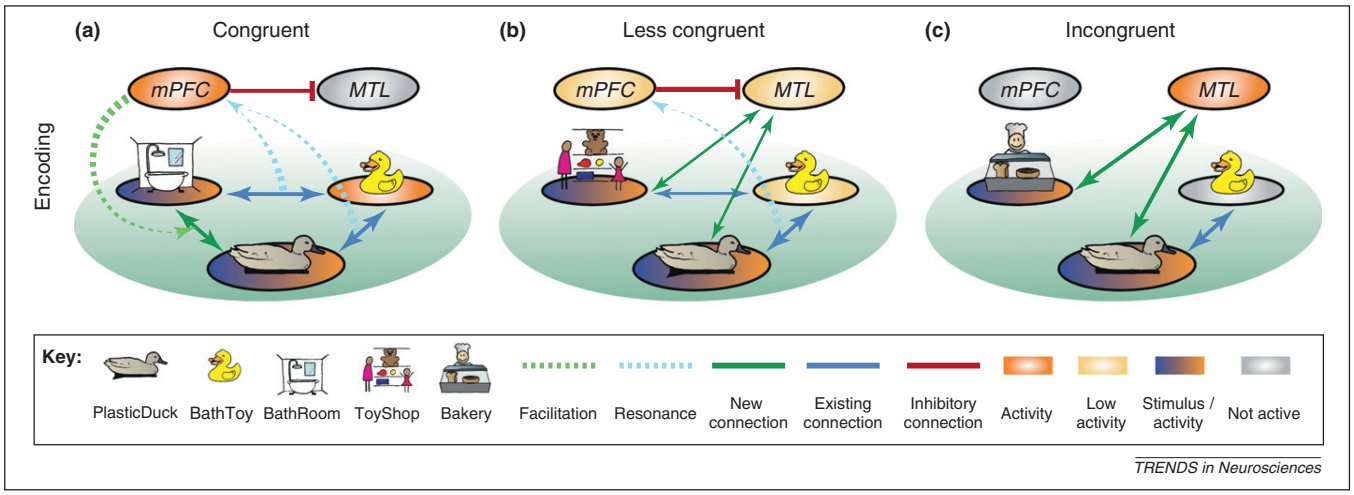

Figure 2. Schematic depiction of the SLIMM model during encoding. Interactions between the mPFC, MTL and neocortex (indicated by the grey-green plane) during encoding of associations between a familiar object (PlasticDuck) and a familiar environment, which is either **(a)** a BathRoom, providing a congruent schema by virtue of a similar BathToy kept there, **(b)** a ToyShop, for which the schema is less congruent, in that a BathToy is only loosely related, or **(c)** a Bakery, in which a BathToy is not part of the schema. In the congruent case, the neocortical representations of PlasticDuck and BathRoom are activated by their perception, and BathToy is activated by its existing associations to both. The mPFC is activated by the resonance (synchronous co-activity) of these representations, and therefore potentiates neocortical connections between all of them, resulting in a new direct connection between PlasticDuck and BathRoom representations (whereas other connections in the schema may have already reached maximum strength). The mPFC additionally inhibits the MTL (indicated by the inhibitory connection). In the incongruent case, the lack of resonance between activated neocortical representations means that mPFC is not activated, MTL is not inhibited and the new association between PlasticDuck and Bakery is stored instead via a separate instance in MTL. In the less congruent case, in which there is only a weak connection between BathToy and ToyShop representations (indicated by a weaker arrow), there is only partial activation of the BathToy representation, and hence partial resonance, and greater inhibition from mPFC to MTL is required to resolve this intermediate state. Hence, memory encoding is less effective.

information can be related via a schema. Only when there is low resonance (or high prediction error; Box 2), as occurs for incongruent information, will the MTL bind those elements into an instance (e.g. via a unique index in hippocampus [74], given its pattern separation capability).

Encoding

As an example, imagine that you encounter a model duck (PlasticDuck) in your bathroom (Figure 2a) that resembles your favourite bath rubber duck (BathToy) but that has not been encountered in your bathroom before. According to SLIMM, your memory for this new (congruent) pairing of PlasticDuck and BathRoom is likely to be good because you already possess an association between BathToy and PlasticDuck and between BathToy and BathRoom (the schema) to which the new PlasticDuck can be related. The simultaneous perception of PlasticDuck and BathRoom activates their corresponding neocortical representations, and this activity spreads to other strongly connected neocortical representations, such as BathToy, owing to previously learned associations. These strong connections mean that the BathToy, BathRoom and PlasticDuck representations resonate (e.g. via synchronous oscillations [75]). This resonance is detected by the mPFC [8,76], which then potentiates the strengthening of neocortical connections between the resonating representations, leading specifically to fast learning of a new, direct connection between BathRoom and PlasticDuck (i.e. good learning). The high activity in mPFC also inhibits activity in MTL [6] such that no indirect association is made between PlasticDuck and BathRoom via a new MTL instance.

