Please cite this article in press as: Samiei et al., The schema spectrum: Emergent structures and levels of abstraction in AI and the brain, Neuron (2026), https://doi.org/10.1016/j.neuron.2026.05.007

Neuron

CellPress

Perspective

The schema spectrum: Emergent structures and levels of abstraction in AI and the brain

Mandana Samiei,[1,2,*] Doina Precup,[1,2,3] and Blake A. Richards[1,2,3,4,5,*]
[1]School of Computer Science, McGill University, Montreal, QC, Canada
[2]Mila - Quebec Artificial Intelligence Institute, Montreal, QC, Canada
[3]CIFAR Learning in Machines and Brains Program, Toronto, ON, Canada
[4]Department of Neurology & Neurosurgery, McGill University, Montreal, QC, Canada
[5]Montreal Neurological Institute, McGill University, Montreal, QC, Canada
*Correspondence: mandana.samiei@mail.mcgill.ca (M.S.), blake.richards@mila.quebec (B.A.R.)
https://doi.org/10.1016/j.neuron.2026.05.007

SUMMARY

There is a long history of interplay between the brain sciences and AI in the area of schema theory. Schemas are typically defined as abstract mental structures representing prior knowledge, experiences, and concepts that capture how events unfold in different contexts and that alter how we learn new information. Classical models have treated schemas as being distinct from both detail-rich episodic memories and general semantic knowledge. Motivated by learning phenomena in modern generative AI, we propose that earlier connectionist theories that did not articulate any strict division between schemas and other declarative memories should be revived. According to this perspective, schemas are not a distinct set of mnemonic objects in the brain; rather, they are a conceptual tool that scientists use to describe how existing knowledge can exist along a spectrum of abstraction.

INTRODUCTION TO SCHEMA THEORY

Schema theory has a long and influential history across psychology, cognitive science, and neuroscience.[1–9] It also played a pivotal role in early artificial intelligence (AI) research,[10,11] providing a foundational formalism for structured knowledge representation through frame-based architectures[10] and script theory.[11] The concept of schemas was originally introduced by Piaget[1] in his work on cognitive development. For Piaget, schemas were frameworks for knowledge—cognitive structures that allow individuals to organize and interpret information as they develop.

He proposed that learning occurs through two complementary processes of structural adaptation.[1,12] The process of assimilation refers to situations in which new experiences or knowledge are re-interpreted to fit existing schemas. For example, a child with a schema for "dogs" may refer to any furry four-legged creature as a dog. In contrast, accommodation refers to situations in which existing schemas must be modified to account for novel information that violates current structural expectations. For example, a child learning that furry four-legged creatures that meow are not dogs, but cats, will modify their schema of furry four-legged animals. This dynamic adaptation of schemas remains a cornerstone of developmental psychology.[13] It also relates to more recent theories that frame schemas' impact on new information based on a spectrum of expertise.[8]

Shortly after Piaget's developmental work, Frederic Bartlett[2] extended the concept of schemas into the domain of memory research. He shifted the focus from structural growth to the reconstructive nature of recall, demonstrating that the pre-existing frameworks can distort the storage and retrieval of specific episodic details. Specifically, his experiments on story recall showed that participants unconsciously reshaped narratives to fit their cultural and cognitive expectations, often omitting or altering details that did not align with their schemas. For example, people of a European background who read an indigenous North American text would recall that the characters went "fishing" or "sailing" even though text stated that they went "seal hunting." Bartlett concluded from these studies that memory recall is, in part, a reconstructive process that depends on individual's schemas and not an exact recall of past information. This idea foreshadowed later discussions on memory distortion, constructive retrieval, and the role of schemas in shaping perception and recall.[14]

By the late 20th century, schema theory was formalized in cognitive science through models of structured knowledge representation. Schank and Abelson[11] introduced script theory, a specialized form of schemas that encode typical event sequences in human cognition. For example, we have a "restaurant script" that allows us to know how a visit to a restaurant will unfold. This script would include typical objects, roles, and scenes for different events such as being seated, ordering, eating, and paying. Schank and Abelson used script theory to guide the development of AI systems, such as the Script Applier Mechanism (SAM), designed to understand and recall stories.[15] For example, SAM could read a news story about a man ordering food and infer that he probably paid at the end—even if that step was never mentioned. Psychological research has generally supported aspects of Schank and Abelson's script theory; for instance, people tend to recall events from a story based on their

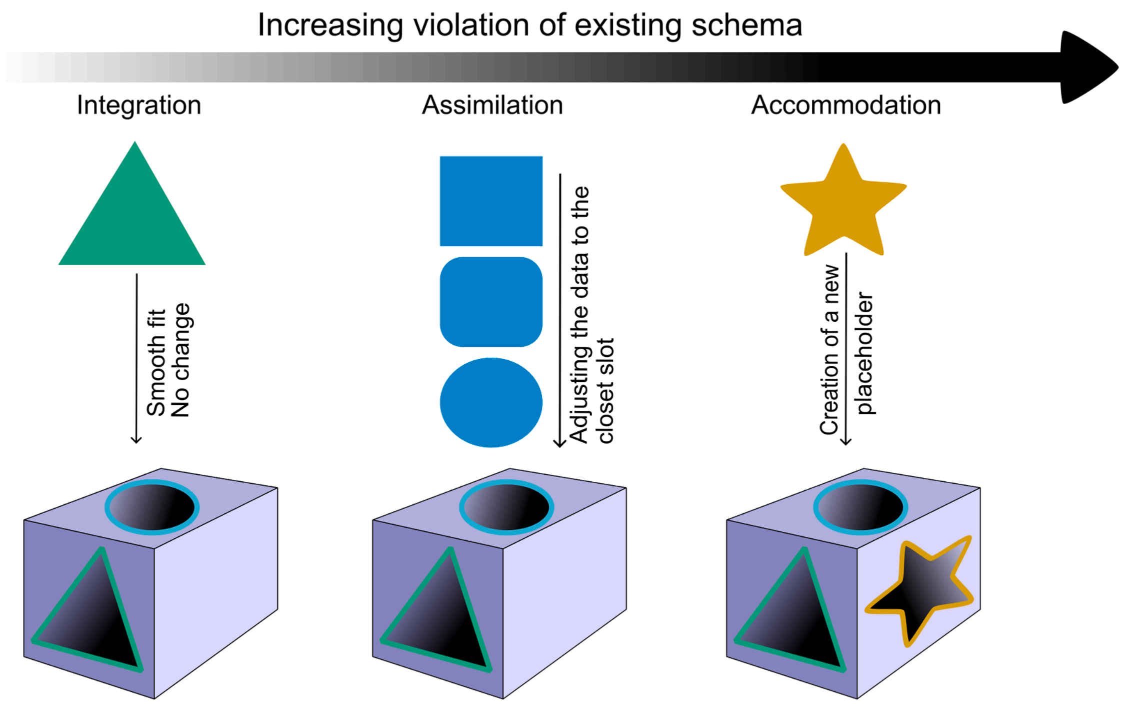

Figure 1. Schematic representation of cognitive processing based on schema theory
New information can be integrated through three main processes depending on its complexity and novelty: (1) integration occurs when incoming data fit smoothly into an existing schema without requiring changes; (2) assimilation requires adjusting the new data to match the closest existing schema; and (3) accommodation involves creating a new schema when the information cannot fit into any existing structure. The requirement for additional computation and learning increases from integration to accommodation.

familiar order according to typical scripts rather than the order in which the events necessarily occurred.[16]

Later, in 1980, Rumelhart proposed that schemas are "hierarchical knowledge structures" that organize all levels of understanding, allowing for generalization and inference.[3] According to Rumelhart's conception, schemas provide "slots" or "variables" that specify the components or attributes of a given concept and "values" or "fillers" that provide specific information for an instantiation of that concept. For instance, we may have a schema for houses that would contain variables such as "number of bedrooms," "neighborhood," or "type of heating," which can be readily filled when encountering a new house. According to Rumelrhart's framework, new information is "integrated" into a schema when we map a given value to a given variable.

