ESSAY

Memory instability as a gateway to generalization

Edwin M. Robertson *

Institute of Neuroscience & Psychology, Centre for Cognitive Neuroimaging, University of Glasgow, Glasgow, United Kingdom

* edwin.robertson@glasgow.ac.uk

Abstract

Our present frequently resembles our past. Patterns of actions and events repeat throughout our lives like a motif. Identifying and exploiting these patterns are fundamental to many behaviours, from creating grammar to the application of skill across diverse situations. Such generalization may be dependent upon memory instability. Following their formation, memories are unstable and able to interact with one another, allowing, at least in principle, common features to be extracted. Exploiting these common features creates generalized knowledge that can be applied across varied circumstances. Memory instability explains many of the biological and behavioural conditions necessary for generalization and offers predictions for how generalization is produced.

Our past experience can aid our current and future performance. For instance, being a skilled tennis player may help when it comes to playing other racquet sports such as squash. Encouraging the transfer of skill from one situation to another also lies at the heart of many brain training and rehabilitative strategies.

The ability to generalize from a specific example to a category or concept is not restricted to the world of actions. The location of a reward in a navigation task can be found more quickly in subsequent versions of the task even though key aspects of the task, including the location of the reward, change [1]. Similarly, generalization can occur across different objects, facts, or events to create categories. Generalization therefore plays a key role in a wide array of cognitive functions. Yet, despite its clear importance and adaptive value, how and when generalization occurs is poorly understood.

Instability as opportunity

Generalization requires the identification of features common across experiences. For example, a common element of some navigation tasks is that a food pellet is never visible but is instead always buried within the sand. To discover the common feature requires a comparison and hence an interaction between memories for the different tasks. Interactions between memories can also lead to interference, after which a memory either is lost or becomes
inaccessible [2]. In either case, knowledge is lost, and this loss, due to the formation of another interfering memory, is a disadvantage of a plastic and hence changeable network (for a review, please see [3]).

When memories are formed in quick succession, each is unstable and susceptible to interference [4,5]. The interference between memories may, in part, be due to overlapping or functionally connected patterns of activity [4]. Normally, discrete populations of neurons, or ensembles, are activated within the hippocampus when a memory is formed. These ensembles overlap when memories for different contexts are acquired within a few hours of one another [6,7]. However, when a substantial delay is introduced (i.e., >8 hours), there is no longer an overlap in the memory representations. As a consequence, the temporal profile of the overlap in memory representation matches the time interval that a memory is unstable following its formation [4]. In sum, the overlapping representation of memories may lead to interference, making them unstable for several hours after their formation.

While instability leads to interference, and potentially to the loss of a memory, it also provides an opportunity for memories to interact with one another. Through these interactions, the statistical regularities shared between memories may activate common networks, strengthening them or at least distinguishing them in some way from those networks associated with the detailed differences between the memories. In this way, memories can, at least in principle, be compared, and shared features can be identified and potentially used to enable the creation of generalized knowledge that can be applied in different circumstances (Fig 1). Thus, whilst instability makes a memory vulnerable to forgetting, it is also linked to and perhaps even critical for the construction of generalized knowledge, which is central to so much of behaviour. This is not to suggest that instability is unique in supporting generalization (please see “Alternative mechanisms of generalization,” Box 1). Instead, what is explored here is memory instability as perhaps one gateway to generalization. Within this scenario, instability is adaptive, rather than purely detrimental, and offers a time window for the creation of generalizable knowledge (please see “Time window of generalization,” Box 2).

**Common conditions**

It is when memories are unstable that transfer is most prominent. For example, a recently acquired perceptual skill for detecting a stimulus at one location can easily generalize or transfer to other new locations [14]. Similarly, a newly acquired motor skill learnt with one hand can be easily transferred to the other hand (i.e., intermanual transfer; [15–18]). Yet, transfer is much reduced once a memory has been stabilized through consolidation [11]. Even following consolidation, it is only after a memory has once again become unstable that it can be modified and integrated with other memories, which is a prerequisite for subsequent generalization [19]. Together, these studies suggest that for many tasks, memory instability and transfer are found in similar circumstances. Yet, this is only circumstantial evidence. Few studies have measured memory instability and subsequent performance transfer together.

**Behaviour connects instability to generalization**

Recent work has tested the link between instability and generalization [5]. Learning a sequence of actions improves subsequent learning of a sequence of words. Conversely, learning a sequence of words improves the subsequent learning of a sequence of actions. This reciprocal pattern of transfer between different types of knowledge occurs provided 2 conditions are satisfied. Firstly, the motor and word sequences must share a common abstract structure or grammar. What is transferred is the high-level or abstract relationship between elements rather than knowledge of the individual elements themselves (i.e., words versus actions). Secondly,
the memory must be unstable for performance to transfer. The instability of the initial memory is correlated with subsequent transfer, suggesting that transfer is related to the instability of the memory. Yet, the relationship between instability and transfer goes beyond correlation. Stabilizing the initial memory, preventing it from being susceptible to interference, also prevented transfer to the subsequent memory task. Thus, transfer from one task to another was critically dependent upon memory instability.

Modifying stability modifies transfer

A manipulation that modifies memory stability can also modify transfer. Prolonged practice stabilizes a memory [20]. Usually, a new skill memory is so unstable that it can be disrupted by subsequently learning another different skill [4]. Yet, when the amount of training is increased, the newly formed skill memory is no longer susceptible to interference from further learning. The increased training has stabilized the memory. The increased training also reduces subsequent skill transfer. For example, after a short period of initial training on a visual task, participants show substantial transfer to a novel visual task, whereas after prolonged training, there is limited transfer [14,21]. Similarly, the transfer of skill between hands is frequently greater when there is a short rather than a prolonged period of initial training [22]. Extended training can also lead to the formation of habits, which show limited transfer [23]. Yet, not all extended training leads to the formation of a habit. Critically, a habit also requires a reduction in the importance of a goal [24]. The common feature across habits and extended training is the duration of practice, and thus, it seems likely that this is responsible for impairing transfer. Together these studies suggest that a manipulation—specifically, prolonged practice—can both stabilize a memory and reduce transfer from a learnt to a novel task.
Box 1. Alternative mechanisms of generalization

Different mechanisms may be responsible for supporting generalization under different circumstances. Learning may enhance the plasticity of a circuit, which could improve the learning of any subsequent task [9]. For example, learning a sequence of movements can improve how quickly participants adapt their movements to a novel visual environment, as occurs in prism adaptation. As a mechanism for generalization, it is potentially broad because it does not require shared attributes or knowledge between tasks; however, it does require similar or at least partially overlapping circuits to be involved in learning the different tasks. Within this framework, learning primes neuroplastic mechanisms, supporting the transfer of performance to subsequent tasks.

