Artificial intelligence insights into hippocampal processing

Hannah S. Wirtshafter[1,2,3]* and Matthew A. Wilson[1,2,3,4]

[1]Department of Biology, Massachusetts Institute of Technology, Cambridge, MA, United States, [2]Picower Institute for Learning and Memory, Cambridge, MA, United States, [3]Massachusetts Institute of Technology, Cambridge, MA, United States, [4]Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA, United States

Advances in artificial intelligence, machine learning, and deep neural networks have led to new discoveries in human and animal learning and intelligence. A recent artificial intelligence agent in the DeepMind family, muZero, can complete a variety of tasks with limited information about the world in which it is operating and with high uncertainty about features of current and future space. To perform, muZero uses only three functions that are general yet specific enough to allow learning across a variety of tasks without overgeneralization across different contexts. Similarly, humans and animals are able to learn and improve in complex environments while transferring learning from other contexts and without overgeneralizing. In particular, the mammalian extrahippocampal system (eHPCS) can guide spatial decision making while simultaneously encoding and processing spatial and contextual information. Like muZero, the eHPCS is also able to adjust contextual representations depending on the degree and significance of environmental changes and environmental cues. In this opinion, we will argue that the muZero functions parallel those of the hippocampal system. We will show that the different components of the muZero model provide a framework for thinking about generalizable learning in the eHPCS, and that the evaluation of how transitions in cell representations occur between similar and distinct contexts can be informed by advances in artificial intelligence agents such as muZero. We additionally explain how advances in AI agents will provide frameworks and predictions by which to investigate the expected link between state changes and neuronal firing. Specifically, we will discuss testable predictions about the eHPCS, including the functions of replay and remapping, informed by the mechanisms behind muZero learning. We conclude with additional ways in which agents such as muZero can aid in illuminating prospective questions about neural functioning, as well as how these agents may shed light on potential expected answers.

KEYWORDS

hippocampus, replay, navigation, learning, context, DeepMind, muZero, model-based learning

Introduction

A goal of artificial intelligence research is to design programs that can solve problems as well as, or better than humans. Algorithms such as AlphaGo and MuZero can compete with and beat the best trained humans at tasks such as Chess, Go, and Atari (Silver et al., 2016; Schrittwieser et al., 2020). Other artificial intelligence programs can do image classification at accuracy on par with and speeds surpassing humans (Krizhevsky et al., 2012). Ideally, advances in artificial intelligence and machine learning would transfer to understanding human intelligence. However, this often fails to be the case as machines and humans use different mechanisms to learn. For instance, while a supervised image classifier can only distinguish between an elephant and a dog after being presented with 1,000s of images, a child can distinguish between these animals after only a few stimulus exposures (Krizhevsky et al., 2012; Zador, 2019).

However, advances in artificial intelligence, machine learning, and deep neural networks have also led to new discoveries in human and animal learning and intelligence (Richards et al., 2019). Notably, artificial intelligence research has proposed cellular and systems level mechanisms of reinforcement learning (Wang et al., 2018; Dabney et al., 2020), auditory processing (Kell et al., 2018; Jackson et al., 2021), object recognition (Cichy et al., 2016), and spatial navigation (Banino et al., 2018; Bermudez-Contreras et al., 2020; Lian and Burkitt, 2021). In several of these studies, artificial intelligence showed that specific inputs or parameters, such as grid cell coding, were sufficient to create a representation of the environment in which performance on a task closely resembles or surpasses human ability (Banino et al., 2018; Kell et al., 2018).

A recent artificial intelligence agent in the DeepMind family, muZero, can complete a variety of tasks with limited information about the world in which it is operating and with high uncertainty about features of current and future space (Schrittwieser et al., 2020). For instance, unlike in chess, where each move results in a predictable board layout, Atari games (as well as real-world scenarios) involve complex environments with unknown action consequences. MuZero can perform impressively in these environments using only three functions. These functions first encode the environment, then compute potential actions and rewards, and then finally predict an estimation of the future environment. These three functions are both general and specific enough to allow learning across a variety of tasks without overgeneralization across different contexts.

Like muZero, humans and animals can learn context dependent and independent task schema. Decisions by humans and animals are also made in complex spaces and result in unknown consequences, and, like muZero, humans and animals are still able to learn and improve in these environments. In particular, the mammalian extrahippocampal system (eHPCS)

can guide spatial decision making while simultaneously encoding and processing spatial and contextual information (O'Keefe, 1976; Penick and Solomon, 1991; Hafting et al., 2005; Moser et al., 2008; Wirtshafter and Wilson, 2021). The eHPCS is also able to adjust contextual representations depending on the degree and significance of environmental changes and the significance of environmental cues (Hollup et al., 2001; Leutgeb et al., 2005; Lee et al., 2006; Fyhn et al., 2007; Wikenheiser and Redish, 2015b; Boccara et al., 2019; Wirtshafter and Wilson, 2020). In this opinion, we will argue that the muZero functions parallel those of the hippocampal system. We will show that the different components of the muZero model provide a framework for thinking about generalizable learning in the eHPCs.