Dedre Gentner's later work on structure-mapping theory[17] advanced a complementary view of how structured knowledge emerges through experience. Instead of treating schemas as static templates, she proposed that cognition operates through processes of relational alignment and abstraction, whereby structural correspondences between familiar and novel domains support understanding and transfer.[18] Through repeated analogical comparisons, individuals form relational schemas that generalize across contexts—a process she termed "progressive alignment."[19]

Combining Piaget's[12] theories with later conceptions of schemas like those of Rumelhart,[3] we can identify three distinct processes that schema theory would describe when we store or interpret new information. These three processes depend on how well new information matches existing schemas (Figure 1).

- In cases of high fit, we get integration; new information is rapidly and accurately stored using an existing schema.
- In cases of medium fit, we get assimilation; new information is rapidly stored using an existing schema but is modified to fit it.
- In cases of low fit, we get accommodation; the existing schema is updated, or a new schema is created, in order to account for the new information, requiring more time and computation.

Modern schema theory is as diverse as these classic works, though within modern theories, there are consistent ideas related to abstraction, novelty, distributed memory traces, and consolidation.[8,9,20–26] Additionally, modern theory tends to place more emphasis on the neurobiology of schema-related learning.[8,21,24] In what follows, though, we will use these three processes derived from more classical theories (integration, assimilation, and accommodation) to form the foundation of our analysis of how schema theory manifests in both neuroscience and modern AI research.

NEUROSCIENCE OF SCHEMAS

Given that schemas capture existing abstract knowledge formed through repeated experiences, they are deeply intertwined with both episodic and semantic memories.[2,23] Episodic memories are typically understood as detail rich, ‘‘autonoetic’’ memories that enable "mental time travel."[27] But, episodic memories themselves can be encoded and stored with varying levels of detail.[25] Even if we can mentally return to some event, we may not be able to remember every detail, and some aspects of the experience will be filled in with the ‘‘gist’’ of what occurred (though a true episodic memory will, by definition, always include some degree of subjective detail[27]). Interestingly, the gradient of how ‘‘gist-like’’ an episodic memory is reflects anatomical differentiation along the hippocampal long axis: the posterior hippocampus and associated sensory cortices support fine-grained, perceptually rich memories, while anterior hippocampus and associative cortical areas are implicated in more schematized, gist-like representations of past experiences.[28,29] Like schemas, semantic memories are typically understood to be general knowledge that is wholly abstracted from any specific experiences, yet evidence shows that semantic knowledge can also be contextual to some degree.[30] Moreover, semantic memory is typically understood as referencing specific facts, as opposed to schemas, which capture highly general relations.[23] These insights point to the possibility that fully disentangling episodic and semantic memories from schemas may be impossible, even at the level of physiology.[31] Supporting this, declarative memories, in general, seem to implicate a series of connected networks in the temporal lobes and prefrontal cortex (PFC).[32]

Naturally, then, the physiological basis for how schemas impact new learning seems to depend on a similar array of distributed networks. When incoming information aligns well with existing schemas, leading to integration, evidence suggests that the PFC exhibits enhanced activity and strong functional coupling with multiple cortical regions.[33–37] This interaction facilitates rapid integration of new information and may involve local synaptic plasticity within cortical circuits.[38–40] Through this interplay between the PFC and the rest of the brain, schema-consistent information can be incorporated directly into cortical representations, reducing reliance on encoding detailed episodic memory and bypassing prolonged hippocampal consolidation.[4,8,34,39,41,42]

In cases of assimilation, when new inputs partially align with existing schemas, encoding is still facilitated but often biased. Incoming details are adjusted to conform to prior expectations, leading to distortions and less accurate recall.[43] More accurate recall of schema-inconsistent details appears to require enhanced PFC coupling with temporal and parietal regions to overcome the incongruence and store the memory accurately.[42,44,45] Interestingly, interfering with ongoing PFC processing can sometimes prevent such distortions of information during assimilation.[46]

In contrast, accommodation, which involves learning highly novel, schema-inconsistent or -unrelated information, strongly engages the hippocampus and medial temporal lobe (MTL) systems.[22,39,40,47] The engagement of the MTL likely reflects interactions between the hippocampus and neuromodulatory systems in the midbrain and brainstem, most notably, dopaminergic and adrenergic pathways.[48,49] Exposure to novel information may trigger a neuromodulatory state that promotes the formation of entirely new episodic memories in the hippocampus and related structures.[8,49] Over a prolonged consolidation period, and the formation of multiple novel episodic memories, a new schema can be formed or existing schemas can be updated to accommodate this new information.[4,50,51] Paradoxically, this process may enable individuals to remember wholly novel or highly schema-inconsistent information more effectively than slightly schema-consistent information.[43,48,52]

Together, these schema-guided processes of memory formation may support the construction and maintenance of abstract, compositional knowledge structures that can enable flexible behavior and planning. Such organization of experience allows the brain to reuse and recombine learned components when facing new situations, a principle that closely mirrors hierarchical strategies in reinforcement learning.[53–57] But, where do episodic memories, semantic memories, and schemas start and end? To what extent are these concepts referring to wholly distinct processes in the brain? Understanding these correspondences opens the door for computational and AI systems to be related to schema theory.

COMPUTATIONAL MODELS OF SCHEMAS

In early symbolic AI (AI based on logical symbol systems), schemas were formalized as frame-based structures, with clearly defined slots and default values to be filled with contextual information, allowing for flexible yet structured reasoning.[10] This concept was expanded to include the structure of events with ‘‘scripts,’’ i.e., frame-like structures that provided the scaffolding that involved a stereotyped sequence of actions or events in a particular context, such as going to a restaurant or visiting a doctor.[11] While these symbolic representations were highly influential in cognitive science and early AI, they ultimately struggled to scale to the complexity and variability of real-world inputs.

Like symbolic AI practitioners, connectionist researchers also placed emphasis on schemas, but they tended to take a different tack. Rumelhart very specifically highlighted schemas as critical to understanding human cognition, arguing that knowledge is organized into flexible, interconnected mental frameworks via schemas.[3,58,59] His work emphasized that the frameworks that schemas provide are crucial for making sense of the world, understanding language, remembering experiences, and acquiring new knowledge, but this early work was largely couched at higher level.

Later connectionist models from the 1980s, inspired by Rumelhart, re-imagined schemas as emergent phenomena in distributed neural systems. Most notably, harmony theory proposed that sub-symbolic cognitive systems could represent knowledge through patterns of activation across interconnected units. According to this perspective, schemas emerge as stable attractor states in a neural network's dynamics.[60,61] Knowledge atoms, small units of information, are activated through spreading activation, and schemas correspond to coherent, high-‘‘harmony’’ configurations that integrate these atoms. More concretely, harmony theory essentially proposed that schemas were high-probability activity patterns in a distributed neural network and that new information was either integrated, assimilated, or accommodated based on how well it matched these existing high-probability activity patterns.