Forgetting may also drive generalization [10]. Losing information that appears only in specific situations allows a memory to become less tied to a specific circumstance and thus able to be applied generally across a wide range of circumstances (forgetting model). By contrast, rather than losing irrelevant information and diminishing the specificity of a memory, it may be possible to identify relevant information, a pattern, or a feature that reoccurs across a range of situations, enhance knowledge for that feature, and thus increase the efficacy of the memory across a range of situations. Identifying these common features requires an interaction or communication between memories, which can occur when they are unstable (instability model; [5,7,11]). Thus, generalization might be achieved by losing knowledge for specific situations; equally, it may also be achieved by enhancing knowledge for features that are a motif across a family of tasks.

Identifying the information that has to be either strengthened or weakened is an important challenge for both of these models of generalization. This challenge is substantial because it is also potentially a dynamic challenge. Initially, a feature of a task could recur across a family of tasks, and thus, strengthening knowledge of this feature would aid generalization. Yet, later in another circumstance, this same feature might be an idiosyncrasy of one particular situation, and therefore, strengthening knowledge of this feature, rather than aiding generalization, would only serve to increase the specificity of a memory. Another feature that these models share is that they link generalization to the loss of detailed information. For one model, forgetting, the loss of knowledge drives generalization by decreasing the specificity of a memory. Yet, in the other, it is a side effect. The interaction between unstable memories can lead to identifying common features between tasks, producing generalization, but it also leads to the disruption and loss of detailed knowledge [5]. Overall, the different models of forgetting and instability envisage generalization arising by weakening or strengthening different aspects of a memory (i.e., specifics versus recurring motifs, respectively). Both models share the common challenge of how to identify the information that needs to be strengthened or weakened, and both, either directly or indirectly, provide a link between generalization and the loss of detailed knowledge.

These are examples of how generalization may arise. Each mechanism is better suited to, and perhaps can only operate under, specific circumstances, and thus, it seems likely that at least in principle, these mechanisms could act together in a complementary fashion, with the strengths of one compensating for the weaknesses of another.
Prolonged practice leading to reduced transfer may seem counterintuitive. With prolonged practice comes increased proficiency, which might logically be expected to improve transfer. After all, the more knowledge is gained about one task, then the greater facility and hence perhaps the greater the potential to transfer performance to another related task. Yet, transfer does not seem to operate in this way. Prolonged practice appears to prevent rather than support transfer. Thus, transfer is not simply dictated by the accumulation of knowledge or performance. Instead, transfer occurs in the same circumstances as memory instability, is for at least one set of tasks critically dependent on instability, and can be prevented by prolonged practice, which stabilizes a memory. Together, these findings converge to suggest a link between memory instability and transfer.

A common mechanism for stability and transfer

Prolonged practice leading to reduced transfer may seem counterintuitive. With prolonged practice comes increased proficiency, which might logically be expected to improve transfer. After all, the more knowledge is gained about one task, then the greater facility and hence perhaps the greater the potential to transfer performance to another related task. Yet, transfer does not seem to operate in this way. Prolonged practice appears to prevent rather than support transfer. Thus, transfer is not simply dictated by the accumulation of knowledge or performance. Instead, transfer occurs in the same circumstances as memory instability, is for at least one set of tasks critically dependent on instability, and can be prevented by prolonged practice, which stabilizes a memory. Together, these findings converge to suggest a link between memory instability and transfer.

Changes in GABA have been linked to changes in performance transfer. The concentration of GABA within the cortex can be modified using a noninvasive brain stimulation technique called direct current stimulation. In this technique, 2 electrodes (an anode and cathode) are placed on the scalp of a human participant, and a small current is passed between the electrodes (for a review, please see [25]). Placing the anode of a stimulation device over the motor cortex decreases GABA concentrations in this area [26]. Decreasing GABA in this way
contralateral to the trained hand enhances subsequent transfer to the untrained hand [27]. Thus, GABA is linked to performance transfer: an increase in GABA, due to prolonged practice, impairs transfer, while a decrease in GABA, due to current stimulation, enhances transfer. Equally, an increase in GABA has been linked to an increase in memory stability [20,28]. Together, these studies reveal a mechanistic link between stability and transfer. Subsequent studies may further test the nature of this link by using pharmacological methods to specifically modify GABA. However, even with converging evidence to show that GABA mechanistically links stability with transfer, it should not be assumed that this link depends solely on GABA. Potentially, GABA is only one component of what is likely to be shown, in time, as a complex and diverse neurochemical mechanism linking memory stability to performance transfer.

**Circuits of stability**

A link between memory stability and generalization is also present at the level of networks and brain circuits. One part of a network critical for generalization and for the interaction between unstable memories appears to be the prefrontal cortex.

Several studies have shown that the prefrontal cortex makes a critical contribution to generalization. For instance, lesions to the ventromedial prefrontal cortex in humans or disruption to prefrontal function with transcranial magnetic stimulation (TMS) prevents semantic generalization [29,30]. This is when participants learn a list of semantically related words and subsequently incorrectly identify another semantically related word as coming from the list [31]. These errors due to semantic generalization, frequently called false memories, are decreased when the function of the prefrontal cortex is impaired.

Other studies have shown that the prefrontal cortex makes a critical contribution to memory instability. Disrupting the function of the human prefrontal cortex, with TMS, prevents newly formed unstable memories from being susceptible to interference [32]. Similarly in rodents, a lesion to the frontal cortex also prevents the interaction between new unstable memories ([33]; for a review, please see [4]). Together, these studies suggest that the prefrontal cortex is responsible for creating interference between newly acquired memories [4].