Conversely, the same PlasticDuck encountered in a Bakery (Figure 2c) will produce a strong novelty effect (prediction error; Box 2) because such objects are not normally expected there. In this (incongruent) case, SLIMM predicts that you are also likely to remember the pairing of PlasticDuck and Bakery, but for a different reason. The lack of any strong pre-existing connections, direct or indirect, between PlasticDuck and Bakery representations leads to little resonance in the neocortical network. Thus, mPFC is not activated, MTL is not inhibited and the MTL serves to bind the active representations of PlasticDuck and Bakery via a new instance. This leads to good (episodic) encoding [8] that is sensitive to MTL disruption [46,50,54].

Finally, if you encounter PlasticDuck in a ToyShop (Figure 2b), a location assumed to be only loosely related to BathToy (less congruent), neither a specific schema nor a prediction error is likely to be evoked, and memory for that encounter is predicted to be poor. This is because there is weak resonance, requiring increased MTL–mPFC interactions for resolution, as both try to encode the memory [6]. Consequently, neither is strongly activated and there is neither good schematic (mPFC-mediated) nor good instance (MTL-mediated) encoding, leading to poor memory.

Retrieval before consolidation

Imagine walking back into the BathRoom shortly after encoding the congruent case. Activation of the BathRoom representation will lead to processes similar to those at encoding: reactivation of BathToy (schema) and hence PlasticDuck representations, resonance, mPFC activation and further strengthening of the direct neocortical connection between PlasticDuck and BathRoom. Note that concurrent activation of other elements of the schema (e.g. BathToy) can explain the bias towards remembering schematic aspects of PlasticDuck. Similar processes are assumed to happen during replay, when the BathRoom