Similar to harmony theory, adaptive resonance theory provided a powerful sub-symbolic account of knowledge acquisition—and, implicitly, schemas—not as a result of pre-defined symbolic structures but rather as emergent categories or prototypes that a network can learn through unsupervised or supervised interactions with the environment.[62,63]

Bridging symbolic and connectionist paradigms, Gary Drescher's constructivist approach[64] offered a computational realization of Piaget's notions of assimilation and accommodation,[65] proposing that schemas can be incrementally constructed through sensorimotor experience and internal simulation. Drescher's work proposed how abstract knowledge structures could arise from low-level interactions with the environment, a perspective that presaged later cognitive accounts of self-organizing schema representations.[64]

These classical works laid the foundation for later conceptualizations of schemas that built on the neuroscience of memory consolidation. Complementary learning systems (CLS) theory proposed that learning about the patterns in episodes of experience requires an interaction between a fast-learning episodic system (i.e., the hippocampus) and a slow-learning neural network that stores semantic knowledge via gradual integration of information (i.e., the neocortex).[66] CLS was later updated to account for humans' and animals' ability to weight experiences based on their goals and to incorporate information into neocortical traces rapidly through schematic integration.[67] Recent computational work has advanced the CLS framework by demonstrating that a rapidly learning "hippocampus" can act as a teacher for a "neocortical" generative model, allowing the system to reconstruct specific episodes or generate entirely new gist-like samples from learned distributions.[9] These ideas provided the background for other more recent approaches in AI that seek to address problems of continual learning, specifically the challenge of catastrophic forgetting, where new information overwrites previously learned skills. For instance, recent work in reinforcement learning uses predictive structures to stabilize agents in non-stationary environments, ensuring that updating for a new task does not overwrite the internal models of past experiences.[68] Furthermore, the emergent "schematization" seen in modern generative models has been proposed as a biological-inspired solution: by consolidating sequential experiences into a deep generative network, AI systems can move beyond simple storage toward generative reconstruction.[9] This allows for the flexible reuse of abstract components across different contexts, mirroring the way human memory construction supports prediction and planning.

Notably, though, many of the connectionist approaches, including harmony theory, CLS, and more modern models of schemas are very unclear as to how schemas fundamentally differ from semantic memory: both are often treated as abstract representations formed via gradual, interleaved training across multiple episodes of experience.[9,26,61,67] We believe that this lack of distinction between different forms of declarative memory may be a feature of connectionist theories and not a bug. Specifically, we believe that when we consider large language models (LLMs), which arguably have nothing but a large semantic knowledge store, we can see many parallels to the phenomena of schema-based learning. Below, we argue that these advances invite a reconsideration of schemas not as distinct cognitive constructs but as the end of a spectrum of abstraction in memory, aligning closely with the vision of early connectionist theories.[61,62]

EVIDENCE FOR SCHEMAS IN LLMs

Large-scale, modern AI models, such as LLMs, are not explicitly designed to have schemas. Rather, they are designed to be flexible sequence processing models, with the ability to identify how different contexts impact the interpretation of items in a sequence.[69,70] Thus, these models are very different from both early symbolic AI models and more recent graph-based models that use explicit, frame-based or graph-based representations to structure and store new knowledge.[10,15,71] Moreover, unlike some of the early connectionist systems,[61,72] LLMs are not generally framed as models of schematic knowledge representation. Yet, LLMs often display many learning behaviors that are reminiscent of the functional hallmarks of schema-driven learning, i.e., integration, assimilation, and accommodation. LLMs can identify schemas within data though,[73–75] and, as we argue below, they seem to possess mnemonic schemas similar to those that humans possess, at least, in terms of how they impact new learning.

When new information provided to an LLM aligns well with the data encountered during pretraining, LLMs can rapidly learn the new information, even small or any updates to their synaptic weight parameters.[76,77] This process is often referred to as "in-context-learning," and it results from changes to the internal activation vectors that mimic the process of learning but with no parameter changes.[78,79] (It should be noted that LLMs will not store that information permanently unless they update their parameters.) Evidence suggests that in-context learning is a result of the models using previously learned statistical and semantic relationships to interpret new data,[80,81] and this can even be framed as a form of implicit Bayesian inference.[78,82] Notably, in-context learning is similar to integration via schemas as observed in humans, because it provides LLMs with the ability to rapidly encode new information using existing knowledge when they are commensurate. Additionally, similar to schema-based learning, in-context learning does not work when new information conflicts with previously learned information.[83] This process, sometimes referred to as "knowledge conflict," can lead to "confirmation bias," wherein the model only keeps those parts of the new information that match its previous data,[84] or "hallucinations," where the model invents new information that better matches its previous data.[85] Arguably, the adaptation of new information to match existing knowledge is similar to the process of assimilation in schema-based learning in humans.

When in-context learning is not sufficient, training modern LLMs on new information involves actual changes to the model's synaptic weight parameters, a process known as "fine-tuning."[77,86,87] Similar to accommodation, fine-tuning can be used to modify existing stored information in order to create scaffolds for new behaviors. But, as with accommodation in humans, fine-tuning can modify existing knowledge or even overwrite it.[88,89]

Altogether, these observations suggest that LLMs have something akin to schemas, even if they are not storing information in exactly the same way that our brains do. Moreover,

Please cite this article in press as: Samiei et al., The schema spectrum: Emergent structures and levels of abstraction in AI and the brain, Neuron (2026), https://doi.org/10.1016/j.neuron.2026.05.007

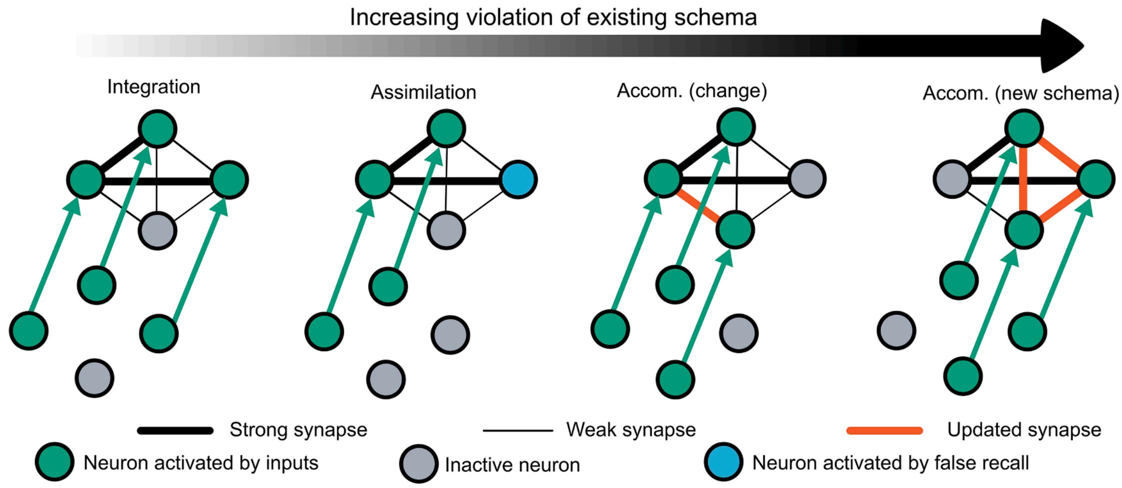

Figure 2. Schema theory mechanisms in neural networks
When new inputs are fully consistent with existing high-probability patterns (integration), activation flows along already strong synapses without significant change. With minor inconsistencies relative to existing patterns (assimilation), inputs can still be mapped onto existing connections, though false recall may occur, e.g., a neuron that wasn't originally active during storage could get activated at recall (blue neuron). Greater mismatches require accommodation (change), where specific synaptic weights are updated (shown in orange) to incorporate the new information. When violations of existing high-probability patterns are too extensive (accommodation), the system engages in more widespread updating of synaptic connections to support a reorganized representation.