The prefrontal cortex may support interference between memories by affecting their representation. Unstable newly formed memories have an overlapping representation within the hippocampus [6,7]. The prefrontal cortex exerts an influence upon the representation of motor skill memories in the primary motor cortex, and thus, at least in principle, it may have a similar role in influencing the representation of memories in the hippocampus [34,35]. Disruption to the prefrontal function could then transform the overlapping representation to a set of independent representations. Without an overlap, there may be no communication or interference between the memories, which is consistent with the work on rodents and humans [32,33]. Envisaging the overlapping representations as providing communication between memories would explain their vital contribution to transfer [7]. Information from one memory, for example, about the fear associated with one context has to be communicated to another memory for fear to transfer to a previously neutral context. Thus, instability and susceptibility to interference may be achieved by the prefrontal cortex creating overlapping memory representations, which are critical for transfer. This may explain the critical contribution of the prefrontal cortex to semantic generalization.

A similar mechanism may also allow generalization between old and recently formed memories. Previously formed memories are reactivated when related new information is being encoded into a memory. Reactivation of the memory is due to a dialogue between the medial prefrontal cortex and the hippocampus ([36–38]; for a review, please see [39]). When
reactivated, an old memory becomes once again represented within the hippocampus, and it reverts to an unstable state, which allows it to be modified [40]; it can be then strengthened or integrated with a new memory [19,41]. The instability, modifiability, and capacity to be integrated with new memories suggest a communication between old and new memories, which may be achieved by the old reactivated memory sharing an overlapping representation with the new recently formed memory ([6]; for a review, please see [10]). Such communication between the memories may allow the identification of common elements, or motifs, which in turn supports the creation of generalizable knowledge. Once the common elements have been identified, the newly acquired memory may quickly cease to be represented within the hippocampus, instead becoming part of a cortical representation of the common properties shared across the old and new memory (i.e., part of a schema; [1]). Overall, the prefrontal cortex mediates the reactivation of an old memory, leading to it becoming once again unstable and able to have an overlapping representation with a new related memory, which provides perhaps the basis for communication between memories and generalization across them.

Prefrontal circuits are not unique in making a critical contribution to generalization. The circuit can alter depending upon the nature of the common characteristic or repeating regularity that is being generalized. For instance, a circuit that includes the ventromedial prefrontal cortex is critical for semantic generalization [29,30]. By contrast, when the common feature is no longer semantic but instead, for instance, spatial position, then another brain area, the angular gyrus, is critical for generalization across tasks [42,43]. Similarly, the circuit critical for a newly formed unstable memory alters depending upon the type of information learnt [44–47]. While the circuit dedicated to memory stability and generalization may vary, what does not vary perhaps is the overlapping relationship between stability and generalization.

Instability: A trade-off between detail and generalization

Detailed knowledge can easily be lost when a newly formed memory is unstable and susceptible to disruption. For example, rather than recalling a complete list of 12 words, a person might only recall 10 words [5,32,48]. However, this loss of detailed knowledge can come with a benefit. There is a positive correlation between the loss of detailed knowledge and transfer of performance to a subsequent related task. For example, skill learnt in performing a sequence of actions is lost in direct proportion to the performance transferred to learning a sequence of words [5].

Similarly, knowledge of word sequence is lost in direct proportion to the performance transferred to an action sequence. This pattern of reciprocal transfer is observed when the sequence of words or actions share a common abstract structure. Converging with this behavioural work showing a trade-off between detailed knowledge and generalization is more recent functional imaging work (Fig 1, [8]).

Patterns of activation within the human hippocampus have also revealed a trade-off between detailed knowledge and generalization [8]. Maintaining a detailed knowledge of a learnt association between an object and a scene was measured as the match between the pattern of neural activation at the initial encoding and its subsequent retrieval. A tight match between the pattern of activation at encoding and retrieval indicated retention of detailed knowledge. Each of the 128 objects was uniquely paired with one of only 4 scenes. This allowed memories to be related to one another through a common shared scene. Similarity of activation between those memories with a shared common scene provided a measure of generalization. Using these analysis techniques provided measures of both detailed knowledge retention and generalization, which were negatively correlated to one another, revealing a trade-off between detailed knowledge retention and generalization. Overall, for both the behavioural
and the functional imaging work, the greater the loss of detailed knowledge is, the greater the ability to generalize. This trade-off can be explained by memory instability being necessary for generalization.

Instability makes a memory susceptible to interference. The greater this instability is, the greater the interference, and the greater the loss of detailed knowledge. However, instability also increases the interaction between memories, potentially allowing the identification of shared common features, which can be exploited to allow transfer and generalization between different but related situations. Thus, instability may well explain the trade-off between the loss of detailed knowledge and generalization (Fig 1).

**Reward as a modifying factor of memory stability and generalization**

After its formation, a memory is stabilized over several hours. Such offline processing can be affected by modifying factors, one of which appears to be reward. For example, providing a reward following learning for the retrieval of specific items enhances the subsequent recall of those items [49]. Similarly, rewarding the acquisition of a motor skill enhances the skill improvements that develop offline during consolidation [50].

The connectivity of circuits and what happens within those circuits following memory formation are both affected by reward. An increase in connectivity between the visual cortex and both the anterior hippocampus and the ventral tegmental area is associated with a reward [51]. There is also an increase in the frequency with which the pattern of neural activity present at memory formation is replayed offline during consolidation [52,53]. Together, these studies provide evidence that reward can affect offline memory processing during consolidation.

A wide range of memory changes occur during offline processing [54]. Stabilization is but one example of these changes. Other examples of offline processing, such as memory enhancement, are clearly affected by reward [50]; yet, what remains less clear is whether specifically memory stabilization is affected by reward. It is conceivable that a reward can affect memory stability. After all, reward leads to the release of dopamine, affecting synaptic plasticity mechanisms, which, at least in principle, may shorten the interval that a memory is unstable for [55,56]. This would suggest that reward might stabilize a memory. Increasing the stability of a memory with a reward may reduce the propensity to transfer learning between related tasks.

At first, the idea that reward will impair a behaviour, in this case the transfer of learning, may seem counterintuitive; however, it may have important adaptive benefits. Reward may be sculpting behaviour so that high performance is focused precisely on those tasks that yield a reward, and not upon related tasks that may not yield any reward. Overall, it seems highly likely that memory stability can be manipulated with reward; yet, what remains to be tested is whether this manipulation affects subsequent transfer between related tasks.