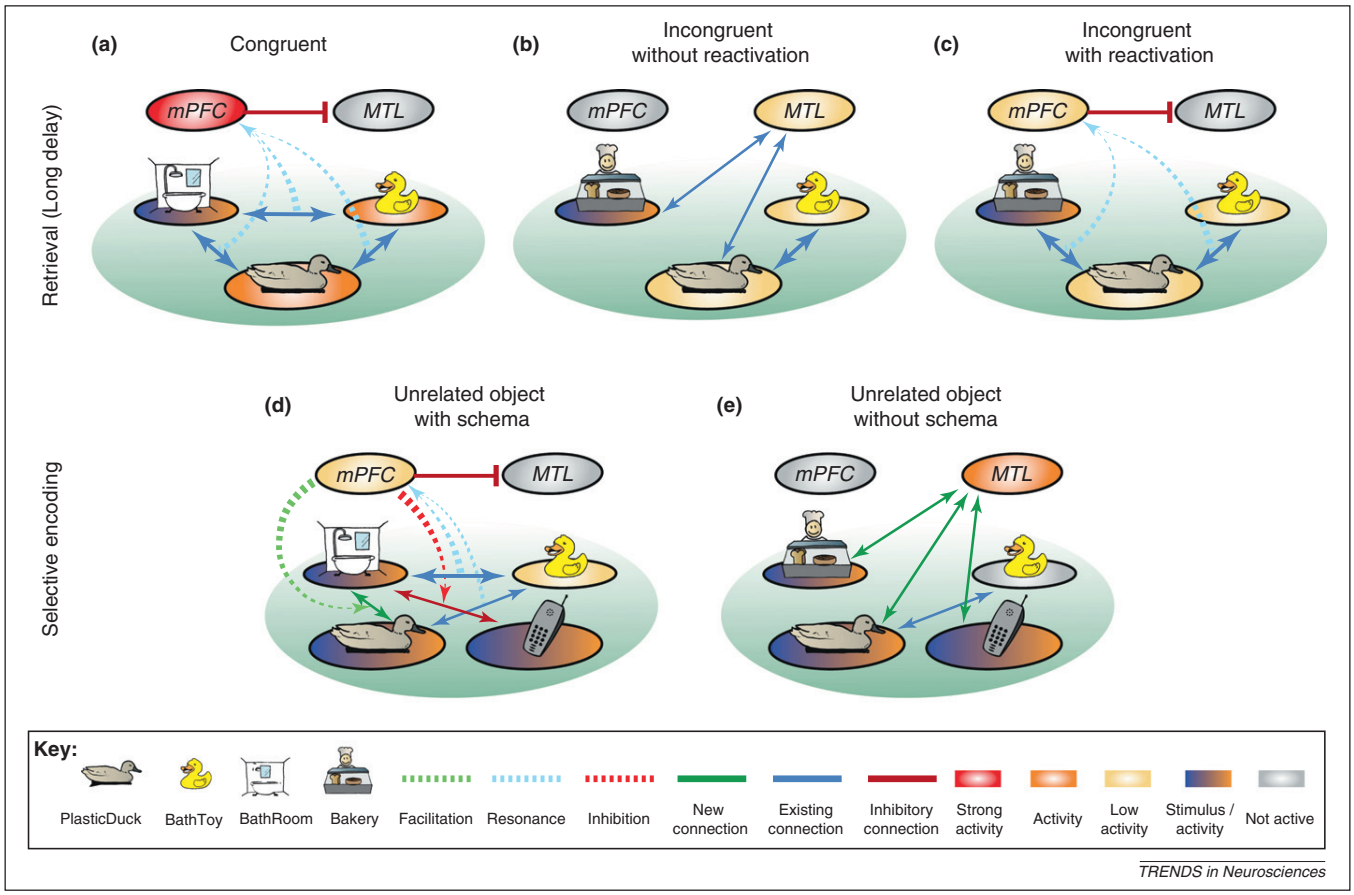

Figure 3. Schematic depiction of the SLIMM model during **(a–c)** memory retrieval and **(d,e)** selective encoding. **(a–c)** mPFC, MTL and neocortical interactions during retrieval of associated objects after consolidation when cued by perception of the familiar BathRoom or Bakery representations from Figure 2. **(a)** In the congruent case, PlasticDuck is likely to be recalled owing to high activity of its representation following activation spread from the BathRoom (and indirectly from the BathToy) representation. **(b)** In one incongruent case, recall of PlasticDuck can occur through retrieval of the MTL instance (episodic recall), although this may be rare (see the text). **(c)** Alternatively, if there have been repeated reactivations of the PlasticDuck and Bakery representations during the delay, for example by their repeated co-occurrence in the environment, recall of PlasticDuck can occur owing to a direct connection from the Bakery representation (i.e. PlasticDuck has now become part of the Bakery schema). **(d)** Associations with incidental, unrelated events (e.g. mobile phone ringing) are suppressed when not consistent with the dominant (e.g. BathRoom) schema in the congruent case, and are hence not well encoded. **(e)** By contrast, in the incongruent case, when all activated representations are bound into the same instance, such associations are encoded by the MTL.

representation is reactivated by internal processes rather than by sensory input [30].

In the incongruent case, walking into the Bakery leads to retrieval of an instance from the MTL, which entails reactivation of not only the PlasticDuck representation but also other incidental (episodic) representations that were present at encoding (see below and Figure 3d). Walking into the ToyShop, by contrast, only leads to weak reactivation of PlasticDuck, given only weak neocortical connections and the low likelihood of the MTL having encoded an instance. Note, however, that if PlasticDuck and ToyShop are repeatedly experienced together, gradual strengthening of neocortical–neocortical connections can eventually lead to effective storage in long-term memory (see the next section), as in standard consolidation theory.

Retrieval after consolidation

After a longer delay, the outcome depends on whether consolidation has occurred, that is, whether there has been repeated reactivation of the crucial representations, either by re-exposure to both or by offline replay. Such reactivation is particularly likely for the congruent case, resulting in the connection between the PlasticDuck and BathRoom representations reaching an asymptote (Figure 3a). In this case, cueing by Bathroom still activates the mPFC through resonance [4,7,8], but because no further neocortical strengthening is needed, the mPFC is not necessary for retrieval.

For the incongruent case, there are two possibilities. If PlasticDuck and Bakery have been repeatedly reactivated (Figure 3c), multiple instances will be encoded by the MTL (in addition to gradual cortical learning). The greater number of such MTL instances also increases the likelihood of offline reactivation of these representations, allowing commonalities across instances to effectively be extracted by gradual learning [30,77]. Thus, eventually the PlasticDuck is no longer incongruent but has become part of the Bakery schema. However, even if PlasticDuck and Bakery have not been reactivated, such that no direct neocortical–neocortical connection exists (Figure 3b), retrieval of PlasticDuck can still occur after a long delay via retrieval of the MTL instance. Although not predicted by standard consolidation theory, this possibility is consistent with other theories and evidence that some remote episodic memories are MTL-dependent [19,28].

