A learning mechanism completed in milliseconds capable of transitioning to stabilizable forms can generate working, short and long-term memories - A verifiable mechanism

Kunjunmon I. Vadakkan
Neurosearch Center, 76 Henry Street, Toronto, ON M5T1X2
November 8, 2018

Abstract

Multiple associative learning events can take place within sub-second time and the “completed” mechanism can then be used for specific memory retrieval without any lapse of time. This indicates that a biological process is completed within the matching time-scales of milliseconds that can be used for retrieving specific memory. Since qualia of working, short-term and long-term memories are same except for degradation of features in long-term memory and since every long-term memory had the capability to induce working memory immediately after learning, all memories are anticipated to get induced from a mechanism formed at the time of learning. When memories are viewed as first-person internal sensations, a derived mechanism fulfills the “completion” requirement within milliseconds that can be used to induce working memory and can be transitioned to stabilizable forms to induce short-term and long-term memories.

1 Introduction

Higher brain functions such as learning, memory, and thought processes are first-person internal sensations. However, it is not possible for a third-person observer to sense these functions using direct methods. In this context, current investigations are limited to studies of third-person observations at different levels such as biochemical, cellular, electrophysiological, systems, behavior, and imaging. For example, current studies examine biochemical changes occurring at the level of transcription of DNA, translation of mRNA, post-translational modifications of proteins, and functional aspects of different proteins. These functions are then correlated with motor actions responsible for either behavior or speech to make correlations. Even though first-person internal sensations cannot be directly examined, indirect methods can be applied towards understanding its operational mechanism. In this context, present work explains the importance of examining changes taking place at time-scales equal to or faster than first-person reports or external manifestations of inner sensations of memory from a completed learning mechanism.

Among different brain functions, memory has the advantage that experiments can be carried out to associatively teach the system, examine changes generated by learning and then explore
how these changes are used during memory retrieval. Even though pioneering thoughts about the importance of understanding the inner sensations were made by Wilhelm Wundt (Wundt, 1874), there were no methods available to make an approach towards it until recently. Therefore, we have been assessing memory using motor changes such as speech or behavior occurring at the time of memory retrieval. In this line of approach, at one point, several findings led to the view that memory storage requires a “tritrace” system (McGaugh, 1966) based on experimental results using protein synthesis inhibitors (Flexner et al., 1963; Agranoff et al., 1965; Deutsch et al., 1966). In tests to measure capacity to hold information that were then used to support different abilities such as reasoning and motor activities, short-term memory was distinguished from long-term memory (Waugh and Norman, 1965; Baddeley and Warrington, 1970). Later, the concept of working memory was introduced to explain temporary information storage that can perform a wide range of complex cognitive tasks (Baddeley and Hitch, 1974; Baddeley, 2003). Most of the associative learning changes that last only for a short duration of seconds were classified as working memory and those that last for comparatively long durations were viewed as long-term memories. Ability to block long-term memories using protein synthesis inhibitors provided the idea that long-term memory formation dependent on protein synthesis may have a different mechanism than that of working memory. Due to continued lack of methods to explore the non-sensible nature of first-person internal sensations of memory, the classification of memories based on behavioral findings continued. In a perspective on the man-made origin of different types of memories, it was stressed that separating short-term and long-term memories is an example of classification by dissociation (Tulving, 1987).

Studies of amnesic patients with medial temporal lobe damage showed that their working memory were intact, even though those patients had markedly impaired performance on tasks requiring long-term memory (Drachman and Arbit, 1966; Milner, 1972). This also matched with the view that working memory is independent of medial temporal lobe structures (Milner, 1972). It is to be noted that interpretations from these experimental results did not rule out the possibility whether the mechanism of working memory at the medial temporal lobe can progress towards long-term memories. In other words, these studies did not provide any clues to spark the questions whether a) the mechanism of working memory can be maintained for long duration of time for long-term memory, and b) if can be maintained, how does it depend on protein synthesis.