**Stressing memory**

As is the case for a rewarding stimulus, a stressful or aversive stimulus can affect the offline processing of a memory. For example, in rodent studies, applying a foot shock immediately after memory formation can enhance subsequent consolidation [57]. Stress may act to stabilize a memory or at least reduce its susceptibility to interference by altering the connectivity of circuits so that memory processing becomes isolated or independent from other processing [58]. With this reduced vulnerability to interference, a memory is stabilized, and retention improved.

Whether these changes in the memory stability due to stress affect the development of generalizable knowledge is not currently clear. Some studies have suggested that stress increases
generalization, others that it makes minimal difference, and others that it decreases generalization [59–61]. For some of these studies, stress did not increase knowledge retention, and thus, there is no evidence that stress affected consolidation or memory stability. When knowledge retention is increased, indicating enhanced memory stabilization, there is a decrease in generalization. The latter observations are consistent with the model proposed here, with stability favouring accurate detailed retention whilst instability favours generalization.

Memory instability is also implicated in the transfer of fear from one context to another. Fear paired within one context will transfer to another neutral context when, and only when, the acquisition of a memory for both contexts is separated by only a few hours (≈5 hours; [7]). There is no transfer when the interval between acquiring the memories is increased to 1 week. A similar time course is followed by memory instability, with a memory being unstable and susceptible to interference within the first hours of its formation and stable within a day (and certainly within 1 week) [4]. Instability and subsequent transfer are present over a similar time window. Thus, it is conceivable that the capacity to transfer fear between contexts is related to and, consistent with the current model, potentially relies upon memory instability.

Transfer is also linked to instability through how memories are represented. When acquired in quick succession, memories share a neural circuit, or ensemble. When the overlap between memory ensembles is decreased through aging, transfer of fear to a new context is impaired [7]. Conversely, when the overlap in aging rodents is rescued through experimental manipulation, the transfer of fear to a new context is once again possible. What this beautifully illustrates is that an overlapping representation is necessary for subsequent transfer. The overlapping representation, on the one hand, provides a means for different memories to interact, common features to be identified, and transfer to happen. Yet, on the other hand, it makes memories susceptible to interference and perhaps makes them unstable.

Overall, (A) reducing memory instability with aversive stimuli impairs transfer; (B) memory instability is transient, lasting for only a matter of hours, and it is only during this time window that transfer of fear from one context to another is possible; and (C) instability may be due to an overlapping memory representation, which is critical for transfer. Together, these results converge to suggest that memory instability may have an important role to play in transfer and the creation of generalizable knowledge (please see “Predictions,” Box 3).

**Brain state**

Sleep has been linked to supporting generalization (for a review, please see [64]). During sleep, memories continue to be processed, are enhanced, and are reorganized [54]. The reorganization of past events potentially allows hidden patterns to be uncovered.

For instance, infants demonstrate knowledge for an artificial grammar of nonsense letter strings only after sleep during a nap. This is achieved by identifying repeating patterns—in this example, the grammatical structure common to letter sequences. Specifically, the first syllable (“PEL”) predicts the final syllable, and as a consequence, both “PELwadimRUD” and “PELchilaRUD” are valid grammatical structures for these nonsense letter strings [65,66].

Sleep has also been implicated in transitive inference when the high-order structure of the relationship between arbitrary symbols (such as fractal patterns) is uncovered based solely upon exposure to low-order relationships [67]. For example, when participants are exposed to simple pairings such as A > B, B > C, and C > D, the appropriate inference from exposure to these is that A > D, which is enhanced over sleep. Thus, sleep provides an environment that promotes the extraction of rules, the identification of repeating abstract patterns, and generalization across tasks.
Generalization during sleep may be linked to memory instability. Memories are reactivated during sleep, which may cause the memory to become unstable. The pattern of neural activity present during memory formation can be found again during sleep. For example, the neuronal firing patterns during a motor learning task are replayed again in the motor cortex of a rodent.

Box 3. Predictions

Memory instability may provide one gateway to the development of generalizable knowledge. As a consequence, modifying memory stability through fear, reward, or even prolonged practice could modify subsequent transfer of performance across related tasks or situations. Reward and fear are both predicted to decrease transfer because both enhance consolidation and thus are assumed to increase memory stability. Evidence is accumulating that is consistent with this view; however, as yet, there is very little direct evidence, because no single study has modified memory stability using fear or reward, measured that change, and examined the subsequent effects, if any, in transfer. Such studies are not without challenges. For instance, the shift in processing from a goal-based to a habit-based strategy that fear promotes could couple performance more tightly to a particular context and thus impair transfer, regardless of changes in memory stability [57,62,63].

Another approach to testing the link between memory instability and transfer is to better understand the conditions necessary for transfer. Fear can only transfer from one context to another neutral context when the memories for each context are formed within a few hours of one another [7]. The transfer in these circumstances could be related to memory instability. Memories are unstable for a few hours after their formation, the same time interval during which transfer is possible. This suggests a link between memory instability and subsequent transfer. Certainly, the instability of a memory for a sequence is related to the subsequent transfer to a different type of sequence (actions versus words; [5]). However, these are very different types of transfer—the latter case requires identifying and using the common sequential attributes to transfer performance, whereas in the former, there is no common element; instead, fear is being misattributed to a neutral context. Despite these differences, both may be dependent upon memory instability to enable transfer; alternatively, these differences may translate onto different mechanisms (see Box 1).

Transfer is also dependent upon the nature of the memory representation. When learnt within hours of one another, memories have overlapping representations [6,7]. Elegant work has shown that these overlapping representations are critical for transfer [7]. Manipulations that modify transfer would therefore be predicted to alter this memory representation and stability. For example, prolonged practice may diminish transfer and increase memory stability by promoting the creation of nonoverlapping or independent representations. The rise in GABA during prolonged practice may be responsible for diminishing the excitability within a shared overlapping representation and split it into independent representations [20]. Increasing excitability between independent representations can rescue the capacity to transfer fear from one context to another [7]. Manipulating memory representations may provide a way to test for a mechanistic link between instability and transfer.