Selective encoding and interference

If memories can be retrieved via indices within the MTL system, as in Figure 3c, why is an additional mPFC system needed? Our proposal is that an additional learning system is necessary to overcome the high levels of interference resulting from multiple MTL instances sharing common elements. The function of the mPFC is then to select the most relevant elements of an experience (those congruent with existing schema) during both encoding and retrieval. Thus, the mPFC not only detects resonance but also amplifies activity in congruent representations by additionally suppressing activity in representations inconsistent with the dominant schema (possibly through an attractor-type mechanism [78]). Imagine, for example, that on encountering PlasticDuck in your BathRoom, your mobile phone rings (Figure 3d). Because telephone calls are not particularly related to the BathRoom schema, any connections between the MobilePhone representation and the other active representations are de-potentiated (Figure 3d). In this way, only information that is related to the dominant (active) schema is effectively selected for direct neocortical learning. This automatic highlighting of schema-relevant information is likely to maximise the efficiency of learning of new information [79]. By contrast, when experiencing the PlasticDuck and MobilePhone call in the Bakery, where there is no dominant schema activated (Figure 3e), all of these elements are bound into a single instance by the MTL (i.e. incidental, episodic details, such as the phone call, are better remembered in the incongruent case). This mPFC amplification is also important for reducing interference during retrieval by focusing on representations congruent with existing knowledge. This might explain why patients with mPFC lesions often confabulate, retrieving semantic or episodic information not directly relevant to the retrieval cue [56].

Predictions of SLIMM

SLIMM provides several predictions for future experiments in both healthy subjects and subjects with MTL or mPFC damage. Foremost, it predicts that memory performance in healthy subjects can be a non-linear function of congruency, with better (schematic) memory for congruent items mediated by mPFC and better (instance) memory for incongruent items mediated by MTL. However, because the nature of the memories underlying performance at either end of this congruency dimension differs, the precise shape of this function will depend on the nature of the retrieval test. Free recall or cued recall, for example, may show only an advantage for congruent items (particularly if a generate-and-recognise strategy is used; Box 1), whereas tests of incidental episodic detail (unrelated to a schema), such as recognition or source memory tests, may show an advantage for incongruent items.

For future neuroimaging experiments, the framework predicts that MTL and mPFC will show differential activity patterns and functional coupling (both between each other and with neocortical regions representing components of the memory) as a function of congruency during encoding, offline replay and retrieval. During encoding and replay, mPFC activity is predicted to increase with congruency, MTL activity is predicted to decrease with congruency, and mPFC–MTL coupling is predicted to be maximal for partially congruent conditions when mPFC and MTL are both partially activated (Figure 2b). After consolidation, initially incongruent information will engage mPFC (because it has effectively become incorporated in the schema), whereas successful retrieval of unconsolidated incongruent information will still engage the MTL (Figure 3).

Damage to either the mPFC or the MTL is expected to disrupt the balance between the two types of learning described above. Selective MTL damage is predicted to disrupt episodic encoding and produce complete retrograde amnesia for instances [19], along with temporally graded retrograde amnesia sparing those memories that have already been consolidated [18]. However, the still-functioning mPFC will continue to encode information congruent with prior knowledge (producing congruency effects [46]) via strengthening of neocortical connections between novel information and existing schemas. Conversely, mPFC damage will disrupt schematic encoding of information and hence lead to absence of a congruency effect, because all memories will be stored as instances by the MTL. This will result in difficulties in integrating new information into a schema and increased interference during retrieval of information (confabulation). For information acquired shortly before the mPFC lesion (recent memories), there will still be a congruency effect, because congruent information has been consolidated into neocortical networks in an accelerated manner relative to incongruent information. However, there may be a brief period of retrograde amnesia for highly congruent information acquired very shortly before the mPFC lesion, when no instances were likely to be encoded and consolidation has not yet occurred. For more remote memories already consolidated in neocortex, the mPFC lesion should have no effect (unlike, e.g., damage to anterior, lateral temporal lobes [80]), nor should mPFC lesions affect long-term instances still indexed by the intact MTL. In sum, MTL and mPFC lesions will produce specific problems encoding new-instance and schematic memories respectively, and differential retrograde amnesia gradients for recent and remote memories as a function of congruency.

Conclusions and future directions

Our aim has been to integrate research and theories on schema, novelty and the contributions of the MTL and mPFC to memory formation within a single framework. SLIMM is broadly consistent with a number of other consolidation theories [11–13,19], but makes the role of schema, mPFC and mPFC–MTL interactions more explicit. We accept that the framework is simplistic (e.g. when assuming mechanisms that are not yet fully empirically tested, such as resonance detection by mPFC), and faces problems with some existing data (Box 3). Nonetheless, at a minimum, SLIMM should help the understanding and interrelation of previous, sometimes paradoxical, findings in the neuroscientific and psychological literature. We hope it will also prompt future behavioural, neuroimaging and lesion studies that test the predictions outlined above. We believe that these developments will be of fundamental importance for optimising life-long learning and education, and for treating learning and memory disorders.