Both the short lasting feature of working memory and the protein synthesis dependence of long-term maintenance of learning-induced changes maintained the view that cellular mechanism for working memory is different from long-term memories. This was reinforced by long-term potentiation (LTP), an electrophysiological finding that has shown several correlations with the ability to learn and retrieve memory by examining behavior of animals. Since routine stimulation protocols can induce LTP only after a delay of 20 to 30 seconds (Gustafsson and Wigström, 1990) or even after a delay of more than a minute (Escobar and Derrick, 2007), LTP was primarily correlated with long-term memory. Delay of induction of LTP roughly matches with the time-scales at which protein synthesis associated with long-term memory takes place. Furthermore, protein synthesis inhibition was shown to block LTP both in slices and in vivo (Stanton and Sarvery 1984; Krug et al., 1984). In summary, several observations allowed to maintain the idea that mechanisms of working and long-term memories are different.

In the broader context that we have not yet understood how the brain operates, it was necessary to re-examine the previous conclusions. In the renewed approach, it is necessary to keep replica-
tion of the mechanism in an engineered system as the gold standard proof of understanding it. Even though, this is not immediately achievable, it is what is required to properly understand the operations of the system. This necessitates us to take several important steps. First, it is necessary to view memories as first-person internal sensation within the system because that is what is truly happening within the nervous system and it is what we expect to happen within the engineered system. When this approach is used, it can be seen that qualia (features of internal sensations) of working, short-term and long-term memories are same except that long-term memories are likely to lose clarity of sensory features of retrieved memory. Secondly, every long-term memory had the capability to induce working memory immediately after learning, indicating that all memories can get induced from the same mechanism formed at the time of learning. With this much insight, it is necessary to search for a very fast learning-mechanism from which internal sensations can be induced. It should be capable of both reversing back quickly and at the same time have the capability to transition to more stable structural changes for short-term and long-term memories. The stabilized changes should have the capability to reverse back if not used or reactivated by related learning or memory retrieval events having some shared features. The learning-mechanism should also have provision to get augmented by motivation.

Viewing memories are first-person internal sensations raises the issue of access problem to empirically test the formation of first-person properties by third-person observers. Since both mathematics and physics have developed methods to access items that are not ordinarily accessible to our sensory systems, we can use the principle behind their approaches to understand the unknown mechanism of induction of internal sensations. Based on the deep underlying principle used in the above basic science fields, a non-sensible unknown mechanism within a system can be derived if we have information about all other features of that system. In other words, it is possible to hypothesise a mechanism of learning from which internal sensation of memory can be induced, provided that mechanism can interconnect all the findings of that system at different levels. Following this, both retrodiction and prediction can be carried out to verify the derived mechanism.

2 Time-scale matched “completion” requirement of the mechanism

Humans have the ability to associatively learn rapidly and retrieve memories from learning-induced changes in the presence of a cue stimulus. This can be demonstrated by the ability to learn multiple associations during a rapid fire series of associative visual learning tasks that can be completed within a second. Ability to associate more than one pair of associative learning tasks and retrieve specific memory in response to a specific cue stimulus indicates that the biological mechanism of learning is “completed” at corresponding or faster time-scales of milliseconds. The completed mechanism is necessary for a specific cue stimulus to induce internal sensation of memory of associatively learned item. If this completed minimum mechanism is not subjected to any stabilization process, it will reverse back quickly. The ability to retrieve memories during this short period of time before its retrieval can explain working memory. Whatever is the mechanism for working memory, it is merely holding the association of between two learned stimuli very briefly. Such a transient mechanism should become evident when the actual mechanism is discovered. If learning mechanism can be stabilized, then it should be able to persist for long period of time responsible for long-term memories. Here, one should note that when net memory results from an integral of
Figure 1: Comparison between times required for generation of a “completed” mechanism necessary for long-term memory that takes only milliseconds (in red, on the left end) from which memory can be retrieved and that based on current investigations that takes minutes (entire length of the line). Lengths of lines (red ≈ 50 milliseconds, red + blue ≈ 1 minute) are only for comparison and not to scale. Huge mismatch between the durations indicates that current methods are missing the completed mechanism occurring at the time of learning.

large number of operational units, then presence of surplus number of units permits losing certain number of units without compromising net memory. In summary, a learning mechanism needs to fulfill a “completion” requirement of milliseconds and it should be directly responsible for working memory and deeply ingrained within the mechanism for long-term memories. This necessitates the mechanism of learning capable of inducing memories to be completed much faster than what is being studied currently by a factor of $10^3$ (Figure 1).