Generalization during sleep may be linked to memory instability. Memories are reactivated during sleep, which may cause the memory to become unstable. The pattern of neural activity present during memory formation can be found again during sleep. For example, the neuronal firing patterns during a motor learning task are replayed again in the motor cortex of a rodent.
during sleep, and this replay is correlated with the subsequent sleep-dependent performance improvements [13,68]. These reactivations may lead to the memory becoming unstable [69]. When retrieved during wakefulness, a memory is rendered unstable, vulnerable to interference, just as the memory had been soon after its initial formation. Similarly, the reactivation of a memory during sleep may also make it unstable.

Elegant work has demonstrated that memories can be artificially reactivated during sleep [70,71]. For instance, a memory formed while a sensory cue, such as an odour, is presented can be reactivated when that same sensory cue is represented during sleep [70]. The same pattern of functional activation found during learning is found again when the sensory cue is represented during sleep.

Yet, when a specific memory is reactivated during sleep, it remains invulnerable to interference [72]. Interference from further learning is only one measure of memory instability, and a memory may be unstable, despite not being susceptible to interference. Changes in brain state during sleep may make a memory, even an unstable memory, invulnerable to interference.

During large parts of sleep, the effective functional connectivity of the human brain is markedly reduced. For instance, the waveform evoked by applying TMS to the motor cortex travels a substantially shorter distance when applied during slow-wave sleep than during wakefulness [73,74]. Along with this decrease in functional connectivity, there is a change in brain organization. Specifically, the brain becomes more modular. Functionally connected circuits remain, but these circuits are smaller and more constrained, lacking the widespread connections present during wakefulness [75].

Thus, memories may well become unstable over sleep but remain protected from interference because of the poor functional connection amongst brain areas. Unstable memories are constrained within functionally discrete, independent circuits, and therefore, interference between memories is minimized. Thus, sleep may provide an ideal environment for memories to become unstable because they are protected from interference. However, this environment may also restrict the generalization that is possible during sleep.

With connectivity limited during sleep, generalization may only occur between those memories represented within restricted circuits. This may mean that generalization can only occur across memories with certain properties such as having the same or similar content. By contrast, generalization can occur between memories with different content during wakefulness (i.e., between actions and words; [5]). Yet, transient restorations in long-range connectivity associated, for example, with sleep spindles may be sufficient to allow communication across brain areas to support generalization across diverse memories [76–78]. Alternatively, interludes of rapid eye movement (REM) sleep may be sufficient to restore connectivity when, or if, coordinated with episodes of memory reactivation that predominately occur when connectivity is reduced during slow-wave sleep [64]. Sleep may restrict the damaging effects of interference upon memories by having generally limited connectivity whilst simultaneously having brief restorations in connectivity to allow communication and potentially generalization across memories.

**Conclusions and beyond**

Broadly, there appear to be at least 2 contrasting perspectives on memory instability, which are unlikely to be mutually exclusive. One perspective sees instability as arising from the unique requirements of biology. For instance, it takes time to synthesize the protein necessary to stabilize a memory, and thus, an interval of instability follows. From this perspective, memory instability is simply the inevitable consequence of having algorithmic processes implemented within a biological substrate. Implement those same processes within a different substrate,
such as silicon, and memory instability may well vanish without any loss of memory function. In an alternative perspective, memory instability—and potentially the offline processing of memories more broadly—may make an indispensable contribution to the algorithms necessary for memory function. Instability may provide an opportunity for a particular form of computation or algorithm that is critical for memory function.

Instability may be critical to uncovering patterns common across different memories. It provides an opportunity for a comparison between different memories, allowing common features to be identified, extracted, and exploited. Once stabilized, a memory becomes invulnerable to interference; yet, it may also lose its ability to interact with other memories, and thus, features common to the memories can no longer be identified. Consistent with this idea, the transfer of performance from one task to another is most prominent in those circumstances that favour memory instability [11,14–18,20]. Yet, this is more than simply a circumstantial link. The transfer of performance across related sequential tasks is correlated with memory instability [5]. Stabilizing those memories, either through subtle changes to the tasks or inserting a time interval to allow consolidation to take place, prevents transfer [5]. Similarly, prolonged practice stabilizes perceptual memories and is associated with decreased transfer [14,20,21]. Conversely, reversing those neurochemical changes associated with memory stabilization enhances transfer [20,26–28]. Instability then may be critical for generalization. Instability also explains the trade-off between the loss of detailed knowledge and the creation of generalizable knowledge found in behavioural and functional imaging work (Fig 1; [5,8]). This relationship between instability and generalization may remain even in other brain states. Sleep has been widely associated with promoting generalization, and it is during sleep that the patterns of neural activity present during memory formation are replayed. Memories are rendered unstable through replay yet are protected from interference because of the changes in brain organization and connectivity that take place during sleep. Overall, from across a diverse array of studies, a consistent link emerges, connecting memory instability to generalization.

Acknowledgments
I am grateful to Philippe Schyns for encouraging me to produce this work.

References
1. Tse D, Langston RF, Kakeyama M, Bethus I, Spooner PA, Wood ER, et al. Schemas and memory consolidation. Science. 2007; 316(5821):76–82. https://doi.org/10.1126/science.1135935 PMID: 17412951.
2. Ryan TJ, Roy DS, Pignatelli M, Arons A, Tonegawa S. Memory. Engram cells retain memory under retrograde amnesia. Science. 2015; 348(6238):1007–13. PMID: 26023136
3. Abraham WC, Robins A. Memory retention—the synaptic stability versus plasticity dilemma. Trends in neurosciences. 2005; 28(2):73–8. https://doi.org/10.1016/j.tins.2004.12.003 PMID: 15667929.
4. Robertson EM. New insights in human memory interference and consolidation. Current biology: CB. 2012; 22(2):R66–71. https://doi.org/10.1016/j.cub.2011.11.051 PMID: 2280913.
5. Mosha N, Robertson EM. Unstable Memories Create a High-Level Representation that Enables Learning Transfer. Current biology: CB. 2016; 26(1):100–5. https://doi.org/10.1016/j.cub.2015.11.035 PMID: 26711483.
6. Rashid AJ, Yan C, Mercaldo V, Hsiang HL, Park S, Cole CJ, et al. Competition between engrams influences fear memory formation and recall. Science. 2016; 353(6297):383–7. https://doi.org/10.1126/science.aaf0594 PMID: 27463673.
7. Cai DJ, Aharoni D, Shuman T, Shobe J, Biane J, Song W, et al. A shared neural ensemble links distinct contextual memories encoded close in time. Nature. 2016; 534(7605):115–8. https://doi.org/10.1038/nature17955 PMID: 27251287
8. Tompary A, Davachi L. Consolidation Promotes the Emergence of Representational Overlap in the Hippocampus and Medial Prefrontal Cortex. Neuron. 2017; 96(1):228–41 e5. https://doi.org/10.1016/j.neuron.2017.09.005 PMID: 28957671

9. Seidler RD. Multiple motor learning experiences enhance motor adaptability. Journal of cognitive neuroscience. 2004; 16(1):65–73. https://doi.org/10.1162/08989290432755566 PMID: 15006037.