Box 3. Outstanding questions

- How precisely does the mPFC detect resonance, amplify congruent representations and suppress less congruent representations? How does it then potentiate synaptic changes between resonating neocortical representations (e.g. in terms of synaptic tagging or neurotransmitters)?
- How precisely do the mPFC and the hippocampus interact: does the mPFC only inhibit the MTL, or is there mutual competition? Do they interact differently during online experience and during offline replay (when instances may be retrieved from MTL)?
- Why do temporary lesions of the hippocampus impair encoding of schema-congruent information [94]? Does this happen only when the schemas are spatial, given evidence that hippocampus also represents spatial information in the rodent?
- How does the role of MTL in SLIMM relate to other theories of MTL function, such as scene construction [95] and future simulation [96]? Is reconsolidation related to integration of new information into schemas, and will blocking of reconsolidation thus also affect schema-related memories [12]?
- Do different subparts of the mPFC have different functions? And how are the mPFC subparts in the rodent related to those in human mPFC?
- What are the precise memory deficits following mPFC lesions, for example in terms of interference, transient retrograde amnesia (for congruent information) and possibly even encoding of greater episodic detail than in controls?
- How do the effects of schemas vary across development? For example, given the relatively slow maturation of PFC relative to other brain regions [97], does the ability to use schema change from childhood to early adulthood? And does healthy ageing reduce the mPFC efficiency in learning new schema-congruent information?
- How are more complex or structured schemas represented in the brain, and how can their formation be manipulated (during encoding, consolidation and retrieval) to optimise learning and education?

Acknowledgements

We wish to thank Bernhard Staresina, Pierre Gagnepain, Maria Wimber, Andrea Greve, Atsuko Takashima, Marijn Kroes, and three anonymous reviewers for helpful comments on the manuscript. We also wish to thank Vincent Schoots and Simon Strangeways for graphical assistance. M.T.R.v.K. is funded by Radboud University Medical Centre, and this work was made possible by the UK Experimental Psychology Society. R.N.H. is funded by the UK Medical Research Council (MC_US_A060_0046).