Learning-induced changes can be qualified as fast (in milliseconds), complete (completed so that it can allow memory retrieval), reversible, stabilizable (for different durations for short-term and long-term memory retrievals) and re-activatable (passive re-activation for memory retrieval). This can be explained in terms of the working of a separating zipper of a jacket (Figure 2). Inserting the insertion pin into the retainer box at the bottom of the jacket connects the two zipper chains. This initial step is similar to the “completed” learning-induced change in milliseconds. During all further actions of zipping upwards towards the neck, attachment between the zipper chains remains. All the delayed processes taking place after the “completed” minimum required changes after learning are likely mechanisms to a) stabilize the learning change, b) replenish used up molecules, and c) initiate downstream cascades of changes triggered by original learning-induced change that took place at time-scales of milliseconds. Examining individual delayed changes following learning is unlikely to provide information regarding the learning-induced changes.

3 Limitations of current methods of studies

1) Since third-person observers cannot sense first-person internal sensations of retrieved memories, investigations have been limited to third-person observable findings. For example, higher brain functions are being assessed by examining motor activities such as behaviour and speech.

2) Manipulations of proteins by over-expression, blocking of expression and recombinant expression under desired conditions have been carried out to alter cellular changes occurring at time-scales much slower than the time-scales at which mechanism of learning takes place. Results from these experiments have been used to assess higher brain functions. For example, manipulations of expression of immediate early genes (IEGs) *arc* (activity-regulated cytoskeleton-associated protein) *c-fos* (fos proto-oncogene), *zif268* (zinc finger protein 268), and *egr-1* (early growth response-1) have been used to assess different higher brain functions. These findings independently do not provide any indirect means to solve the system. Even though IEG expression is the earliest third-person observation that can be utilized to manipulate the cell function, this delayed event does not allow us to understand the mechanism occurring at time-scales of milliseconds. Even though mRNA for IEGs may appear in cells within minutes after stimulation (Bahrami and Drabløs, 2016), it
Figure 2: A comparison of the “completed” learning-induced change with that of the working of a separating zipper of a jacket. The “completed” mechanism during learning can be compared to that of the step of inserting the insertion pin ($I_p$) into the retainer box ($R$) at the bottom of a jacket. This will connect two zipper chains. Say for example, if the zippers are made of metal, the initial step will be sufficient to make an electrical connection between the zipper chains to close an electric circuit to light a small bulb at the front side of a jacket (a nice design for a night jacket!). This connection will remain when the pull tab is pulled all the way up and lock it. It will even remain intact during the unzipping stage until the insertion pin is removed from the retainer box. Similarly, completed learning mechanism taking place in milliseconds is sufficient to induce internal sensation of memory of the same qualia for any length of time that it can be maintained to persist. Note that inserting insertion pin into retainer box only for a short period of time before taking it out can be compared to the short-lasting learning mechanism during which working memory is retrieved.

is occurring at time-scales that differ by the order of $10^3$ than what is expected to be completed in milliseconds of time at the time of learning. Studies that use dendritic targeting elements of arc mRNA (Koybayashi et al., 2005; Hayashi-Takagi et al., 2015) use expression of a protein that takes place in time-scales of minutes and it does not use a “completion” criterion. Furthermore, it does not explain how specificity required of a mechanism is implemented when that mechanism is used.

Studies that use expression of IEGs to control neuronal firing and make correlations with various higher brain functions are based on the observation that neuronal firing triggers expression of those proteins. Since there is huge redundancy of inputs in firing a neuron (Vadakkan, 2018), expression of IEGs can occur by activation of a gigantic number of very small subsets of combinations of inputs arriving at the dendritic spines of a neuron. This non-specificity in expression of IEGs further confounds the problem of not examining a learning mechanism occurring at appropriate time-scales.
4 Necessary features of minimum “completed” mechanism