these schema-like behaviors in LLMs echo earlier connectionist accounts, such as harmony theory, which proposed that schemas could emerge naturally in sub-symbolic systems as high-probability activation states.[61] Indeed, given that LLMs are pretrained to predict high-probability words in sentences, it is perhaps not surprising that their internal systems develop something like the schemas of harmony theory. Another way of phrasing this is that in LLMs, schemas exist and correspond to activation patterns in the neural network that map to high-probability sequences.

To summarize, large-scale AI models are not built with schemas by design; they have neither explicit frame-like nor graph-like structures for storing information. Yet, these models exhibit the hallmarks of schemas:

- When new information matches pre-existing knowledge, they can rapidly and accurately store the new information (integration).
- When new information only partially matches existing knowledge, they can rapidly store new information, but they often distort it (assimilation).
- When new information doesn't match existing knowledge, they require more extensive training to learn it, and this changes how they process future information (accommodation).

LESSONS TO BE DRAWN IN NEUROSCIENCE

What does the presence of schema-like information storage in LLMs tell us? The apparent presence of schema-like effects in modern AI models with no explicit, built-in schemas suggests the original ideas of harmony theory and adaptive resonance theory were closer to the mark than other computational theories of schematic storage.[61,63] It is not unreasonable, then, to take a few lessons from the conception of schemas that harmony theory initially proposed (Figure 2):

- When new information is well aligned to existing high-probability activation patterns (based on strong synaptic connections), learning can occur with almost no changes to synapses (integration).
- When new information partially aligns with existing high-probability activation patterns, learning can occur with few changes to synapses, but this could distort the new information when it is recalled (assimilation).
- When new information doesn't match existing high-probability activation patterns, learning requires large updates to synapses, potentially involving changes to existing high-probability patterns or construction of new high-probability activation patterns altogether (accommodation).

However, none of this is unique to schemas, per se. Any memory stored in a distributed, parallel processing system will exhibit similar behaviors to these.[60,61] As such, if we take this conception of schemas seriously, then schemas are probably not really distinct from other forms of memory—rather, they are part of a spectrum of mnemonic abstraction (Figure 3). At one end of this spectrum are highly detailed episodic memories that will depend significantly on the posterior hippocampus and its connections to sensory regions of the brain.[25] As mnemonic structures become more abstract, shifting from less-detailed autobiographical memory, to general semantic knowledge, and eventually to schematic knowledge, the underlying physiological substrates will shift, but the core principle of how existing connections and activity patterns impact new learning will remain the same.

Therefore, we believe that the distinction between "schematized" and "detailed" memories likely just relates to how abstract the information encoded by the relevant activation patterns is.

Please cite this article in press as: Samiei et al., The schema spectrum: Emergent structures and levels of abstraction in AI and the brain, Neuron (2026), https://doi.org/10.1016/j.neuron.2026.05.007

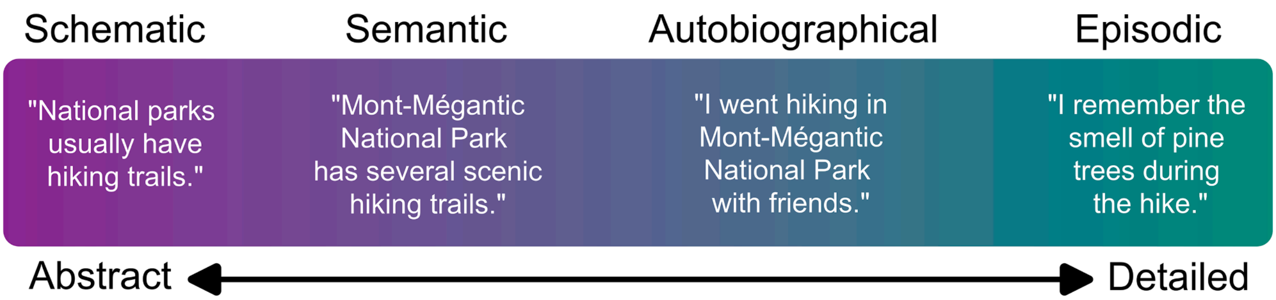

Figure 3. Spectrum of declarative memory types across levels of abstraction
We propose that different forms of declarative memories are best viewed not as being wholly distinct but as sitting along a spectrum of abstraction, with schemas occupying one extreme

Data showing that schematic processing depends on the anterior axis of the medial temporal lobes and the PFC may simply reflect the fact that these brain regions tend to be concerned with more abstract representations of the world, not that they support a fundamentally different form of memory structure. Moreover, the data showing that the hippocampus is particularly important when new information clashes with existing schemas may reflect the hippocampus' capacity for large amounts of rapid synaptic plasticity[90] and the requirement for the brain to update synapses when new information does not match existing high-probability activation patterns. The importance of neuromodulators in this process likely reflects both their role in detecting novelty and the importance of reinforcement learning for shaping activation patterns in the brain more generally.[53] Given these considerations, we suggest two implications for neuroscience.

The first implication is that "schemas" are conceptual tools that we scientists use to better understand processes in the brain that are otherwise hard to describe with language. In other words, there is no hard dividing line between "schemas" and other forms of memory, since all memory is likely to be a result of high-probability activation patterns. Schemas are distinct only so far as they engage with a very abstract representational space. Put differently, other forms of declarative memory are also likely to be captured by high-probability activation patterns, but they will be concerned with less abstract spaces. Indeed, even low-level sensory information can be shaped by previous experience, depending on how well it matches that previous experience.[91] It may simply be a quirk of scientific nomenclature and the history of schema theory (which is rooted in more abstract psychological studies) that we often don't consider how low-level, detailed memories can impact new learning in similar ways to "schemas."

The second implication is that experiments on how schemas impact memory consolidation should not assume that there is any anatomical or physiological dividing line that can neatly distinguish "schematized" and "non-schematized" memories. In line with this, an interesting shift in the memory consolidation literature in recent years has been a growing recognition that there is not a simple breakdown whereby episodic memories exist in the hippocampus and schematized memories exist in the neocortex.[25,41] Rather, when new information is stored, the role of the hippocampus likely has more to do with whether the new information matches existing synaptic connections, regardless of whether those connections encode abstract or detailed information. If there is a match, then learning can proceed rapidly with few changes to the networks in the brain, and thereby, there is less need for hippocampal involvement at encoding.[8] According to this perspective, the process of schematization of memories[4,24,33] may actually have less to do with transferring memories to the neocortex and more to do with learning abstract relationships that can be inferred only by averaging over many unique experiences.[50]

Altogether, this perspective invites a shift in how we conceptualize and discuss "schemas" within neuroscience. It suggests that we should be cautious of any model whereby schemas are viewed as distinct from other forms of memory and knowledge in the brain. Instead, it brings us to a perspective, closer to harmony theory and adaptive resonance theory, in which schemas are viewed as conceptual tools that we scientists apply to help us account for the ways in which previous, abstract knowledge stored in a neural network can impact the storage of new information. Schemas may simply be at the far end of a spectrum of abstraction that includes other forms of declarative memory.