References

1 Wang, S.H. and Morris, R.G. (2010) Hippocampal-neocortical interactions in memory formation, consolidation, and reconsolidation. *Annu. Rev. Psychol.* 61, 49–79
2 Morris, R.G. (2006) Elements of a neurobiological theory of hippocampal function: the role of synaptic plasticity, synaptic tagging and schemas. *Eur. J. Neurosci.* 23, 2829–2846
3 Squire, L.R. (1992) Memory and the hippocampus: a synthesis from findings with rats, monkeys, and humans. *Psychol. Rev.* 99, 195–231
4 Takashima, A. *et al.* (2006) Declarative memory consolidation in humans: a prospective functional magnetic resonance imaging study. *Proc. Natl. Acad. Sci. U.S.A.* 103, 756–761
5 Frankland, P.W. and Bontempi, B. (2006) Fast track to the medial prefrontal cortex. *Proc. Natl. Acad. Sci. U.S.A.* 103, 509–510
6 van Kesteren, M.T. *et al.* (2010) Persistent schema-dependent hippocampal–neocortical connectivity during memory encoding and postencoding rest in humans. *Proc. Natl. Acad. Sci. U.S.A.* 107, 7550–7555
7 van Kesteren, M.T. *et al.* (2010) Retrieval of associative information congruent with prior knowledge is related to increased medial prefrontal activity and connectivity. *J. Neurosci.* 30, 15888–15894
8 Tse, D. *et al.* (2011) Schema-dependent gene activation and memory encoding in neocortex. *Science* 333, 891–895
9 van Kesteren, M.T. *et al.* (2012) Schema-dependent neocortical connectivity during information processing. In *Encyclopedia of the Sciences of Learning* (Seel, N., ed.), pp. 2953–2956, Springer
10 Ruiter, D.J. *et al.* (2010) How to achieve synergy between medical education and cognitive neuroscience? An exercise on prior knowledge in understanding. *Adv. Health Sci. Educ. Theory Pract.* DOI: 10.1007/s10459-010-9244-5
11 Squire, L.R. and Bayley, P.J. (2007) The neuroscience of remote memory. *Curr. Opin. Neurobiol.* 17, 185–196
12 McKenzie, S. and Eichenbaum, H. (2011) Consolidation and reconsolidation: two lives of memories? *Neuron* 71, 224–233
13 Winocur, G. *et al.* (2010) Memory formation and long-term retention in humans and animals: convergence towards a transformation account of hippocampal-neocortical interactions. *Neuropsychologia* 48, 2339–2356
14 Marr, D. (1970) A theory for cerebral neocortex. *Proc. R. Soc. Lond. B: Biol. Sci.* 176, 161–234
15 Rasch, B. and Born, J. (2007) Maintaining memories by reactivation. *Curr. Opin. Neurobiol.* 17, 698–703
16 Skaggs, W.E. and McNaughton, B.L. (1996) Replay of neuronal firing sequences in rat hippocampus during sleep following spatial experience. *Science* 271, 1870–1873
17 Wilson, M.A. and McNaughton, B.L. (1994) Reactivation of hippocampal ensemble memories during sleep. *Science* 265, 676–679
18 Scoville, W.B. and Milner, B. (1957) Loss of recent memory after bilateral hippocampal lesions. *J. Neurol. Neurosurg. Psychiatry* 20, 11–21
19 Moscovitch, M. *et al.* (2006) The cognitive neuroscience of remote episodic, semantic and spatial memory. *Curr. Opin. Neurobiol.* 16, 179–190
20 Frankland, P.W. and Bontempi, B. (2005) The organization of recent and remote memories. *Nat. Rev. Neurosci.* 6, 119–130
21 Maviel, T. *et al.* (2004) Sites of neocortical reorganization critical for remote spatial memory. *Science* 305, 96–99
22 Bontempi, B. *et al.* (1999) Time-dependent reorganization of brain circuitry underlying long-term memory storage. *Nature* 400, 671–675
23 Wierzynski, C.M. *et al.* (2009) State-dependent spike–timing relationships between hippocampal and prefrontal circuits during sleep. *Neuron* 61, 587–596
24 Takehara-Nishiuchi, K. and McNaughton, B.L. (2008) Spontaneous changes of neocortical code for associative memory during consolidation. *Science* 322, 960–963
25 Euston, D.R. *et al.* (2007) Fast-forward playback of recent memory sequences in prefrontal cortex during sleep. *Science* 318, 1147–1150
26 Ongur, D. and Price, J.L. (2000) The organization of networks within the orbital and medial prefrontal cortex of rats, monkeys and humans. *Cereb. Cortex* 10, 206–219
27 Pandya, D.N. *et al.* (1981) Efferent connections of the cingulate gyrus in the rhesus monkey. *Exp. Brain Res.* 42, 319–330
28 Maguire, E.A. (2001) Neuroimaging studies of autobiographical event memory. *Philos. Trans. R. Soc. Lond. B: Biol. Sci.* 356, 1441–1451
29 Tse, D. *et al.* (2007) Schemas and memory consolidation. *Science* 316, 76–82
30 Lewis, P.A. and Durrant, S.J. (2011) Overlapping memory replay during sleep builds cognitive schemata. *Trends Cogn. Sci.* 15, 343–351
31 Staresina, B.P. *et al.* (2009) Event congruency enhances episodic memory encoding through semantic elaboration and relational binding. *Cereb. Cortex* 19, 1198–1207
32 Atienza, M. *et al.* (2011) Semantic congruence enhances memory of episodic associations: role of theta oscillations. *J. Cogn. Neurosci.* 23, 75–90
33 Crespo-Garcia, M. *et al.* (2010) Functional neural networks underlying semantic encoding of associative memories. *Neuroimage* 50, 1258–1270
34 Siapas, A.G. *et al.* (2005) Prefrontal phase locking to hippocampal theta oscillations. *Neuron* 46, 141–151
35 Benchenane, K. *et al.* (2010) Coherent theta oscillations and reorganization of spike timing in the hippocampal–prefrontal network upon learning. *Neuron* 66, 921–936
36 Peyrache, A. *et al.* (2009) Replay of rule-learning related neural patterns in the prefrontal cortex during sleep. *Nat. Neurosci.* 12, 919–926
37 Paz, R. *et al.* (2008) Theta synchronizes the activity of medial prefrontal neurons during learning. *Learn. Mem.* 15, 524–531