10. Schlichting ML, Frankland PW. Memory allocation and integration in rodents and humans. Current opinion in behavioral sciences. 2017; 17:90–8. https://doi.org/10.1016/j.cobeha.2017.07.013

11. Stockinger C, Thurer B, Focke A, Stein T. Intermuscular transfer characteristics of dynamic learning: direction, coordinate frame, and consolidation of interlimb generalization. Journal of neurophysiology. 2015; 114(6):3166–76. https://doi.org/10.1152/jn.00727.2015 PMID: 26424581

12. Lutz ND, Diekelmann S, Hinse-Stern P. Born J, Rauss K. Sleep Supports the Slow Abstraction of Gist from Visual Perceptual Memories. Scientific reports. 2017; 7:42950. https://doi.org/10.1038/srep42950 PMID: 28211489

13. Genzel L, Robertson EM. To Replay, Perchance to Consolidate. PLoS Biol. 2015; 13(10):e1002285. https://doi.org/10.1371/journal.pbio.1002285 PMID: 26496145

14. Harris H, Gilksberg M, Sagi D. Generalized perceptual learning in the absence of sensory adaptation. Current biology: CB. 2012; 22(19):1813–7. https://doi.org/10.1016/j.cub.2012.07.059 PMID: 22921366.

15. Wachs K, Pascual-Leone A, Grafman J, Hallett M. Intermuscular transfer of implicit knowledge of sequential finger movements. Neurology. 1994; 44(Supplement 2):805S.

16. Verwey W, Clegg B. Effector-dependent sequence learning in the serial RT task. Psychological Research. 2005; 69:42–51. https://doi.org/10.1007/s00426-004-0181-x PMID: 15235913

17. Verwey W, Wright D. Effector-independent and effector-dependent learning in the discrete sequence production task. Psychological Research. 2004; 68:64–70. https://doi.org/10.1007/s00426-003-0144-7 PMID: 12955505

18. Cohen DA, Pascual-Leone A, Press DZ, Robertson EM. Off-line learning of motor skill memory: A double dissociation of goal and movement. PNAS. 2005; 102(50):18237–41. https://doi.org/10.1073/pnas.0506072102 PMID: 16330773

19. Lee JL. Memory reconsolidation mediates the strengthening of memories by additional learning. Nature neuroscience. 2008; 11(11):1264–6. https://doi.org/10.1038/nn.2205 PMID: 18949987.

20. Shibata K, Sasaki Y, Bang JW, Walsh EG, Machizawa MG, Tamaki M, et al. Overlearning hyperstabilizes a skill by rapidly making neurochemical processing inhibitory-dominant. Nature neuroscience. 2017; 20(3):470–5. https://doi.org/10.1038/nn.4490 PMID: 28135242

21. Liang J, Zhou Y, Fahle M, Liu Z. Limited transfer of long-term motion perceptual learning with double training. Journal of vision. 2015; 15(10):1. https://doi.org/10.1167/15.10.1 PMID: 26230916.

22. Rand MK, Hikosaka O, Miyachi S, Lu X, Miyashita K. Characteristics of a long-term procedural skill in the monkey. Experimental brain research. 1998; 118(3):293–7. PMID: 9497137.

23. Wood W, Tam L, Witt MG. Changing circumstances, disrupting habits. Journal of personality and social psychology. 2005; 88(6):918–33. https://doi.org/10.1037/0022-3514.88.6.918 PMID: 15982113.

24. Robbins TW, Costa RM. Habits. Current biology: CB. 2012; 22(19):1813–7. https://doi.org/10.1016/j.cub.2012.07.059 PMID: 22921366.

25. Makkar SR, Zhang SQ, Cranney J. Behavioral and neural analysis of GABA in the acquisition, consolidation, reconsolidation, and extinction of fear memory. Neuropsychopharmacology: official publication of the American College of Neuropsychopharmacology. 2010; 35(6):1625–52. https://doi.org/10.1038/npp.2010.53 PMID: 20410874

26. Warren DE, Jones SH, Duff MC, Tranel D. False recall is reduced by damage to the ventromedial prefrontal cortex: implications for understanding the neural correlates of schematic memory. The Journal of...
neuroscience: the official journal of the Society for Neuroscience. 2014; 34(22):7677–82. https://doi.org/10.1523/JNEUROSCI.0119-14.2014 PMID: 24872571

30. Berkers RM, van der Linden M, de Almeida RF, Muller NC, Bovy L, Dresler M, et al. Transient mediofrontal perturbation reduces false memory formation. Cortex; a journal devoted to the study of the nervous system and behavior. 2017; 88:42–52. https://doi.org/10.1016/j.cortex.2016.12.015 PMID: 28068640.

31. Roediger HL, McDermott KB. Creating False Memories: Remembering Words Not Presented in Lists. Journal of Experimental Psychology: Learning, Memory and Cognition. 1995; 21(4):803–14.

32. Cohen DA, Robertson EM. Preventing interference between different memory tasks. Nature neuroscience. 2011; 14(8):953–5. Epub 2011/06/28. https://doi.org/10.1038/nn.2840 PMID: 21706019

33. Costanzi M, Sarauldi D, Rossi-Arnaud C, Aceti M, Cestari V. Memory impairment induced by an interfering task is reverted by pre-frontal cortex lesions: a possible role for an inhibitory process in memory suppression in mice. Neuroscience. 2009; 158(2):503–13. https://doi.org/10.1016/j.neuroscience.2008.08.026 PMID: 18790014.