CONCLUSION

The perspective we have articulated here is not entirely novel. The view that we arrive at from examining LLMs, where schema-like processes during new learning appear to exist despite no explicit attempt to build schemas, suggests that something akin to older connectionist theories of schemas (e.g., harmony theory) may be correct. Second, other authors have pointed out that schemas likely exist in a spectrum. For example, Alonso et al.[8] argued that the ways in which schemas can impact downstream learning is related to the degree of expertise of an individual—the role of the hippocampus and PFC will depend on just how naive or expert a subject is on that particular area of knowledge. Similarly, Tarder-Stoll et al.[25] argued that memories can exist along a spectrum of how much they contain gist elements and how much reconstruction of the memory is involved at recall—they propose that this spectrum relates to a broader anatomical spectrum along the posterior (more detailed) to anterior (more gist-like) axis. Finally, the idea that memory consolidation strengthens repeated, overlapping memory traces and that this process helps to produce ever more abstract understanding is well aligned with the ideas we articulated above and has been proposed by several prior studies.[23,51,92]

For neuroscience, the experimental implications of this perspective mostly relate to the need to think about how previous memories impact new learning more broadly, with abstraction as a continuous spectrum rather than a hard line. For example, perhaps experiments need to not assume that subjects will have a "schema" and instead investigate simply how detailed versus abstract their memories seem to be.[93] We may predict that subjects who extract more gist and more abstract relations will show more "schematic" processing, i.e., a higher reliance on anterior hippocampus at some degrees of abstraction, and even more reliance on PFC at very high degrees of abstraction.[25,41,94]

For AI systems, an interesting conclusion may be that current models have schemas, but they may not use them effectively. In particular, schemas, thanks to their abstraction, should help with learning tasks at multiple levels of abstraction (related to concepts in hierarchical reinforcement learning[95]). Recent work attempts to leverage the internal "schemas" of transformer models precisely to achieve this form of hierarchical behavioral control.[96] As well, the perspective we argued for here suggests that achieving more human-like memory systems in AI may not require having separate memory modules for each type of declarative memory.[97] Related to this, perhaps the best way to think about building in episodic memory into AI systems may not be to use a completely separate "bank" of non-parametric memory[98–100] but to consider methods for rapidly wiring new synaptic connections when novel, detail-rich information is received. Such mechanisms could help AI systems to learn continuously from new experiences while forming meaningful semantic connections across the existing ones.

Altogether, neuroscientists and AI researchers can and should continue to think about schema theory to help guide our ideation and research. However, we should avoid the temptation to think of schemas as distinct, explicit structures in either natural or artificial brains. Schemas may simply be part of a larger mnemonic spectrum.

ACKNOWLEDGMENTS

We acknowledge support by NSERC (Discovery Grant: RGPIN-2020-05105; Discovery Accelerator Supplement: RGPAS-2020-00031), CIFAR (Canada AI Chair; Learning in Machine and Brains Fellowship) and Fonds de recherche du Québec – Nature et technologies (FRQNT #2023-2024-B2X-336394). Additionally, we would like to thank Annik Carson for valuable discussions during the early stages of this work.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

1. Piaget, J. (1926). The Language and Thought of the Child (Kegan Paul, Trench & Trubner).
2. Bartlett, F.C. (1932). Remembering: A Study in Experimental and Social Psychology (Cambridge University Press).
3. Rumelhart, D.E. (1980). Schemata: The building blocks of cognition. In Theoretical Issues in Reading Comprehension, R.J. Spiro, B.C. Bruce, and W.F. Brewer, eds. (Lawrence Erlbaum), pp. 33–58.
4. Tse, D., Langston, R.F., Kakeyama, M., Bethus, I., Spooner, P.A., Wood, E.R., Witter, M.P., and Morris, R.G.M. (2007). Schemas and memory consolidation. Science *316*, 76–82.
5. Van Kesteren, M.T.R., Ruiter, D.J., Fernández, G., and Henson, R.N. (2012). How schema and novelty augment memory formation. Trends Neurosci. *35*, 211–219. https://doi.org/10.1016/j.tins.2012.02.001.
6. Kumaran, D. (2013). Schema-driven facilitation of new hierarchy learning in the transitive inference paradigm. Learn. Mem. *20*, 388–394. https://doi.org/10.1101/lm.030296.113.
7. Baldassano, C., Hasson, U., and Norman, K.A. (2018). Representation of real-world event schemas during narrative perception. J. Neurosci. *38*, 9689–9699. https://doi.org/10.1523/JNEUROSCI.0251-18.2018.
8. Alonso, R., van der Meij, J., Tse, D., and Genzel, L. (2020). Naïve to expert: Considering the role of previous knowledge in memory. Brain Neurosci. Adv. *4*, 2398212820948686. https://doi.org/10.1177/2398212820948686.
9. Spens, E., and Burgess, N. (2024). A generative model of memory construction and consolidation. Nat. Hum. Behav. *8*, 526–543. https://doi.org/10.1038/s41562-023-01799-z.
10. Minsky, M. (1975). A Framework for Representing Knowledge (McGraw-Hill). http://hdl.handle.net/1721.1/6089.
11. Schank, R.C., and Abelson, R.P. (1977). Scripts, Plans, Goals, and Understanding: An Inquiry into Human Knowledge Structures. Artificial Intelligence Series (Lawrence Erlbaum Associates). https://doi.org/10.4324/9780203781036.
12. Piaget, J. (1952). The Origins of Intelligence in Children (W. W. Norton & Company).
13. Goswami, U. (2008). Principles of learning, implications for teaching: A cognitive neuroscience perspective. J. Philos. Educ. *42*, 381–399. https://doi.org/10.1111/j.1467-9752.2008.00639.x.
14. Schacter, D.L., Coyle, J.T., Fischbach, G.D., Mesulam, M.M., and Sullivan, L.E. (1997). Memory Distortion (Harvard University Press).
15. Schank, R.C. (1975). SAM–a Story Understander. Research Report No. 43. (Yale University, Department of Computer Science) https://eric.ed.gov/?id=ED161024.
16. Bower, G.H., Black, J.B., and Turner, T.J. (1979). Scripts in memory for text. Cogn. Psychol. *11*, 177–220. https://doi.org/10.1016/0010-0285(79)90009-4.
17. Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. Cogn. Sci. *7*, 155–170. https://doi.org/10.1207/s15516709cog0702_3.
18. Gentner, D., Loewenstein, J., and Thompson, L. (2003). Learning and transfer: A general role for analogical encoding. J. Educ. Psychol. *95*, 393–408. https://doi.org/10.1037/0022-0663.95.2.393.
19. Gentner, D., and Loewenstein, J. (2002). Relational language and relational thought. In Language, literacy, and cognitive development: The development and consequences of symbolic communication, E. Amsel and J.P. Byrne, eds. (Psychology Press), pp. 87–120. https://api.semanticscholar.org/CorpusID:14863021.
20. Fernández, G., and Morris, R.G.M. (2018). Memory, novelty and prior knowledge. Trends Neurosci. *41*, 654–659.
21. Takeuchi, T., Tamura, M., Tse, D., Kajii, Y., Fernández, G., and Morris, R.G.M. (2022). Brain region networks for the assimilation of new associative memory into a schema. Mol. Brain *15*, 24.
22. Sekeres, M.J., Schomaker, J., Nadel, L., and Tse, D. (2024). To update or to create? the influence of novelty and prior knowledge on memory networks. Phil. Trans. R. Soc. B. *379*, 20230238. https://doi.org/10.1098/rstb.2023.0238.
23. Ghosh, V.E., and Gilboa, A. (2014). What is a memory schema? a historical perspective on current neuroscience literature. Neuropsychologia *53*, 104–114. https://doi.org/10.1016/j.neuropsychologia.2013.11.010.
24. Gilboa, A., and Marlatte, H. (2017). Neurobiology of schemas and schema-mediated memory. Trends Cognit. Sci. *21*, 618–631. https://doi.org/10.1016/j.tics.2017.04.013.