Associative learning events capable of generating long-term memories are also capable of generating working memory of the same qualia immediately following learning of those associative learning events. This indicates the possibility that mechanism occurring at the time of learning can be used to retrieve internal sensation of memory at different time intervals after learning. This mechanism should be able to explain how majority of learning-induced changes reverse back within a few seconds, responsible for short life-span of working memory. The learning-induced change while continuing to generate internal sensation of memories, is also expected to get stabilized for varying duration of time for generating short-term and long-term memories. All the above features show that working, short-term and long-term memories can have the same “completed” learning-induced change and that it can be used to induce same internal sensation of memory at different times, if the learning-induced change can be maintained. When net memory results from an integral of large number of operational units, then the strength of memory depends on the number of such units available at a given time. In summary, examining memories in their true nature as first-person internal sensations occurring at physiological time-scales of milliseconds provides an opportunity to understand the operational mechanism.

5 A time-scale matched cellular mechanism generated at the time of learning

Since both learning and memory retrieval occur in milliseconds, their mechanisms should take place close to the time-scales at which depolarization spread occurs. Using both a) method of logical derivation (Vadakkan, 2007, 2013), and b) principle of methods used in linear algebra to find a unique solution that binds all the linear equations within a system (Vadakkan 2016b), it was possible to arrive at a cellular change that can take place at physiological time-scales during learning that has a unique potential mechanism residing within it to induce internal sensation of memory. Accordingly, a time-scale matched interaction at the level of converging inputs between dendritic spines that belong to different neurons, namely inter-postsynaptic functional LINK (IPL) was derived as a mechanism occurring at the time of learning. IPL can progress through different stages (Figure 3). The initial stage consists of overcoming the energy barriers between the membranes (Martens and McMahon, 2008) of interacting dendritic spine heads, which can enable propagation of potentials between the spines within the physiological time-scale of milliseconds. This initial stage is expected to enable one of the associatively learned stimuli to propagate across this IPL and induce units of internal sensations at the inter-LINKed spine (Vadakkan, 2013, 2016a). This non-stabilized initial stage of IPL formation will be sufficient to achieve the “completion” requirement expected from a mechanism of learning.

At baseline state, extracellular matrix space between neuronal processes is very thin allowing presence of a microenvironment similar to that of cerebrospinal fluid with proteoglycans (Nicholson and Syková, 1998) to separate the membranes preventing depolarization crossing between the spines. Using artificial membranes, it was shown that very high energy is required to overcome a high energy barrier between membranes to initiate membrane fusion (Cohen and Melikyan, 2004). Since the repulsive hydration force increases steeply as the distance between the two bilayers falls below 20Å, fusion between two membranes is a high energy requiring process (Rand and Parsegian, 1984; Harrison, 2015). Using elastic models, it was found that the high curvature of
Figure 3: Potential mechanism for minimum “completed” change during learning and its further stabilization. A) Two abutted synapses AB and CD. Presynaptic terminals A and C are shown with synaptic vesicles (in blue color). Postsynaptic terminals (dendritic spines or spines) B and D have membrane-bound vesicles containing subunits of AMPA receptor inside them. Note the presence of a hydrophilic region (in green) separating spines B and D, which prevents interaction between them. Very high energy is required for bringing together repulsive membranes and overcome energy barriers (Martens and McMahon, 2008) between the postsynaptic membranes. B) Hydration exclusion and overcoming of repulsive membrane charges are the changes necessary to occur during learning that will enable memory retrieval (see Vadakkan, 2013). Simultaneous activation of spines B and D results in membrane reorganization that leads these changes to occur allowing close contact between their membranes at this region. Since high energy is required to overcome the membrane repulsive factors, this interaction reverses back very quickly. Therefore, this forms only a transient inter-postsynaptic LINK (IPL) that lasts for a short period of time. This is similar to the example (in Fig.2) of insertion of the insertion pin into the retainer box only for a short period of time before taking it out. Factors such as dopamine released to the spines can lead to latter’s expansion can promote IPL formation. During this short period of time, a cue stimulus-generated action potential arriving at synapse AB can reactivate IPL B-D, propagate to spine D and induce units of internal sensation at inter-LINKed spine D. This can explain working memory. C) Formation of a partial inter-spine hemifusion. AMPA receptor subunits are located at the extra-synaptic locations extending up to 25nm beyond the synaptic specialization (Jacob and Weinberg, 2014). This indicates that activity arriving at the synapse can lead to exocytosis of vesicles containing AMPA GluA1 receptor-subunits abutted to the cell membranes of the spines at physiological time-scales. During exocytosis, vesicle membrane segments reorganize the spine membrane by getting incorporated into it. This can facilitate inter-spine partial membrane hemifusion. D) Stage of partial hemifusion progresses to complete hemifusion. Some of the hemifusion changes can get stabilized for different lengths of time and are responsible for long-term memory (Figure modified from Vadakkan, 2016a).