Parallels in muZero with hippocampal function can help explain mechanisms of context independent learning

The ability to generalize learning to unknown contexts is a key component of human and animal intelligence. For instance, a driver is able to generalize that a red light means stop, regardless of the location of the intersection, the car being driven, or the time of day. While the neural mechanisms behind generalization remain largely unknown, most artificial intelligence models generalize using either model-free or model-based learning. Model-free systems, which learn through experience, are inflexible, fast to run but slow to adapt to change, and estimate the best action based on cached values. Conversely, model-based systems, which aim to learn a model of the environment, are flexible but inefficient, highly sensitive to details of task representations, and easily adaptive to environmental changes (Stachenfeld et al., 2017; Gershman, 2018; Vértes and Sahani, 2019). Although model-based, muZero improves on many shortcomings of model-based and model-free systems by modeling only the elements of the environment important for decision making: the value (the strength of the correct position), the policy (the best action to take), and the reward (the benefit of the previous action). It models these three components using a representation function, a dynamics function, and a predictions function (Schrittwieser et al., 2020). Below, we argue these functions parallel functions found in the hippocampal system, and can be used to provide further insight into and build on current theories of hippocampal function during learning generalization.

The representation function

The first step of muZero processing is to encode the actual environment to a hidden state using a representation function

(Schrittwieser et al., 2020). This hidden state is a rendition of the environment that contains crucial information for performing the task. Similarly, several structures in the eHPCS flexibly represent environmental cues, with some regions placing emphasis on cues and locations of significance. Specifically, the entorhinal cortex (EC), CA regions of the hippocampus (HPC), and the lateral septum (LS) all contain cells whose firing rate is modulated by the location of the animal (termed grid cells, place cells, and "place-like" cells, respectively) (O'Keefe, 1976; Hafting et al., 2005; Moser et al., 2008; Wirtshafter and Wilson, 2019). Spatial representations in all three regions can be restructured by the presence of reward or cues/landmarks of interest, in which cells show a bias for coding for these significant locations (Hollup et al., 2001; Lee et al., 2006; Boccara et al., 2019; Wirtshafter and Wilson, 2020). When there are large changes to the environment, HPC and EC cell populations change representations and are said to "remap" (Fyhn et al., 2007; Colgin et al., 2008). It is likely that this phenomenon also occurs in the LS but it has not been studied.

Selectively coding and over-representing causal features of the environment is important for generalization of learning across contexts (Collins and Frank, 2016; Lehnert et al., 2020). The refinement of this coding allows the inference of critical salient environmental features that may allow or preclude generalization to other contexts—if the salient features are shared among two contexts, learning may be generalizable between the contexts (Collins and Frank, 2016; Abel et al., 2018; Lehnert et al., 2020). This muZero processing parallels the theory that the hippocampus can generalize across different contexts during learning by treating states that have equivalent actions or rewards as equivalent states. By compressing environmental representations into lower dimensionality abstractions, learning can be sped up and more easily transferred between environments with shared features (Lehnert et al., 2020). As such, the representation function provides a means by which learning can be generalized through contextual learning (Figure 1).

The dynamics and prediction functions and the Monte Carlo tree search algorithm

In muZero, a simulation always begins with the representation of the current state (environment), as created using the representation function. From this initial state, a number of hypothetical new states and their corresponding potential actions and rewards are derived using the dynamics and prediction functions, which both rely on the use of a Monte Carlo tree search (MCTS) algorithm.

A MCTS allows a guided search of possible future actions. In brief, at a certain stage of the environment (such as a position in a game), subsequent possible actions and the resulting rewards are evaluated. Using these potential actions and rewards, a new