38 Ranganath, C. *et al.* (2005) Functional connectivity with the hippocampus during successful memory formation. *Hippocampus* 15, 997–1005
39 Kumaran, D. *et al.* (2009) Tracking the emergence of conceptual knowledge during human decision making. *Neuron* 63, 889–901
40 Bar, M. (2009) The proactive brain: memory for predictions. *Philos. Trans. R. Soc. Lond. B: Biol. Sci.* 364, 1235–1243
41 Summerfield, C. and Koechlin, E. (2008) A neural representation of prior information during perceptual inference. *Neuron* 59, 336–347
42 Kumaran, D. and Maguire, E.A. (2009) Novelty signals: a window into hippocampal information processing. *Trends Cogn. Sci.* 63, 889–901
43 Knight, R. (1996) Contribution of human hippocampal region to novelty detection. *Nature* 383, 256–259
44 Strange, B.A. *et al.* (2000) Brain mechanisms for detecting perceptual, semantic, and emotional deviance. *Neuroimage* 12, 425–433
45 Gabrieli, J.D. (1998) Cognitive neuroscience of human memory. *Annu. Rev. Psychol.* 49, 87–115
46 Kan, I.P. *et al.* (2009) Contribution of prior semantic knowledge to new episodic learning in amnesia. *J. Cogn. Neurosci.* 21, 938–944
47 Rosenbaum, R.S. *et al.* (2005) The case of K.C.: contributions of a memory-impaired person to memory theory. *Neuropsychologia* 43, 989–1021
48 Vargha-Khadem, F. *et al.* (1997) Differential effects of early hippocampal pathology on episodic and semantic memory. *Science* 277, 376–380
49 O'Kane, G. *et al.* (2004) Evidence for semantic learning in profound amnesia: an investigation with patient H.M. *Hippocampus* 14, 417–425
50 Tramoni, E. *et al.* (2011) Long-term consolidation of declarative memory: insight from temporal lobe epilepsy. *Brain* 134, 816–831
51 Stark, C. *et al.* (2005) New semantic learning and generalization in a patient with amnesia. *Neuropsychology* 19, 139–151
52 Bayley, P.J. *et al.* (2008) New semantic learning in patients with large medial temporal lobe lesions. *Hippocampus* 18, 575–583
53 Duff, M.C. *et al.* (2006) Development of shared information in communication despite hippocampal amnesia. *Nat. Neurosci.* 9, 140–146
54 Sharon, T. *et al.* (2011) Rapid neocortical acquisition of long-term arbitrary associations independent of the hippocampus. *Proc. Natl. Acad. Sci. U.S.A.* 108, 1146–1151
55 Goshen-Gottstein, Y. *et al.* (2000) Intact implicit memory for newly formed verbal associations in amnesic patients following single study trials. *Neuropsychology* 14, 570–578
56 Schnider, A. (2003) Spontaneous confabulation and the adaptation of thought to ongoing reality. *Nat. Rev. Neurosci.* 4, 662–671
57 Gilboa, A. *et al.* (2009) Ventromedial prefrontal cortex lesions produce early functional alterations during remote memory retrieval. *J. Neurosci.* 29, 4871–4881
58 Bird, C.M. *et al.* (2004) The impact of extensive medial frontal lobe damage on 'theory of mind' and cognition. *Brain* 127, 914–928
59 Wang, S.H. *et al.* (2009) The precision of remote context memories does not require the hippocampus. *Nat. Neurosci.* 12, 253–255
60 Kirchhoff, B.A. *et al.* (2000) Prefrontal–temporal circuitry for episodic encoding and subsequent memory. *J. Neurosci.* 20, 6173–6180
61 Nyberg, L. (2005) Any novelty in hippocampal formation and memory? *Curr. Opin. Neurol.* 18, 424–428
62 Detterman, D.K. (1975) The von Restorff effect and induced amnesia: production by manipulation of sound intensity. *J. Exp. Psychol. Hum. Learn.* 1, 614–628
63 Kumaran, D. and Maguire, E.A. (2007) Match mismatch processes underlie human hippocampal responses to associative novelty. *J. Neurosci.* 27, 8517–8524
64 Fenker, D.B. *et al.* (2008) Novel scenes improve recollection and recall of words. *J. Cogn. Neurosci.* 20, 1250–1265
65 Wang, S.H. *et al.* (2010) Relevance of synaptic tagging and capture to the persistence of long-term potentiation and everyday spatial memory. *Proc. Natl. Acad. Sci. U.S.A.* 107, 19537–19542
66 Lisman, J.E. and Grace, A.A. (2005) The hippocampal–VTA loop: controlling the entry of information into long-term memory. *Neuron* 46, 703–713