34. Tunovic S, Press DZ, Robertson EM. A physiological signal that prevents motor skill improvements during consolidation. The Journal of neuroscience: the official journal of the Society for Neuroscience. 2014; 34(15):5302–10. https://doi.org/10.1523/JNEUROSCI.3497-13.2014 PMID: 24719108

35. Breton J, Robertson EM. Flipping the switch: mechanisms that regulate memory consolidation. Trends in cognitive sciences. 2014; 18(12):629–34. https://doi.org/10.1016/j.tics.2014.08.005 PMID: 25240672.

36. Place R, Farovik A, Brockmann M, Eichenbaum H. Bidirectional prefrontal-hippocampal interactions support context-guided memory. Nature neuroscience. 2016; 19(8):992–4. https://doi.org/10.1038/nn.4327 PMID: 27322417

37. Kaplan R, Bush D, Bonnefond M, Bandettini PA, Barnes GR, Doeller CF, et al. Medial prefrontal theta phase coupling during spatial memory retrieval. Hippocampus. 2014; 24(6):656–65. https://doi.org/10.1126/science.aaa0781 PMID: 24497013

38. Brown TI, Carr VA, LaRocque KF, Favila SE, Gordon AM, Bowles B, et al. Prospective representation of navigational goals in the human hippocampus. Science. 2016; 352(6291):1323–6. https://doi.org/10.1126/science.aaf0784 PMID: 27284194.

39. Morton NW, Sherrill KR, Preston AR. Memory integration constructs maps of space, time, and concepts. Current opinion in behavioral sciences. 2017; 17:161–8. https://doi.org/10.1016/j.cobeha.2017.08.007 PMID: 28924579

40. Lee JL. Reconsolidation: maintaining memory relevance. Trends in neurosciences. 2009; 32(8):413–20. https://doi.org/10.1016/j.tins.2009.05.002 PMID: 19640595

41. Hupbach A, Gomez R, Hardt O, Nadel L. Reconsolidation of episodic memories: a subtle reminder triggers integration of new information. Learning & memory. 2007; 14(1–2):47–53. https://doi.org/10.1101/lm.365701 PMID: 17202429

42. Amorapanth P, Kranjec A, Bromberger B, Lehet M, Widick P, Woods AJ, et al. Language, perception, and the schematic representation of spatial relations. Brain and language. 2012; 120(3):226–36. https://doi.org/10.1016/j.bandl.2011.09.007 PMID: 22070948

43. Kranjec A, Ianni G, Chatterjee A. Schemas reveal spatial relations to a patient with simultanagnosia. Cortex; a journal devoted to the study of the nervous system and behavior. 2013; 49(7):1983–8. https://doi.org/10.1016/j.cortex.2013.03.005 PMID: 23643246.

44. Muellbacher W, Ziemann U, Wissel J, Dang N, Kofler M, Facchinis S, et al. Early consolidation in human primary motor cortex. Nature. 2002; 415(6872):640–4. https://doi.org/10.1038/nature712 PMID: 11807497.

45. Robertson EM, Press DZ, Pascual-Leone A. Off-line learning and the primary motor cortex. The Journal of neuroscience: the official journal of the Society for Neuroscience. 2005; 25(27):6372–8. https://doi.org/10.1523/JNEUROSCI.1851-05.2005 PMID: 16006277.

46. Galea JM, Albert NB, Ditye T, Miall RC. Disruption of the dorsolateral prefrontal cortex facilitates the consolidation of procedural skills. Journal of cognitive neuroscience. 2010; 22(6):1158–64. https://doi.org/10.1162/jocn.2009.21259 PMID: 19413472.

47. Breton J, Robertson EM. Dual enhancement mechanisms for overnight motor memory consolidation. Nat Hum Behav. 2017; 1: 0111.

48. Brown RM, Robertson EM. Off-Line Processing: Reciprocal Interactions between Declarative and Procedural Memories https://doi.org/10.1523/JNEUROSCI.2799-07.2007 J Neurosci. 2007; 27(39):10468–75. PMID: 17898218
49. Patil A, Murty VP, Dunsmoor JE, Phelps EA, Davachi L. Reward retroactively enhances memory consolidation for related items. Learning & memory. 2017; 24(1):65–9. https://doi.org/10.1101/lm.042978.116 PMID: 27980078

50. Abe M, Schambra H, Wassermann EM, Luckenbaugh D, Schweighofer N, Cohen LG. Reward improves long-term retention of a motor memory through induction of offline memory gains. Current biology: CB. 2011; 21(7):557–62. https://doi.org/10.1016/j.cub.2011.02.030 PMID: 21419628

51. Murty VP, Tompany A, Adcock RA, Davachi L. Selectivity in Postcoding Connectivity with High-Level Visual Cortex Is Associated with Reward-Motivated Memory. The Journal of neuroscience: the official journal of the Society for Neuroscience. 2017; 37(3):537–45. https://doi.org/10.1523/JNEUROSCI.4032-15.2016 PMID: 28100737

52. Lansink CS, Golstein PM, Lankelma JV, Joosten RN, McNaughton BL, Pennartz CM. Preferential reactivation of motivationally relevant information in the ventral striatum. The Journal of neuroscience: the official journal of the Society for Neuroscience. 2008; 28(25):6372–82. https://doi.org/10.1523/JNEUROSCI.1054-08.2008 PMID: 18562607

53. Singer AC, Frank LM. Rewarded outcomes enhance reactivation of experience in the hippocampus. Neuron. 2009; 64(8):910–21. https://doi.org/10.1016/j.neuron.2009.11.016 PMID: 20064396

54. Robertson EM. From creation to consolidation: a novel framework for memory processing. PLoS Biol. 2009; 7(1):e19. https://doi.org/10.1371/journal.pbio.1000019 PMID: 19175290

55. Kentros CG, Agnihotri NT, Streater S, Hawkins RD, Kandel ER. Increased attention to spatial context increases both place field stability and spatial memory. Neuron. 2004; 42(2):283–95. PMID: 15091343.