hypothetical internal state is created and evaluated. Potential future states of *this* states are then considered and evaluated. This process occurs iteratively, forming expansive nodes off of the actual initial state, to allow for deeper search to evaluate more distant outcomes of potential actions. After expansion, the computed rewards are back-propagated up the tree to the current state and the mean value of an expansion is computed. Only then does the agent choose the optimal move or action and progress to the next state.
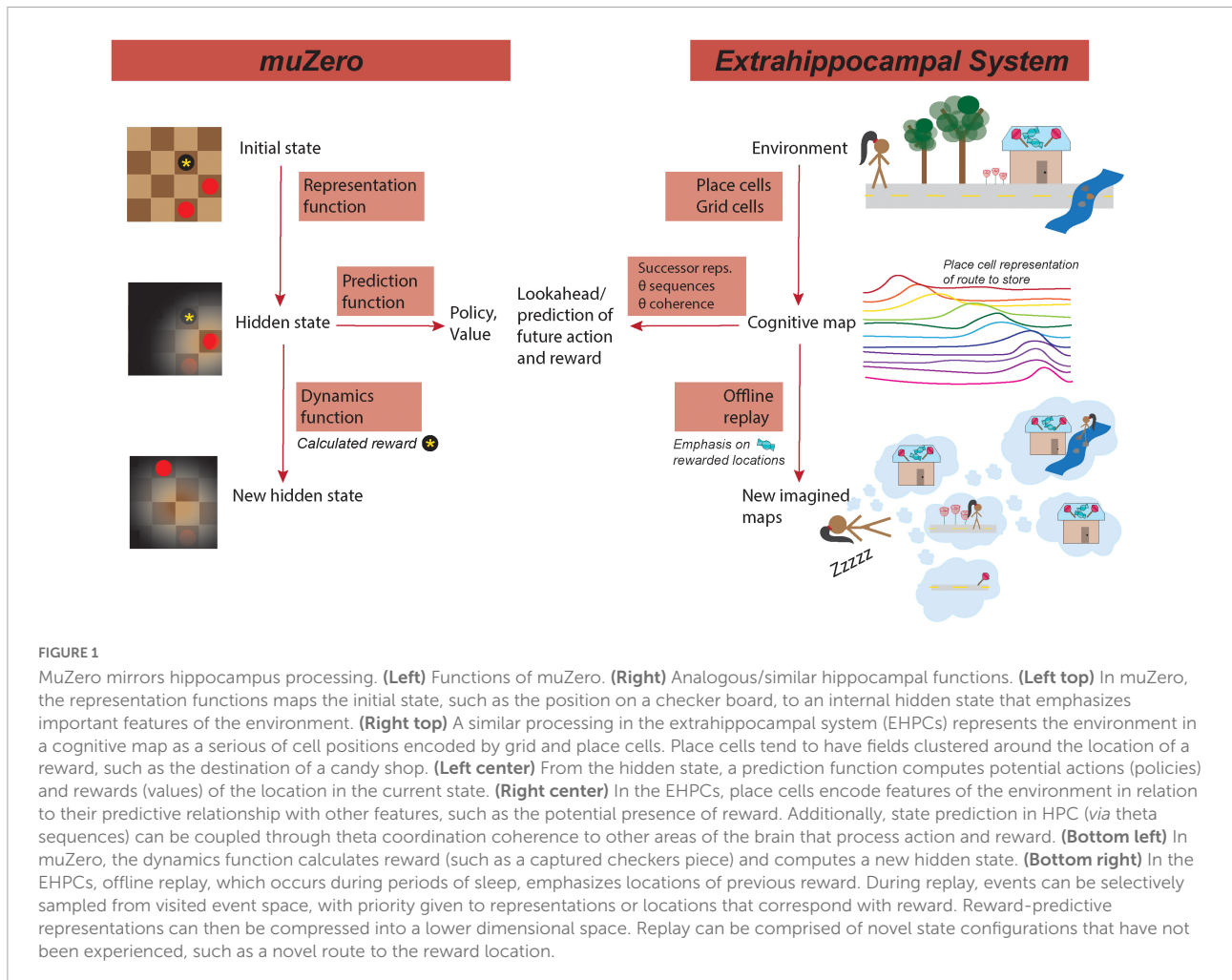
Unlike AlphaGo, in which the rules of the game are known, MuZero learns the mechanics of the game and environment, and can relatively reliably only expand nodes that will likely result in rewarding actions. This expansion is first done using the dynamics function, progressive states are mapped, in a reoccurring process, based on different possible action choices. At each hypothetical state, the function computes a reward and a potential new internal hidden state (the future environment). Both the reward and the new state are computed based on the previous hidden state (Schrittwieser et al., 2020).

Following the creation of each new state in muZero, a prediction function computes potential actions (a policy) and rewards (value) (Schrittwieser et al., 2020). The policy represents how the agent might act in this hidden state, while the value is an estimate of future rewards based on the current state. The value function is updated at each iteration of the dynamics function, as discussed.

This iterative process, of predicting and evaluating future actions in an environment has multiple corollaries in hippocampal processing: predictive place cell coding (the "successor representation"), theta sequences, and replay (Figure 1).