25. Tarder-Stoll, H., Sekeres, M.J., Levine, B., and Moscovitch, M. (2026). Adaptive episodic memory: how multiple memory representations drive behavior in humans and nonhumans. Physiol. Rev. *106*, 841–889. https://doi.org/10.1152/physrev.00005.2025.

26. Spens, E., and Burgess, N. (2025). Consolidation of sequential experience into a deep generative network explains human memory, prediction and planning. Preprint at bioRxiv. https://doi.org/10.1101/2024.11.04.621950.

27. Tulving, E. (2002). Episodic memory: From mind to brain. Annu. Rev. Psychol. *53*, 1–25. https://doi.org/10.1146/annurev.psych.53.100901.135114.

28. Poppenk, J., Evensmoen, H.R., Moscovitch, M., and Nadel, L. (2013). Long-axis specialization of the human hippocampus. Trends Cognit. Sci. *17*, 230–240. https://doi.org/10.1016/j.tics.2013.03.005.

29. Robin, J., and Moscovitch, M. (2017). Details, gist and schema: hippocampal–neocortical interactions underlying recent and remote episodic and spatial memory. Curr. Opin. Behav. Sci. *17*, 114–123. https://doi.org/10.1016/j.cobeha.2017.07.016.

30. Kumar, A.A. (2021). Semantic memory: A review of methods, models, and current challenges. Psychon. Bull. Rev. *28*, 40–80. https://doi.org/10.3758/s13423-020-01792-x.

31. Irish, M., and Piguet, O. (2013). The pivotal role of semantic memory in remembering the past and imagining the future. Front. Behav. Neurosci. *7*, 27. https://doi.org/10.3389/fnbeh.2013.00027.

32. Eichenbaum, H. (2000). A cortical–hippocampal system for declarative memory. Nat. Rev. Neurosci. *1*, 41–50. https://doi.org/10.1038/35036213.

33. Audrain, S., and McAndrews, M.P. (2022). Schemas provide a scaffold for neocortical integration of new memories over time. Nat. Commun. *13*, 5795. https://doi.org/10.1038/s41467-022-33517-0.

34. Sommer, T., Hennies, N., Lewis, P.A., and Alink, A. (2022). The assimilation of novel information into schemata and its efficient consolidation. J. Neurosci. *42*, 5916–5929. https://doi.org/10.1523/JNEUROSCI.2373-21.2022.

35. Van Kesteren, M.T.R., Rijpkema, M., Ruiter, D.J., and Fernández, G. (2010). Retrieval of associative information congruent with prior knowledge is related to increased medial prefrontal activity and connectivity. J. Neurosci. *30*, 15888–15894. https://doi.org/10.1523/JNEUROSCI.2674-10.2010.

36. Shohamy, D., and Adcock, R.A. (2010). Dopamine and adaptive memory. Trends Cognit. Sci. *14*, 464–472. https://doi.org/10.1016/j.tics.2010.08.002.

37. Brod, G., Lindenberger, U., Werkle-Bergner, M., and Shing, Y.L. (2015). Differences in the neural signature of remembering schema-congruent and schema-incongruent events. Neuroimage *117*, 358–366. https://doi.org/10.1016/j.neuroimage.2015.05.086.

38. Tse, D., Takeuchi, T., Kakeyama, M., Kajii, Y., Okuno, H., Tohyama, C., Bito, H., and Morris, R.G.M. (2011). Schema-dependent gene activation and memory encoding in neocortex. Science *333*, 891–895. https://doi.org/10.1126/science.1205274.

39. Bonasia, K., Sekeres, M.J., Gilboa, A., Grady, C.L., Winocur, G., and Moscovitch, M. (2018). Prior knowledge modulates the neural substrates of encoding and retrieving naturalistic events at short and long delays. Neurobiol. Learn. Mem. *153*, 26–39. https://doi.org/10.1016/j.nlm.2018.02.017.

40. Van Kesteren, M.T.R., Beul, S.F., Takashima, A., Henson, R.N., Ruiter, D.J., and Fernández, G. (2013). Differential roles for medial prefrontal and medial temporal cortices in schema-dependent encoding: From congruent to incongruent. Neuropsychologia *51*, 2352–2359. https://doi.org/10.1016/j.neuropsychologia.2013.05.027.

41. Sekeres, M.J., Winocur, G., and Moscovitch, M. (2018). The hippocampus and related neocortical structures in memory transformation. Neurosci. Lett. *680*, 39–53. https://doi.org/10.1016/j.neulet.2018.05.006.

42. Van der Linden, M., Berkers, R.M.W.J., Morris, R.G.M., and Fernández, G. (2017). Angular gyrus involvement at encoding and retrieval is associated with durable but less specific memories. J. Neurosci. *37*, 9474–9485. https://doi.org/10.1523/JNEUROSCI.3603-16.2017.

43. Greve, A., Cooper, E., Tibon, R., and Henson, R.N. (2019). Knowledge is power: Prior knowledge aids memory for both congruent and incongruent events, but in different ways. J. Exp. Psychol. Gen. *148*, 325–341. https://doi.org/10.1037/xge0000498.

44. Guo, D., Chen, G., and Yang, J. (2023). Effects of schema on the relationship between post-encoding brain connectivity and subsequent durable memory. Sci. Rep. *13*, 8736. https://doi.org/10.1038/s41598-023-34822-4.

45. Webb, C.E., Turney, I.C., and Dennis, N.A. (2016). What's the gist? the influence of schemas on the neural correlates underlying true and false memories. Neuropsychologia *93*, 61–75. https://doi.org/10.1016/j.neuropsychologia.2016.09.023.

46. Berkers, R.M.W.J., van der Linden, M., de Almeida, R.F., Müller, N.C.J., Bovy, L., Dresler, M., Morris, R.G.M., and Fernández, G. (2017). Transient medial prefrontal perturbation reduces false memory formation. Cortex *88*, 42–52. https://doi.org/10.1016/j.cortex.2016.12.015.

