Neuroscience: Memory modification without catastrophe

Mircea van der Plas¹², Alberto Failla¹², and Edwin M. Robertson¹,* ¹Institute of Neuroscience and Psychology, Centre for Cognitive Neuroimaging, University of Glasgow, Glasgow G12 8QB, UK ²Equal contribution *Correspondence: edwin.robertson@glasgow.ac.uk https://doi.org/10.1016/j.cub.2024.02.068

Adaptive behaviour is supported by changes in neuronal networks. Insight into maintaining these memories — preventing their catastrophic loss — despite further network changes occurring due to novel learning is provided in a new study.

Dilemmas abound in biology. One such is between stability and change. A stable genome is required to send desirable characteristics down generations; conversely, a genome also needs to change (genetic mutation) to make evolutionary adaptation possible. Similarly, neuronal networks need to be stable, maintaining their functional properties, to support adaptive behaviours; yet, for example, as environments change, networks must also change for behaviour to adapt to a new environment¹. Failing to solve this dilemma could result in newly acquired behaviours creating such profound change within networks that they catastrophically interfere with existing adaptive behaviours²³. Work described in this issue of Current Biology by Losey et al.⁴ provides fresh insights into how networks solve the challenging dilemma between stability and change.

A new experience guides the formation of a memory. Perhaps the simplest way to prevent a newly formed memory from affecting earlier memories is to have any individual memory allocated to only a tiny fraction of an available network. This minimizes the overlap between the memory networks, limiting interference. Such sparse coding occurs within the hippocampus where potentially as few as twenty neurons are critical for a newly acquired adaptive behaviour (spatial navigation⁵). However, this is not the only solution to catastrophic interference between memories. Memories are prevented from interfering with one another — even when allocated to the same network — because their patterns of activity are independent and do not correlate⁶. In each case, memories are isolated from one another within a particular dimension (space or time, respectively), which prevents any interaction, and also prevents the loss of an adaptive behaviour. Yet, new work shows that isolation may not always be necessary to prevent catastrophic interference.

Actions frequently need to adapt to new environments. For example, the stride change from running on rough terrain to running on sand, and the change in arm movements as an arm is moved through water, rather than air, during swimming. It is possible to mimic these adaptions by manipulating the relationship between the firing of motor cortical neurons, and the movement of a cursor on a screen (Figure 1A). This can be achieved using a brain–computer interface (BCI), which samples brain...
activity and analyses this activity to produce commands, which are used by an output device to carry out desired actions. In the new work from Losey et al., using a BCI a monkey learnt to control the screen position of a cursor by modifying the activity of motor cortical neurons sampled using an array of electrodes. The precise relationship between the firing pattern of the sampled neuronal population and the cursor movement was determined by the BCI settings. Thus, changing the BCI settings required the monkey to adapt how they performed a movement by using a different pattern of neuronal activity. Different environments were mimicked by the settings of the BCI (a or b). The first of these (setting a) determined the relationship between motor cortical neuronal firing and cursor movements. By learning this mapping (map A) monkeys controlled the cursor, which allowed them to capture the targets quickly and accurately (low spatial error). Each monkey achieved close to perfect performance. By contrast, when the second setting (setting b) was introduced performance declined substantially, and many subsequent trials were required to adapt to the new setting. Adaptation to the new setting (setting b) required the monkeys to learn a new mapping between neuronal firing and cursor movements (map B). Overall, different settings on the BCI (a or b) triggered the learning of different relationships between neuronal firing and cursor movement (maps A or B).

Learning different maps may be supported by entirely independent memories with different activity patterns (in space or time). Alternatively, learning one memory may alter another, and so modify the neural activity pattern supporting adaption to the earlier map. To distinguish between these possibilities motor cortical activity was first measured in map A, then subsequently during learning of map B, and finally again within map A. The change between the different maps was not cued, which required the maps to be sufficiently similar to prevent the monkeys' performance becoming so poor that they would no longer engage with the task. The pattern of activity for map A showed a substantial shift after learning map B (Figure 1B). Learning B had modified the activity patterns supporting map A, which suggests that the memory (for map A) was modified. Memories are prone to being modified when there is a lack of distinct cue or boundary between learning episodes1. Yet, in this current case the memory modification occurs without any substantial change in performance (of A; Figure 1B). Thus, the learnt map A was modified without any catastrophic interference in performance (of A).

Changing a memory without modifying performance is perhaps puzzling. Yet, this is not as counterintuitive as it may at first appear. A memory is supported by changes in synaptic weights, altering the transmission of information through a network, to create newly acquired
adaptive behaviours. There is not simply one combination of synaptic weights that produce a specific adaptive behaviour. Instead, many different combinations of synaptic weights can modify a network activity in such a way as to produce the required learnt performance. This is an example of another important biological principle — degeneracy — in which different routes (combination of synaptic weights in a network) can be taken to achieve the same result (a learnt behaviour; see Breton and Robertson for example). Many different configurations of synaptic weights and activity patterns can lead to the same learnt performance. Thus, activity and synaptic weights can change without any change in memory performance.