Place cells and the successor representation

In an animal model, it is possible that place cell coding plays a predictive role in cognition: in the "successor representation" (SR) view of place cells, place cells encode features of the environment in relation to their predictive relationship with other features, including the potential presence of reward or potential future trajectories (Stachenfeld et al., 2017; Gershman, 2018). The SR view of place and grid cells explain the aforementioned bias toward significant features in coding. This view would also explain why cells in the HPC and EC maintain stable environmental representations in environments that vary in small ways that do not change the predictive value of contextual clues. In noisy or partially observable environments, models of SR place cells function better than pure cognitive map place cell models. Additionally, in the SR model, state relationships can be learned latently, while reward values are learned locally, allowing for fast updating of changes in reward structure in different contexts (Stachenfeld et al., 2017; Gershman, 2018; Vértes and Sahani, 2019).

FIGURE 1
MuZero mirrors hippocampus processing. **(Left)** Functions of muZero. **(Right)** Analogous/similar hippocampal functions. **(Left top)** In muZero, the representation functions maps the initial state, such as the position on a checker board, to an internal hidden state that emphasizes important features of the environment. **(Right top)** A similar processing in the extrahippocampal system (EHPCs) represents the environment in a cognitive map as a serious of cell positions encoded by grid and place cells. Place cells tend to have fields clustered around the location of a reward, such as the destination of a candy shop. **(Left center)** From the hidden state, a prediction function computes potential actions (policies) and rewards (values) of the location in the current state. **(Right center)** In the EHPCs, place cells encode features of the environment in relation to their predictive relationship with other features, such as the potential presence of reward. Additionally, state prediction in HPC (*via* theta sequences) can be coupled through theta coordination coherence to other areas of the brain that process action and reward. **(Bottom left)** In muZero, the dynamics function calculates reward (such as a captured checkers piece) and computes a new hidden state. **(Bottom right)** In the EHPCs, offline replay, which occurs during periods of sleep, emphasizes locations of previous reward. During replay, events can be selectively sampled from visited event space, with priority given to representations or locations that correspond with reward. Reward-predictive representations can then be compressed into a lower dimensional space. Replay can be comprised of novel state configurations that have not been experienced, such as a novel route to the reward location.

Theta sequences

The HPC also has an additional "lookahead" mechanism to evaluate future trajectory choices after an animal has experienced an environment and formulated a "map" with place cells. This hippocampal phenomenon, termed theta sequences, is a time-compressed and sequential firing of HPC place cells as if the animal were traversing a trajectory (Foster and Wilson, 2007). Because theta sequences can represent future possible trajectories, they are believed to play a role in planning (Wikenheiser and Redish, 2015a,b). Consistent with this idea, theta sequences incorporate information about future goal locations, and can use phase coupling of these rhythms to coordinate processing with other brain areas, much like the prediction function in muZero (Wikenheiser and Redish, 2015b; Penagos et al., 2017) (**Figure 1**).

In addition, the eHPC work in coordination with other structures to link present and future states with potential future actions. For instance, state prediction in HPC (*via* theta sequences) can be coupled through theta coordination coherence to other areas of the brain that process action and reward, including executive systems such as the prefrontal cortex (policy function) and dopaminergic areas like the VTA (reward function). Disruption of this coherence may, therefore, interrupt functions that require accurate predictions and may thus impact learning (Sigurdsson et al., 2010; Lesting et al., 2013; Igarashi et al., 2014).

Replay

In muZero, sampling of future states occurs without the performance of an action and is most analogous to hippocampal offline processing that occurs during non-REM sleep and quiet wake. During these periods, cells in the hippocampus engage in replay: highly compressed (100s of ms) place cell sequences that form a trajectory (Wilson and McNaughton, 1994; Diba and Buzsaki, 2007). It is believed that hippocampal replay plays a fundamental role in memory consolidation and may also be important for planning (Gupta et al., 2010; O'Neill et al., 2010; Pfeiffer and Foster, 2013; Singer et al., 2013).

In one theory of hippocampal function, hippocampal replay also plays an important role in generalization by aiding in the offline construction of compressed environmental representations (Vértes and Sahani, 2019). During replay, events can be selectively sampled from visited event space, with priority given to representations that are reward predictive (Pfeiffer and Foster, 2013; Olafsdottir et al., 2015; Michon et al., 2019). Reward-predictive representations can then be compressed into a lower dimensional space that can be reused to plan future novel trajectories (Penagos et al., 2017). Replay may then use saved states to construct novel trajectories without the need for additional sampling or experience. Consistent with this idea, HPC replay selectively reinforces rewarded and significant locations, and it is possible that the shifting of place fields toward rewarded locations (a modification of representation) is dependent on replay. Additionally, replay can be comprised of novel sequences that have not been experienced, suggesting an environmental map can be generalized beyond prior experience (Gupta et al., 2010; Olafsdottir et al., 2015).