Intermediates during fusion process can create void spaces that require high activation energies (Siegel, 1993). Later studies have shown that several factors are involved in determining the properties of hydration water between the membranes (Disalvo et al., 2008; Song et al., 2014; Dreier et al., 2018). The high energy barrier for membrane fusion can be overcome by the action of certain specific proteins (Kozlovsky et al., 2004; Martens and McMahon, 2008). SNARE (soluble NSF (N-ethylmaleimide sensitive fusion protein) attachment protein receptor) proteins are known to provide energy for bringing together repulsive membrane charges and overcome energy barriers related to curvature deformations during hemifusion formation between the abutted membranes (Martens and McMahon, 2008; Oelkers et al., 2016). SNARE proteins also generate force for pulling
the abutted membranes together as tightly as possible (Hernandez et al., 2012). These factors can in turn cause inter-spine membrane interaction allowing propagation of potentials across them responsible for inducing working memory. Along with this, additional potentials arriving at the inter-LINKed spines are expected to cause firing of additional neurons in response to one of the stimuli following learning (in comparison with before learning) (Vadakkan, 2013, 2016a, 2018). Since high energy is required for this process, it is expected to persist only for a short period of time of few seconds.

6 Learning-generated mechanism transitions to stabilizable yet reversible states

For a learning-generated mechanism to remain for long period of time, it is necessary for the IPLs formed by hydration exclusion to be maintained for long period of time. At the same time, these mechanisms should also be reversible. One method is that membranes undergo further interaction that can lead to their fusion. Hemifusion is a reversible intermediate stage of fusion process. Such a change is possible due to the fact that hemifusion intermediates are characteristic of SNARE proteins, including that of neuronal SNAREs (Lu et al., 2005; Liu et al., 2008). SNARE proteins are known to mediate fusion of vesicles containing AMPARs with the spine membrane (Lu et al., 2001; Kennedy et al., 2010). Furthermore, SNARE proteins are important for supplying energy for fusion (Jahn and Scheller, 2006).

When two abutted spines are activated simultaneously, GluR1 subunit containing vesicle exocytosis leads to incorporation of their vesicle membrane segments in the lateral spine head region and cause membrane reorganization at those locations. Presence of AMPA GluR1 subunits at the lateral spine head region up to 25 nm away from the synaptic cleft border (Jacob and Weinberg, 2014) indicates that this is the most probable region where GluR1 subunit vesicle exocytosis takes place. This finding makes the lateral spine head region the most probable location for inter-spine hemifusion. Similar to the presence of hemifused synaptic vesicles at the presynaptic terminals of central nervous system synapses (Zampighi et al., 2006; Zampighi et al., 2011), docked vesicles can be expected to be present at the lateral aspect of spine head in the resting state. This can facilitate spontaneous vesicle exocytosis when activity arrives at this location. Vesicle exocytosis reorganizing the lateral spine head region along with the action of SNARE protein overcome the energy barrier between the abutted spines of different neurons that are simultaneously activated and facilitate inter-spine hemifusion. This feature that locations of exocytosis are favorable locations of inter-membrane fusion is a common finding. For example, multiple intracellular vesicle fusion with the cell membrane of sperm precedes inter-cellular membrane fusion with the egg (Wassarman and Litscher, 2008). While still maintaining the functions of IPL propagating depolarization across it, the inter-neuronal inter-spine interaction can progress towards the stages of reversible partial and complete hemifusion between the lipid bilayers of interacting spine heads (Figure 3).