47. Kumaran, D., and Maguire, E.A. (2007). Match–mismatch processes underlie human hippocampal responses to associative novelty. J. Neurosci. *27*, 8517–8524. https://doi.org/10.1523/JNEUROSCI.1677-07.2007.

48. Lisman, J.E., and Grace, A.A. (2005). The hippocampal-vta loop: controlling the entry of information into long-term memory. Neuron *46*, 703–713. https://doi.org/10.1016/j.neuron.2005.05.002.

49. Duszkiewicz, A.J., McNamara, C.G., Takeuchi, T., and Genzel, L. (2019). Novelty and dopaminergic modulation of memory persistence: A tale of two systems. Trends Neurosci. *42*, 102–114. https://doi.org/10.1016/j.tins.2018.10.002.

50. Richards, B.A., Xia, F., Santoro, A., Husse, J., Woodin, M.A., Josselyn, S.A., and Frankland, P.W. (2014). Patterns across multiple memories are identified over time. Nat. Neurosci. *17*, 981–986. https://doi.org/10.1038/nn.3736.

51. Lewis, P.A., and Durrant, S.J. (2011). Overlapping memory replay during sleep builds cognitive schemata. Trends Cogn. Sci. *15*, 343–351. https://doi.org/10.1016/j.tics.2011.06.004.

52. Takeuchi, T., Duszkiewicz, A.J., Sonneborn, A., Spooner, P.A., Yamasaki, M., Watanabe, M., Smith, C.C., Fernández, G., Deisseroth, K., Greene, R.W., and Morris, R.G.M. (2016). Locus coeruleus and dopaminergic consolidation of everyday memory. Nature *537*, 357–362. https://doi.org/10.1038/nature19325.

53. Bein, O., and Niv, Y. (2025). Schemas, reinforcement learning and the medial prefrontal cortex. Nat. Rev. Neurosci. *26*, 141–157. https://doi.org/10.1038/s41583-024-00893-z.

54. Botvinick, M.M., Niv, Y., and Barto, A.G. (2009). Hierarchically organized behavior and its neural foundations: A reinforcement learning perspective. Cognition *113*, 262–280. https://doi.org/10.1016/j.cognition.2008.08.011.

55. Collins, A.G.E., and Frank, M.J. (2013). Cognitive control over learning: Creating, clustering, and generalizing task-set structure. Psychol. Rev. *120*, 190–229. https://doi.org/10.1037/a0030852.

56. Frank, M.J., and Badre, D. (2012). Mechanisms of hierarchical reinforcement learning in corticostriatal circuits 1: computational analysis. Cereb. Cortex *22*, 509–526. https://doi.org/10.1093/cercor/bhr114.

57. Badre, D., and Nee, D.E. (2018). Frontal cortex and the hierarchical control of behavior. Trends Cognit. Sci. *22*, 170–188. https://doi.org/10.1016/j.tics.2017.11.005.

58. Rumelhart, D., and Ortony, A. (1977). The representation of knowledge in memory. In Schooling and the Acquisition of Knowledge (Lawrence Erlbaum Associates), pp. 99–135.

59. Rumelhart, D.E., and Norman, D.A. (1976). Accretion, Tuning and Restructuring: Three Modes of Learning. Tech. Rep. 7602 (Center for Human Information Processing and University of California). https://apps.dtic.mil/sti/tr/pdf/ADA030406.pdf.

60. Rumelhart, D.E., and McClelland, J.L. (1986). Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Volume 1: Foundations (MIT Press).

61. Smolensky, P. (1986). Information processing in dynamical systems: Foundations of harmony theory. In Parallel Distributed Processing, Volume 1: Explorations in the Microstructure of Cognition: Foundations, D.E. Rumelhart and J.L. McClelland, eds. (MIT Press), pp. 194–281. https://doi.org/10.7551/mitpress/5236.003.0009.

62. Carpenter, G.A., and Grossberg, S. (1987). A massively parallel architecture for a self-organizing neural pattern recognition machine. Comput. Vis. Graph Image Process *37*, 54–115. https://doi.org/10.1016/S0734-189X(87)80014-2.

63. Grossberg, S. (1987). Competitive learning: From interactive activation to adaptive resonance. Cogn. Sci. *11*, 23–63. https://doi.org/10.1016/S0364-0213(87)80025-3.

64. Drescher, G.L. (1991). Made-Up Minds: A Constructivist Approach to Artificial Intelligence (MIT Press). https://doi.org/10.7551/mitpress/4378.001.0001.

65. Drescher, G.L. (1987). A Mechanism for Early Piagetian Learning. In Proceedings of the Sixth National Conference on Artificial Intelligence - Volume 1. AAAI'87 (AAAI Press), pp. 290–294. https://cdn.aaai.org/AAAI/1987/AAAI87-052.pdf.

66. McClelland, J.L., McNaughton, B.L., and O'Reilly, R.C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. Psychol. Rev. *102*, 419–457. https://doi.org/10.1037/0033-295X.102.3.419.

67. Kumaran, D., Hassabis, D., and McClelland, J.L. (2016). What learning systems do intelligent agents need? complementary learning systems theory updated. Trends Cognit. Sci. *20*, 512–534. https://doi.org/10.1016/j.tics.2016.05.004.

68. Anand, N., and Precup, D. (2023). Prediction and control in continual reinforcement learning. Adv. Neural Inf. Process. Syst. *36*, 63779–63817. https://doi.org/10.5555/3666122.3668908.

69. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, Ł., and Polosukhin, I. (2017). Attention is all you need. In Proceedings of the 31st International Conference on Neural Information Processing Systems (NIPS'17) (Curran Associates Inc), pp. 6000–6010.

70. Yao, S., Yu, D., Zhao, J., Shafran, I., Griffiths, T.L., Cao, Y., and Narasimhan, K. (2023). Tree of Thoughts: Deliberate Problem Solving with Large Language Models. In Advances in Neural Information Processing Systems, A. Oh, T. Naumann, A. Globerson, K. Saenko, M. Hardt, and S. Levine, eds. (Curran Associates, Inc.), pp. 11809–11822.

71. Battaglia, P.W., Hamrick, J.B., Bapst, V., Sanchez-Gonzalez, A., Zambaldi, V., Malinowski, M., Tacchetti, A., Raposo, D., Santoro, A., Faulkner, R., et al. (2018). Relational inductive biases, deep learning, and graph networks. Preprint at arXiv. https://arxiv.org/abs/1806.01261.

72. Hinton, G.E. (1986). Learning distributed representations of concepts. In Proceedings of the Eighth Annual Conference of the Cognitive Science Society, pp. 1–12. https://escholarship.org/uc/item/79w838g1.

73. Sadruddin, S., D'Souza, J., Poupaki, E., Watkins, A., Babaei Giglou, H., Rula, A., Karasulu, B., Auer, S., Mackus, A., and Kessels, E. (2025). Llms4schemadiscovery: A human-in-the-loop workflow for scientific schema mining with large language models. In European Semantic Web Conference (Springer), pp. 244–261.

74. Parciak, M., Vandevoort, B., Neven, F., Peeters, L.M., and Vansummeren, S. (2024). Schema matching with large language models: an experimental study. Preprint at arXiv. https://doi.org/10.48550/arXiv.2407.11852.

75. Dagdelen, J., Dunn, A., Lee, S., Walker, N., Rosen, A.S., Ceder, G., Persson, K.A., and Jain, A. (2024). Structured information extraction from scientific text with large language models. Nat. Commun. *15*, 1418.

76. Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J.D., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., et al. (2020). Language models are few-shot learners. Adv. Neural Inf. Process. Syst. *33*, 1877–1901. https://arxiv.org/abs/2005.14165.

77. Wei, J., Tay, Y., Bommasani, R., Raffel, C., Zoph, B., Borgeaud, S., Yogatama, D., Bosma, M., Zhou, D., Metzler, D., et al. (2022). Emergent abilities of large language models. Transactions on Machine Learning Research. https://openreview.net/pdf?id=yzkSU5zdwD.

78. Akyürek, E., Schuurmans, D., Andreas, J., Ma, T., and Zhou, D. (2023). What learning algorithm is in-context learning? Investigations with linear models. The Eleventh International Conference on Learning Representations, ICLR 2023. https://openreview.net/forum?id=0g0X4H8yN4I.

79. Von Oswald, J., Niklasson, E., Randazzo, E., Sacramento, J., Mordvintsev, A., Zhmoginov, A., and Vladymyrov, M. (2023). Transformers learn in-context by gradient descent. In Proceedings of the 40th International Conference on Machine Learning, A. Krause, E. Brunskill, K. Cho, B. Engelhardt, S. Sabato, and J. Scarlett, eds. (PMLR), pp. 35151–35174. https://proceedings.mlr.press/v202/von-oswald23a.html.

80. Olsson, C., Elhage, N., Nanda, N., Joseph, N., DasSarma, N., Henighan, T., Mann, B., Askell, A., Bai, Y., Chen, A., et al. (2022). In-context learning and induction heads. Preprint at arXiv. https://doi.org/10.48550/arXiv.2209.11895.

81. Ren, J., Guo, Q., Yan, H., Liu, D., Zhang, Q., Qiu, X., and Lin, D. (2024). Identifying semantic induction heads to understand in-context learning. In Findings of the Association for Computational Linguistics: ACL 2024, L.W. Ku, A. Martins, and V. Srikumar, eds. (Association for Computational Linguistics), pp. 6916–6932. https://doi.org/10.18653/v1/2024.findings-acl.412.

82. Xie, S.M., Raghunathan, A., Liang, P., and Ma, T. (2022). An explanation of in-context learning as implicit bayesian inference. In The Tenth International Conference on Learning Representations (ICLR) https://openreview.net/forum?id=RdJVFCHjUMI.

83. Xu, R., Qi, Z., Guo, Z., Wang, C., Wang, H., Zhang, Y., and Xu, W. (2024). Knowledge conflicts for llms: A survey. Conference on Empirical Methods in Natural Language Processing. Preprint at arXiv. https://doi.org/10.48550/arXiv.2403.08319.

84. Xie, J., Zhang, K., Chen, J., Lou, R., and Su, Y. (2024). Adaptive chameleon or stubborn sloth: Revealing the behavior of large language models in knowledge conflicts. The Twelfth International Conference on Learning Representations, ICLR 2024. https://openreview.net/forum?id=auKAUJZMO6.

85. Guerreiro, N.M., Alves, D.M., Waldendorf, J., Haddow, B., Birch, A., Colombo, P., and Martins, A.F.T. (2023). Hallucinations in large multilingual translation models. Transactions of the Association for Computational Linguistics *11*, 1500–1517. https://doi.org/10.1162/tacl_a_00615.

86. Han, Z., Gao, C., Liu, J., Zhang, J., and Zhang, S.Q. (2024). Parameter-efficient fine-tuning for large models: A comprehensive survey. Transactions on Machine Learning Research. https://openreview.net/forum?id=lIsCS8b6zj.

87. Zhao, W.X., Zhou, K., Li, J., Tang, T., Wang, X., Hou, Y., Min, Y., Zhang, B., Zhang, J., Dong, Z., et al. (2026). A Survey of Large Language Models. Front. Comput. Sci. *20*, 2012627. https://doi.org/10.1007/s11704-026-60308-3.

88. Wu, C.C., Tam, Z.R., Lin, C.Y., Chen, Y.N.V., Sun, S.H., and Lee, H.y. (2026). Mitigating forgetting in llm fine-tuning via low-perplexity token learning. Adv. Neural Inf. Process. Syst. *38*, 1708–1744.

89. Biderman, D., Portes, J., Ortiz, J.J.G., Paul, M., Greengard, P., Jennings, C., King, D., Havens, S., Chiley, V., Frankle, J., et al. (2024). LoRA learns less and forgets less. Preprint at arXiv. https://doi.org/10.48550/arXiv.2405.09673.

90. Bittner, K.C., Milstein, A.D., Grienberger, C., Romani, S., and Magee, J.C. (2017). Behavioral time scale synaptic plasticity underlies ca1 place fields. Science *357*, 1033–1036. https://doi.org/10.1126/science.aan3846.

91. Snyder, J.S., Schwiedrzik, C.M., Vitela, A.D., and Melloni, L. (2015). How previous experience shapes perception in different sensory modalities. Front. Hum. Neurosci. *9*, 594. https://doi.org/10.3389/fnhum.2015.00594.

92. Frankland, P.W., and Bontempi, B. (2005). The organization of recent and remote memories. Nat. Rev. Neurosci. *6*, 119–130.

93. van Kesteren, M.T.R., Fernández, G., Norris, D.G., and Hermans, E.J. (2010). Persistent schema-dependent hippocampal-neocortical

connectivity during memory encoding and postencoding rest in humans. Proc. Natl. Acad. Sci. USA *107*, 7550–7555.

94. Momennejad, I. (2025). Memory and Planning in Brains and Machines: Multiscale Predictive Representations. Space, Time, and Memory (Oxford University Press). https://doi.org/10.1093/oso/9780192882547.003.001.

95. Botvinick, M.M. (2012). Hierarchical reinforcement learning and decision making. Curr. Opin. Neurobiol. *22*, 956–962.

96. Kobayashi, S., Schimpf, Y., Schlegel, M., Steger, A., Wolczyk, M., von Oswald, J., Scherrer, N., Maile, K., Lajoie, G., Richards, B.A., et al. (2025). Emergent temporal abstractions in autoregressive models enable hierarchical reinforcement learning. Preprint at arXiv. https://doi.org/10.48550/arXiv.2512.20605.

97. Kim, T., Cochez, M., François-Lavet, V., Neerincx, M., and Vossen, P. (2023). A machine with short-term, episodic, and semantic memory systems. Proc. AAAI Conf. Artif. Intell. *37*, 48–56. https://doi.org/10.1609/AAAI.V37I1.25075.

98. Pritzel, A., Uria, B., Srinivasan, S., Badia, A.P., Vinyals, O., Hassabis, D., Wierstra, D., and Blundell, C. (2017). Neural episodic control. In International Conference on Machine Learning (PMLR), pp. 2827–2836.

99. Khosla, S., Zhu, Z., and He, Y. (2023). Survey on memory-augmented neural networks: cognitive insights to AI applications. Preprint at arXiv. https://doi.org/10.48550/arXiv.2312.06141.

100. Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., Küttler, H., Lewis, M., Yih, W.t., Rocktäschel, T., et al. (2020). Retrieval-augmented generation for knowledge-intensive nlp tasks. Adv. Neural Inf. Process. Syst. *33*, 9459–9474.