Questions for context dependent learning

The means by which the HPC is involved in the differential encoding of separate spatial environments is well studied. This process, known as "remapping," occurs when place cells change spatial representations in new environments (Colgin et al., 2008). In addition to the remapping that occurs when moving between starkly different environments, when exposed to continuous environmental change, place cells will gradually morph to remap to the new environment (Leutgeb et al., 2005). Although modeling experiments have shown that no threshold for remapping exists (Sanders et al., 2020), it is still unknown, experimentally if there is a threshold level of difference required to induce remapping between environments.

Importantly, learning also does not appear to be fully transferred between different environments where remapping is presumed to occur. For example, learning in both classical and operant conditioning tasks is nearly always context dependent, although the conditioned or operant response can be "re"-learned more quickly in the second environment (Urcelay and Miller, 2014). However, environmental changes may not affect the performance of navigational tasks to the same extent (Jeffery et al., 2003). Although cells in the HPC can encode conditioned and operant stimuli (as well as spatial location) (McEchron and Disterhoft, 1999; Wirtshafter and Wilson, 2019), it is not yet known whether any representation of the learned stimuli exists between remapped environments. Interestingly, however, animals with hippocampal lesions will transfer associative conditioning tasks across different environments (Penick and Solomon, 1991). It is similarly unknown whether this "failure" of context-dependent learning is due to an inappropriate preservation of stimuli encoding, or a remapping failure in which the representation of the first environmental interferes with the representation of the second.

Although muZero doesn't explicitly perform context discrimination, using the model it has built of the environment, muZero must infer, probabilistically, whether enough contextual elements have changed to require a new map and, potentially, new learning. Studying how probabilistic inference gives rise to context independent learning in muZero may provide insight into human and animal learning and generalization. Establishing what rules muZero uses to determine when a new map is needed and when learning is context dependent may shed light on the mechanisms behind contextual learning in animals.

Conceptual equivalence can result in predictions about hippocampal function

How might the advances made by muZero direct brain research? The evaluation of how transitions in cell representations occur between similar and distinct contexts is an under-researched topic in hippocampal systems neuroscience, and can be informed by advances in artificial intelligence agents such as muZero.

As previously explained, we believe AI can shed important light on the processes involved in contextual generalization. In muZero, there is a representational bias toward significant cues, and then this biased representation is compressed into a lower dimensionality. Representational bias allows for dimensionality reduction by beginning the selection of the most significant contextual elements (Lehnert et al., 2020). This reduction enables generalization by allowing learning inferences between environments with similarly compressed maps; i.e., in muZero, carry over of biased compressed representations allows for generalization (Lehnert et al., 2020).

In muZero, compression, which occurs during the dynamics function, plays an essential role in the generalization of learning. As previously explained, compression in the hippocampus may occur during hippocampal replay during sleep (**Figure 1**) (Penagos et al., 2017), and, like in muZero, this compression may be an important component of learning generalization (Vértes and Sahani, 2019). We might expect, therefore, for disruptions in sleep replay to cause deficits in task generalization (Hoel, 2021), and for increased replay (such as sleep sessions between tasks) to enhance learning generalization (Djonlagic et al., 2009). If replay plays an important role in generalization, we also might expect replay sequences and structure to be similar between two environments with the same task rules. Specifically, replay may be structured around task organization and reward-contingent actions, resulting in a clustering of replay around shared action states between two environments.

The compression of salient cues informs the creation of new hidden states, which may explain mechanisms of remapping between environments (Sanders et al., 2020). muZero's dynamics function does more than calculate a new hidden state based solely on known context; it also makes inferences based on policy (action) and value (reward) mapping. In other words, a state representation not only reflects the environment, but also the actions that lead to environmental consequences. If the hippocampus follows a similar mechanism, it is likely that "knowing" when to remap requires more than a simple evaluation of environmental cues (there may be similar contexts that require different actions) but also necessitates identifying differences in potential policies and rewards, especially during periods of uncertainty (such as when experiencing novelty) (Penagos et al., 2017; Sanders et al., 2020; Whittington et al., 2020). Because a portion of this evaluation of state similarity and difference happens offline, sleep replay may be an essential mechanism for remapping. In this context, remapping may result in the generating or updating of hidden states when there is uncertainty due to change in inputs.

The speed and type (partial or complete) of remapping that occurs may therefore depend on how past experience is evaluated offline during replay (Plitt and Giocomo, 2021). We might also, therefore, expect disrupting replay to change the speed at which remapping occurs, or whether it occurs at all. Failures in remapping would likely also impact the ability to distinguish different environments during learning, thereby causing inability to learn skills in a new environment, conflation of skills learned in different environments, and/or perseverance of skills in the incorrect environment.