The activity change may have been shaped by the demands of the novel environment (map B). In this scenario, learning the novel map would constrain the possible synaptic weight changes. This would imply that information relevant to both maps (A and B) are embedded within the modified activity pattern. This was elegantly shown in the new work by identifying activity during map A, which was also relevant to map B. Performance was simulated using activity during map A — either before or after learning the map B — with the unmatched BCI setting (b for map B). Using activity before learning map B led to an impaired simulated performance — compared to actual performance when activity and BCI setting were matched (map A or B; settings a or b, respectively; Figure 1C). By contrast, using activity after learning map B led to an improved simulated performance. The activity supporting adaptive performance for map A is modified by learning map B, and it no longer exclusively supports adaptive performance for one map (map A); instead, it is a multiplexed signal that can support adaptive performance for both environments. Thus, a memory has been modified, and further information incorporated.

Further information can be added because of how activity is related to performance within each of the learnt maps. Some changes in activity lead to a substantial change in performance (Figure 1D). By contrast, other changes of activity have only a slight impact upon performance. When these latter changes occur, because they are relatively unimportant to map A, they allow performance for map A to be maintained, and simultaneously can be exploited for a novel map (map B) between activity and performance. In this way, a memory is modified — its pattern of activity changes and new information assimilated — without disrupting performance of the memory. Memory modification is shown to occur without interference. The absence of interference may arise because the memories contain similar information (map A versus B); whereas interference has been found between very different serial information (of actions to positions 4-1-3-2-4 versus 2-3-1-4-2) and even between very different types of information (actions versus words). Memories protect subsequently formed memories with a common (serial) structure from disruption. Equally, the initial training (in map A) ensured performance was close to optimal (asymptotic), and such training can prevent disruption of a memory (for map A). In these situations, a memory is not catastrophically lost. What the current work shows is how the memory is maintained despite the new experience (map B) modifying the existing memory (map A). This modification creates a multiplexed memory able to meet the adaptive demands of multiple environments. This may be limited to only meeting the demands of previous experiences. Alternatively, the memories may become linked or integrated together to allow adaptive performance even to an entirely novel experience (generalization). For example, the memories for simple pairing rules (that A>B, B>C, C>D) when linked together enable individuals to correctly respond to a pairing that they have never previously experienced (A>D). Whether the memory modification observed in the recent work supports such generalization (or not) remains an open question.

Potentially the generalization is limited to a simple interpolation between the different maps — predicting improved initial performance but not enhanced learning in novel maps (c.f. Mosha et al.; for a review, see Robertson). Alternatively, perhaps the modified memory is more than a multiplex of different maps and contains information allowing adaptive performance in a rich variety of novel environments. Distinguishing between these possibilities will reveal the links among memory modification, catastrophic interference, and generalization.

DECLARATION OF INTERESTS

Edwin Robertson is a member of Current Biology’s Advisory Board.

REFERENCES

Virtually all organisms are exposed to periodic changes in the environment, including daily light/dark cycles, oscillating temperatures, and predators active at specific times of the day. Not surprisingly, the ability to anticipate these changes is advantageous. Therefore, to be ready for upcoming events, the majority of organisms evolved so-called circadian (from Latin, circa — approximately, dies — day) clocks that are used to keep track of time internally by measuring approximately 24 hours. We usually acknowledge this endogenous clock when we experience jetlag after a long journey, as was illustrated by the impact of sociosexual interactions on the circadian clock. In insects, the main goal for adults is reproduction. While in females this task often includes gathering energy and building material via feeding, for males the activity reduces to the required 24 hours. Under constant conditions (such as in the laboratory), the circadian clock usually does not run exactly at 24 hours. This small imperfection is not a major problem in real life because the clock is synchronized (entrained) with external time on a daily basis. The strongest external stimulus is light, which is detected by photoreceptors, including insects, the clock relies on several interlocked transcription/translation feedback loops. Transcription and translation are fast processes that can be completed in the order of minutes. Importantly, the circadian clock machinery includes several regulatory mechanisms involving protein phosphorylation, degradation, protein–protein interactions, and subcellular localization, which all together delay the entire cycle of interconnected transcription/translation feedback loops to the required 24 hours. Under constant conditions (such as in the laboratory), the circadian clock usually does not run exactly at 24 hours. This small imperfection is not a major problem in real life because the clock is synchronized (entrained) with external time on a daily basis. The strongest external stimulus is light, which is detected by photoreceptors such as cryptochromes and opsins. Other signals influence the clock as well. For example, social interaction is particularly profound in some eusocial insects such as honey bees. And a seminal work by Levine et al. used constant darkness to illustrate that social experience can reset the clock by olfactory cues even in the fruit fly Drosophila melanogaster.

A new study by Ghosh et al. reported in this issue of Current Biology explored the impact of sociosexual interactions on the circadian clock. In insects, the main goal for adults is reproduction. While in females this task often includes gathering energy and building material via feeding, for males the activity reduces to successful mating. The authors use communication between adults of the moth Spodoptera littoralis as an experimental model. In this species, and many other Lepidoptera, females release a mixture of volatile chemical substances, so-called pheromones. Pheromones are known to spread effectively, attracting males even from up to several kilometers. The pheromone release is restricted to a specific time of day, in nocturnal moths to a portion of night. While the mating-inducing role of pheromones is well established,