Since small areas will be preferred for hemifusion due to the need for concentrating the ionic channels for flow of ions across the membranes, area of hemifusion is likely to restrict to approximately 10nm² as observed by studies using artificial membranes (Leikin et al., 1987). It is possible that repeated learning events that result in interaction between same pair of spines can lead to homeostatic changes that promote stabilization of different stages of this inter-spine interaction for varying durations. All the stages of interaction are reversible, once their stabilizing mechanisms are
removed. All along various stages of hemifusion, stabilization, and its reversal, the IPL function is expected to remain uninterrupted since it allows spread of potentials between the inter-LINKed spines across the IPL (Figure 2).

7 Neuronal firing during memory retrieval

7.1 Sustained firing of neurons when working memory is held

Large number of experiments have shown that during the delay period following learning there is stimulus-specific increase in spike activity of neurons during which working memory is held. Different experiments have shown this pattern of firing of neurons in prefrontal, temporal, parietal, auditory, visual, somatosensory and medial premotor cortices when working memory is held (Zylberberg and Strowbridge, 2017). Hydration exclusion between stimulated abutted spines during learning can allow propagation of depolarization between those spines and maintain this transient IPL for few seconds. Arrival of potentials across the IPL can allow several sub-threshold activated neurons to cross the threshold and result in their firing (Vadakkan, 2013, 2016a) while working memory is held.

7.2 Firing of additional neurons during retrieval of fear memory

It is known that additional neurons in lateral amygdala are fired when one of the associatively learned stimuli (cue stimulus) arrives after learning compared to its arrival before learning in fear conditioning experiments (Schoenbaum et al., 1999; Tye et al., 2008). This indicates that some learning-induced change is allowing potential generated by the cue stimulus to reach several sub-threshold activated neurons that were not firing before learning in the presence of that cue stimulus. This is possible to occur only by the generation of a path through which depolarization can propagate. In one study, it was found that fear conditioning drives AMPA (alpha-amino-3-hydroxy-5-methyl-isoxazole propionic acid) subtype of glutamate receptors (AMPARs) into the synapse of a large fraction of postsynaptic neurons in the lateral amygdala (Rumpel et al., 2005). This study also found that memory was reduced if synaptic incorporation of AMPARs was blocked in as few as 10 to 20% of lateral amygdala neurons. It is necessary to explain how a learning mechanism associated with AMPAR trafficking can generate an additional path through which depolarization can propagate to several sub-threshold activated neurons. It is also necessary to identify their locations and examine whether depolarization spreading through that path can explain a mechanism that can generate internal sensation of memory and concurrent behavioral motor activity. In this context, IPL formation by membrane hemifusion explained in section 6 provides a verifiable mechanism.

8 Conclusion

Animals depend on internal sensations of memories generated within their nervous systems at physiological time-scales of milliseconds to respond to threats and to obtain food in a predator-prey environment. It will be possible to understand the correct operational mechanism only by adhering to changes occurring at matching time-scales. Current investigations of learning examine changes such as immediate early gene expression that occur at time-scales slower than the actual learning mechanism by a factor of at least $10^3$. This leads to ignoring the actual mechanism of
learning. When memories were viewed as first-person internal sensations, it was possible to derive a mechanism of learning occurring at physiological time-scales of milliseconds that can then provide a feasible mechanism for inducing first-person internal sensations of working, short-term and long-term memories, also at time-scales of milliseconds. Future efforts to verify this finding can help to speed up our investigations in understanding the operational principle of the nervous system.

Acknowledgements: This work was supported by Neurosearch Center, Toronto.

Conflict of interest: U.S. patent 9477924 pertains to an electronic circuit model of the interpostsynaptic functional LINK.

References

Gustafsson, B, Wigström H (1990) Long-term potentiation in the hippocampal CA1 region,
https://doi.org/10.1016/S0079-6123(08)61252-2

https://doi.org/10.1016/j.virol.2015.03.043


https://doi.org/10.1016/j.cell.2010.02.042


https://doi.org/10.1016/0361-9230(84)90005-4

https://doi.org/10.1016/0022-5193(87)90021-8

https://doi.org/10.1529/biophysj.107.107896

https://doi.org/10.1016/S0896-6273(01)00194-5

https://doi.org/10.1074/jbc.M506862200

https://doi.org/10.1038/nrm2417