The majority of current hippocampal research focuses on representational correspondence between neural activity and context, such as how the brain represents the spatial correlates of an environment through neuronal firing patterns. Absent from much of this research is a larger focus on questions of algorithmic correspondence between neuron activity and context, such as determining the required difference between two environments to induce remapping. We believe advances in AI agents will assist in shifting research toward the investigation of algorithmic correspondence, *via* providing frameworks and predictions (such the necessity for replay in contextual generalization), by which to investigate the expected link between state changes and neuronal firing. Agents such as muZero can aid in illuminating these prospective questions, as well as shedding light on potential expected answers.

Data availability statement

Author contributions

HW and MW: study conception and design and manuscript editing. HW: draft manuscript preparation. Both authors reviewed the results and approved the final version of the manuscript.

Funding

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Abel, D., Arumugam, D., Lehnert, L., and Littman, M. (2018). State abstractions for lifelong reinforcement learning, International Conference on Machine Learning. *PMLR* 80, 10–19.

Banino, A., Barry, C., Uria, B., Blundell, C., Lillicrap, T., Mirowski, P., et al. (2018). Vector-based navigation using grid-like representations in artificial agents. *Nature* 557, 429–433. doi: 10.1038/s41586-018-0102-6

Bermudez-Contreras, E., Clark, B. J., and Wilber, A. (2020). The Neuroscience of Spatial Navigation and the Relationship to Artificial Intelligence. *Front. Comput. Neurosci.* 14:63. doi: 10.3389/fncom.2020.00063

Boccara, C. N., Nardin, M., Stella, F., O'Neill, J., and Csicsvari, J. (2019). The entorhinal cognitive map is attracted to goals. *Science* 363, 1443–1447. doi: 10.1126/science.aav4837

Cichy, R. M., Khosla, A., Pantazis, D., Torralba, A., and Oliva, A. (2016). Comparison of deep neural networks to spatio-temporal cortical dynamics of human visual object recognition reveals hierarchical correspondence. *Sci. Rep.* 6:27755. doi: 10.1038/srep27755

Colgin, L. L., Moser, E. I., and Moser, M. B. (2008). Understanding memory through hippocampal remapping. *Cell* 31, 469–477. doi: 10.1016/j.tins.2008.06.008

Collins, A. G. E., and Frank, M. J. (2016). Neural signature of hierarchically structured expectations predicts clustering and transfer of rule sets in reinforcement learning. *Cognition* 152, 160–169. doi: 10.1016/j.cognition.2016.04.002

Dabney, W., Kurth-Nelson, Z., Uchida, N., Starkweather, C. K., Hassabis, D., Munos, R., et al. (2020). A distributional code for value in dopamine-based reinforcement learning. *Nature* 577, 671–675. doi: 10.1038/s41586-019-1924-6

Diba, K., and Buzsaki, G. (2007). Forward and reverse hippocampal place-cell sequences during ripples. *Nat. Neurosci.* 10, 1241–1242. doi: 10.1038/nn1961

Djonlagic, I., Rosenfeld, A., Shohamy, D., Myers, C., Gluck, M., and Stickgold, R. (2009). Sleep enhances category learning. *Learn. Mem.* 16, 751–755. doi: 10.1101/lm.1634509

Foster, D. J., and Wilson, M. A. (2007). Hippocampal theta sequences. *Hippocampus* 17, 1093–1099. doi: 10.1002/hipo.20345

Fyhn, M., Hafting, T., Treves, A., Moser, M. B., and Moser, E. I. (2007). Hippocampal remapping and grid realignment in entorhinal cortex. *Nature* 446, 190–194. doi: 10.1038/nature05601

Gershman, S. J. (2018). The Successor Representation: Its Computational Logic and Neural Substrates. *J. Neurosci.* 38, 7193–7200. doi: 10.1523/JNEUROSCI.0151-18.2018

Gupta, A. S., van der Meer, M. A., Touretzky, D. S., and Redish, A. D. (2010). Hippocampal replay is not a simple function of experience. *Neuron* 65, 695–705. doi: 10.1016/j.neuron.2010.01.034

Hafting, T., Fyhn, M., Molden, S., Moser, M. B., and Moser, E. I. (2005). Microstructure of a spatial map in the entorhinal cortex. *Nature* 436, 801–806. doi: 10.1038/nature03721

Hoel, E. (2021). The overfitted brain: Dreams evolved to assist generalization. *Patterns* 2:100244. doi: 10.1016/j.patter.2021.100244

Hollup, S. A., Molden, S., Donnett, J. G., Moser, M. B., and Moser, E. I. (2001). Accumulation of hippocampal place fields at the goal location in an annular watermaze task. *J. Neurosci.* 21, 1635–1644. doi: 10.1523/JNEUROSCI.21-05-01635.2001