56. Hamilton TJ, Wheatley BM, Sinclair DB, Bachmann M, Larkum ME, Colmers WF. Dopamine modulates synaptic plasticity in dendrites of rat and human dentate granule cells. Proceedings of the National Academy of Sciences of the United States of America. 2010; 107(42):18185–90. https://doi.org/10.1073/pnas.1011558107 PMID: 20921404

57. de Quervain D, Schwabe L, Roozenendaal B. Stress, glucocorticoids and memory: implications for treating fear-related disorders. Nat Rev Neurosci. 2017; 18(1):7–19. https://doi.org/10.1038/nrr.2016.155 PMID: 27881856.

58. Schwabe L, Joels M, Roozenendaal B, Wolf OT, OtzI MS. Stress effects on memory: an update and integration. Neuroscience and biobehavioral reviews. 2012; 36(7):1740–9. https://doi.org/10.1016/j.neubiorev.2011.07.002 PMID: 21771612.

59. Smeets T, Jelicic M, Merckelbach H. Stress-induced cortisol responses, sex differences, and false recollections in a DRM paradigm. Biological psychology. 2006; 72(2):164–72. https://doi.org/10.1016/j.biopsycho.2005.09.004 PMID: 16289825.

60. Zoladz PR, Peters DM, Kalchik AE, Hoffman MM, Aufdenkampe RL, Woelke SA, et al. Brief, pre-learning stress reduces false memory production and enhances true memory selectively in females. Physiology & behavior. 2014; 128:270–6. https://doi.org/10.1016/j.physbeh.2014.02.028 PMID: 24560841.

61. Pardilla-Delgado E, Alger SE, Cunningham TJ, Kinealy B, Payne JD. Effects of post-encoding stress on performance in the DRM false memory paradigm. Learning & memory. 2016; 23(1):46–50. https://doi.org/10.1101/lm.039354.115 PMID: 26670187

62. LaBar KS, Cabeza R. Cognitive neuroscience of emotional memory. Nat Rev Neurosci. 2006; 7(1):54–64. https://doi.org/10.1038/nrn1825 PMID: 16371950.

63. Schwabe L, Wolf OT. Stress modulates the engagement of multiple memory systems in classification learning. The Journal of neuroscience: the official journal of the Society for Neuroscience. 2012; 32(32):11042–9. https://doi.org/10.1523/JNEUROSCI.1484-12.2012 PMID: 22875937.

64. Stickgold R, Walker MP. Sleep-dependent memory triage: evolving generalization through selective processing. Nature neuroscience. 2013; 16(2):139–45. https://doi.org/10.1038/nn.3303 PMID: 23354387.

65. Gomez RL, Bootzin RR, Nadel L. Naps promote abstraction in language-learning infants. Psychological science. 2006; 17(8):670–4. https://doi.org/10.1111/j.1467-9280.2006.01764.x PMID: 16913948.

66. Hupbach A, Gomez RL, Bootzin RR, Nadel L. Nap-dependent learning in infants. Developmental science. 2009; 12(6):1007–12. https://doi.org/10.1111/j.1467-7687.2009.00837.x PMID: 19840054.

67. Ellenbogen JM, Hu PT, Payne JD, Titone D, Walker MP. Human relational memory requires time and sleep. Proceedings of the National Academy of Sciences of the United States of America. 2007; 104(18):7723–8. https://doi.org/10.1073/pnas.070094104 PMID: 17449637

68. Ramanathan DS, Gulati T, Ganguly K. Sleep-Dependent Reactivation of Ensembles in Motor Cortex Promotes Skill Consolidation. PLoS Biol. 2015; 13(9):e1002263. https://doi.org/10.1371/journal.pbio.1002263 PMID: 26382320

69. Lewis PA, Durrant SJ. Overlapping memory replay during sleep builds cognitive schemata. Trends in cognitive sciences. 2011; 15(8):343–51. https://doi.org/10.1016/j.tics.2011.06.004 PMID: 21764357.
70. Rasch B, Buchel C, Gais S, Born J. Odor cues during slow-wave sleep prompt declarative memory consolidation. Science. 2007; 315(5817):1426–9. https://doi.org/10.1126/science.1138581 PMID: 17347444.

71. Bendor D, Wilson MA. Biasing the content of hippocampal replay during sleep. Nature neuroscience. 2012; 15(10):1439–44. https://doi.org/10.1038/nn.3203 PMID: 2294111

72. Seibold M, Rasch B, Born J, Diekelmann S. Reactivation of interference during sleep does not impair ongoing memory consolidation. Memory. 2017;1–8. https://doi.org/10.1080/09609327.2017.1329442 PMID: 28537468.

73. Massimini M, Ferrarelli F, Huber R, Esser SK, Singh H, Tononi G. Breakdown of cortical effective connectivity during sleep. Science. 2005; 309(5744):2228–32. https://doi.org/10.1126/science.1117256 PMID: 16195466.

74. Massimini M, Ferrarelli F, Esser SK, Riedner BA, Huber R, Murphy M, et al. Triggering sleep slow waves by transcranial magnetic stimulation. Proceedings of the National Academy of Sciences of the United States of America. 2007; 104(20):8496–501. https://doi.org/10.1073/pnas.0702495104 PMID: 17483481.

75. Boly M, Perlberg V, Marrelec G, Schabus M, Laureys S, Doyon J, et al. Hierarchical clustering of brain activity during human nonrapid eye movement sleep. Proceedings of the National Academy of Sciences of the United States of America. 2012; 109(15):5856–61. https://doi.org/10.1073/pnas.111133109 PMID: 22451917.

76. Kaplan R, Adhikari MH, Hindriks R, Mantini D, Murayama Y, Logothetis NK, et al. Hippocampal Sharp-Wave Ripples Influence Selective Activation of the Default Mode Network. Current biology: CB. 2016; 26(5):686–91. https://doi.org/10.1016/j.cub.2016.01.017 PMID: 26898464.

77. Walker MP, Robertson EM. Memory Processing: Ripples in the Resting Brain. Current biology: CB. 2016; 26(6):R239–41. https://doi.org/10.1016/j.cub.2016.02.028 PMID: 27003888.

78. Boutin A, Pinsard B, Bore A, Carrier J, Fogel SM, Doyon J. Transient synchronization of hippocampo-striato-thalamo-cortical networks during sleep spindle oscillations induces motor memory consolidation. NeuroImage. 2017; 169:419–30. https://doi.org/10.1016/j.neuroimage.2017.12.066 PMID: 29277652.