67 Lisman, J. *et al.* (2011) A neoHebbian framework for episodic memory; role of dopamine-dependent late LTP. *Trends Neurosci.* 34, 536–547
68 Alba, J.W. and Hasher, L. (1983) Is memory schematic? *Psychol. Bull.* 93, 203–231
69 Grossberg, S. (1987) Competitive learning – from interactive activation to adaptive resonance. *Cogn. Sci.* 11, 23–63
70 Fell, J. and Axmacher, N. (2011) The role of phase synchronization in memory processes. *Nat. Rev. Neurosci.* 12, 105–118
71 Laroche, S. *et al.* (2000) Plasticity at hippocampal to prefrontal cortex synapses: dual roles in working memory and consolidation. *Hippocampus* 10, 438–446
72 Colgin, L.L. (2011) Oscillations and hippocampal-prefrontal synchrony. *Curr. Opin. Neurobiol.* 21, 467–474
73 Moscovitch, M. (2008) The hippocampus as a 'stupid', domain-specific module: implications for theories of recent and remote memory, and of imagination. *Can. J. Exp. Psychol.* 62, 62–79
74 Murre, J.M. (1996) TraceLink: a model of amnesia and consolidation of memory. *Hippocampus* 6, 675–684
75 Engel, A.K. *et al.* (2001) Dynamic predictions: oscillations and synchrony in top-down processing. *Nat. Rev. Neurosci.* 2, 704–716
76 Maguire, E.A. *et al.* (1999) The functional neuroanatomy of comprehension and memory: the importance of prior knowledge. *Brain* 122, 1839–1850
77 McClelland, J.L. *et al.* (1995) Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. *Psychol. Rev.* 102, 419–457
78 McKenna, T.M. *et al.* (1994) The brain as a dynamic physical system. *Neuroscience* 60, 587–605
79 Sweller, J. (1988) Cognitive load during problem-solving – effects on learning. *Cogn. Sci.* 12, 257–285
80 Patterson, K. *et al.* (2007) Where do you know what you know? The representation of semantic knowledge in the human brain. *Nat. Rev. Neurosci.* 8, 976–987
81 Henson, R.N. and Gagnepain, P. (2010) Predictive, interactive multiple memory systems. *Hippocampus* 20, 1315–1326
82 Poppenk, J. *et al.* (2011) Revisiting the novelty effect: when familiarity, not novelty, enhances memory. *J. Exp. Psychol. Learn. Mem. Cogn.* 36, 1321–1330
83 Bartlett, F.C. (1932) *Remembering: A Study in Experimental and Social Psychology*, Cambridge University Press
84 Piaget, J. (1926) *The Child's Conception of the World*, Littlefield, Adams
85 Bransford, J.D. *et al.* (2000) *How People Learn: Brain, Mind, Experience and School*, National Academy Press
86 Johnson-Laird, P.N. (1983) *Mental Models: Towards a Cognitive Science of Language, Inference, and Consciousness*, Harvard University Press
87 Neisser, U. (1976) *Cognitive Psychology*, Appleton-Century-Crofts
88 Anderson, R.C. (1984) Role of the reader's schema in comprehension, learning, and memory. In *Learning to Read in American Schools: Basic Readers and Content Texts* (Anderson, R.C. *et al.*, eds), pp. 243–257, Routledge
89 Watkins, M.J. and Gardiner, J.M. (1979) Appreciation of generate–recognize theory of recall. *J. Verb. Learn. Verb. Behav.* 18, 687–704
90 McClelland, J.L. and Rumelhart, D.E. (1986) Pyschological and Biological Models, In *Parallel Distributed Processing: Explorations in the Microstructure of Cognition* (Vol. 2), MIT Press
91 Tulving, E. and Kroll, N. (1995) Novelty assessment in the brain and long-term-memory Encoding. *Psychon. Bull. Rev.* 2, 387–390
92 von Restorff, H. (1933) Über die Wirkung von Bereichsbildungen im Spurenfeld. *Psychol. Forsch.* 18, 299–342
93 Friston, K. (2005) A theory of cortical responses. *Philos. Trans. R. Soc. Lond. B: Biol. Sci.* 360, 815–836
94 Bethus, I. *et al.* (2010) Dopamine and memory: modulation of the persistence of memory for novel hippocampal NMDA receptor-dependent paired associates. *J. Neurosci.* 30, 1610–1618
95 Hassabis, D. and Maguire, E.A. (2007) Deconstructing episodic memory with construction. *Trends Cogn. Sci.* 11, 299–306
96 Schacter, D.L. and Addis, D.R. (2009) On the nature of medial temporal lobe contributions to the constructive simulation of future events. *Philos. Trans. R. Soc. Lond. B: Biol. Sci.* 364, 1245–1253
97 Casey, B.J. *et al.* (2000) Structural and functional brain development and its relation to cognitive development. *Biol. Psychol.* 54, 241–257