Igarashi, K. M., Lu, L., Colgin, L. L., Moser, M. B., and Moser, E. I. (2014). Coordination of entorhinal-hippocampal ensemble activity during associative learning. *Nature* 510, 143–147. doi: 10.1038/nature13162

Jackson, R. L., Rogers, T. T., and Lambon Ralph, M. A. (2021). Reverse-engineering the cortical architecture for controlled semantic cognition. *Nat. Hum. Behav.* 5, 774–786. doi: 10.1038/s41562-020-01034-z

Jeffery, K. J., Gilbert, A., Burton, S., and Strudwick, A. (2003). Preserved performance in a hippocampal-dependent spatial task despite complete place cell remapping. *Hippocampus* 13, 175–189. doi: 10.1002/hipo.10047

Kell, A. J. E., Yamins, D. L. K., Shook, E. N., Norman-Haignere, S. V., and McDermott, J. H. (2018). A Task-Optimized Neural Network Replicates Human Auditory Behavior, Predicts Brain Responses, and Reveals a Cortical Processing Hierarchy. *Neuron* 98, 630–644e16. doi: 10.1016/j.neuron.2018.03.044

Krizhevsky, A., Sutskever, I., and Hinton, E. (2012). Imagenet classification with deep convolutional neural networks. *Adv. Neural Inf. Process. Syst.* 25, 1097–1105.

Lee, I., Griffin, A. L., Zilli, E. A., Eichenbaum, H., and Hasselmo, M. E. (2006). Gradual translocation of spatial correlates of neuronal firing in the hippocampus toward prospective reward locations. *Neuron* 51, 639–650. doi: 10.1016/j.neuron.2006.06.033

Lehnert, L., Littman, M. L., and Frank, M. J. (2020). Reward-predictive representations generalize across tasks in reinforcement learning. *PLoS Comput. Biol.* 16:e1008317. doi: 10.1371/journal.pcbi.1008317

Lesting, J., Daldrup, T., Narayanan, V., Himpe, C., Seidenbecher, T., Pape, H. C., et al. (2013). Directional theta coherence in prefrontal cortical to amygdalo-hippocampal pathways signals fear extinction. *PLoS One* 8:e77707. doi: 10.1371/journal.pone.0077707

Leutgeb, J. K., Leutgeb, S., Treves, A., Meyer, R., Barnes, C. A., McNaughton, B. L., et al. (2005). Progressive transformation of hippocampal neuronal representations in "morphed" environments. *Neuron* 48, 345–358. doi: 10.1016/j.neuron.2005.09.007

Lian, Y., and Burkitt, A. N. (2021). Learning an efficient hippocampal place map from entorhinal inputs using non-negative sparse coding. *eNeuro* 8, ENEURO.557–ENEURO.520. doi: 10.1523/ENEURO.0557-20.2021

McEchron, M. D., and Disterhoft, J. F. (1999). Hippocampal encoding of non-spatial trace conditioning. *Hippocampus* 9, 385–396. doi: 10.1002/(SICI)1098-1063(1999)9:4<385::AID-HIPO5>3.0.CO;2-K

Michon, F., Sun, J. J., Kim, C. Y., Ciliberti, D., and Kloosterman, F. (2019). Post-learning Hippocampal Replay Selectively Reinforces Spatial Memory for Highly Rewarded Locations. *Curr. Biol.* 29, 1436–1444e5. doi: 10.1016/j.cub.2019.03.048

Moser, E. I., Kropff, E., and Moser, M. B. (2008). Place cells, grid cells, and the brain's spatial representation system. *Annu. Rev. Neurosci.* 31, 69–89. doi: 10.1146/annurev.neuro.31.061307.090723

O'Keefe, J. (1976). Place units in the hippocampus of the freely moving rat. *Exp. Neurol.* 51, 78–109. doi: 10.1016/0014-4886(76)90055-8

Olafsdottir, H. F., Barry, C., Saleem, A. B., Hassabis, D., and Spiers, H. J. (2015). Hippocampal place cells construct reward related sequences through unexplored space. *eLife* 4:e06063. doi: 10.7554/eLife.06063

O'Neill, J., Pleydell-Bouverie, B., Dupret, D., and Csicsvari, J. (2010). Play it again: Reactivation of waking experience and memory. *Trends Neurosci.* 33, 220–229. doi: 10.1016/j.tins.2010.01.006

Penagos, H., Varela, C., and Wilson, M. A. (2017). Oscillations, neural computations and learning during wake and sleep. *Curr. Opin. Neurobiol.* 44, 193–201. doi: 10.1016/j.conb.2017.05.009

Penick, S., and Solomon, P. R. (1991). Hippocampus, context, and conditioning. *Behav. Neurosci.* 105, 611–617. doi: 10.1037/0735-7044.105.5.611

Pfeiffer, B. E., and Foster, D. J. (2013). Hippocampal place-cell sequences depict future paths to remembered goals. *Nature* 497, 74–79. doi: 10.1038/nature12112

Plitt, M. H., and Giocomo, L. M. (2021). Experience-dependent contextual codes in the hippocampus. *Nat. Neurosci.* 24, 705–714. doi: 10.1038/s41593-021-00816-6

Richards, B. A., Lillicrap, T. P., Beaudoin, P., Bengio, Y., Bogacz, R., Christensen, A., et al. (2019). A deep learning framework for neuroscience. *Nat. Neurosci.* 22, 1761–1770. doi: 10.1038/s41593-019-0520-2

Sanders, H., Wilson, M. A., and Gershman, S. J. (2020). Hippocampal remapping as hidden state inference. *eLife* 9:e51140. doi: 10.7554/eLife.51140

Schrittwieser, J., Antonoglou, I., Hubert, T., Simonyan, K., Sifre, L., Schmitt, S., et al. (2020). Mastering Atari, Go, chess and shogi by planning with a learned model. *Nature* 588, 604–609. doi: 10.1038/s41586-020-03051-4

Sigurdsson, T., Stark, K., Karayiorgou, M., Gogos, J. A., and Gordon, A. (2010). Impaired hippocampal-prefrontal synchrony in a genetic mouse model of schizophrenia. *Nature* 464, 763–767. doi: 10.1038/nature08855

Silver, D., Huang, A., Maddison, C., Guez, A., Sifre, L., van den Driessche, G., et al. (2016). Mastering the game of Go with deep neural networks and tree search. *Nature* 529, 484–489. doi: 10.1038/nature16961

Singer, A. C., Carr, M. F., Karlsson, M. P., and Frank, L. M. (2013). Hippocampal SWR activity predicts correct decisions during the initial learning of an alternation task. *Neuron* 77, 1163–1173. doi: 10.1016/j.neuron.2013.01.027

Stachenfeld, K. L., Botvinick, M. M., and Gershman, S. J. (2017). The hippocampus as a predictive map. *Nat. Neurosci.* 20, 1643–1653. doi: 10.1038/nn.4650

Urcelay, G. P., and Miller, R. R. (2014). The functions of contexts in associative learning. *Behav. Process.* 104, 2–12. doi: 10.1016/j.beproc.2014.02.008

Vértes, E., and Sahani, M. (2019). A neurally plausible model learns successor representations in partially observable environments. *Adv. Neural Inf. Process. Syst.* 32, 13714–13724.

Wang, J. X., Kurth-Nelson, Z., Kumaran, D., Tirumala, D., Soyer, H., Leibo, J. Z., et al. (2018). Prefrontal cortex as a meta-reinforcement learning system. *Nat. Neurosci.* 21, 860–868. doi: 10.1038/s41593-018-0147-8

Whittington, J. C. R., Muller, T. H., Mark, S., Chen, G., Barry, C., Burgess, N., et al. (2020). The Tolman-Eichenbaum Machine: Unifying Space and Relational Memory through Generalization in the Hippocampal Formation. *Cell* 183, 1249–1263e23. doi: 10.1016/j.cell.2020.10.024

Wikenheiser, A. M., and Redish, A. D. (2015a). Decoding the cognitive map: Ensemble hippocampal sequences and decision making. *Curr. Opin. Neurobiol.* 32, 8–15. doi: 10.1016/j.conb.2014.10.002

Wikenheiser, A. M., and Redish, A. D. (2015b). Hippocampal theta sequences reflect current goals. *Nat. Neurosci.* 18, 289–294. doi: 10.1038/nn. 3909

Wilson, M. A., and McNaughton, B. L. (1994). Reactivation of hippocampal ensemble memories during sleep. *Science* 265, 676–679. doi: 10.1126/science. 8036517

Wirtshafter, H. S., and Wilson, M. A. (2019). Locomotor and Hippocampal Processing Converge in the Lateral Septum. *Curr. Biol.* 29, 3177–3192. doi: 10. 1016/j.cub.2019.07.089

Wirtshafter, H. S., and Wilson, M. A. (2020). Differences in reward biased spatial representations in the lateral septum and hippocampus. *eLife* 9:e55252. doi: 10.7554/eLife.55252

Wirtshafter, H. S., and Wilson, M. A. (2021). Lateral septum as a nexus for mood, motivation, and movement. *Neurosci. Biobehav. Rev.* 126, 544–559. doi: 10.1016/j.neubiorev.2021.03.029

Zador, A. M. (2019). A critique of pure learning and what artificial neural networks can learn from animal brains. *Nat. Commun.* 10:3770. doi: 10.1038/ s41467-019-11786-6