

Changes in statistical learning across development

Tess Allegra Forest , Margaret L. Schlichting , Katherine D. Duncan & Amy S. Finn  

Abstract

Statistical learning enables learners to extract the environmental regularities necessary to piece together the structure of their worlds. The capacity for statistical learning and its properties are likely to change across development from infancy to adulthood. Acknowledging this developmental change has broad implications for understanding the cognitive architecture of statistical learning and why children excel in certain learning situations relative to adults. In this Review, we first synthesize empirical work on the development of statistical learning, which indicates that it improves with development only for certain forms of input. Taking inspiration from related cognitive and neural findings, we then consider developmental changes in the properties of statistical learning. Infants and young children might have a broader and less-directed curriculum for learning and represent the outcomes of learning differently from older children and adults. This synthesis offers insight into how developmental changes in statistical learning from infancy through adulthood might fundamentally alter how children interact with, learn about, and remember their experiences.

Sections

Introduction

Development of statistical learning

Neural bases of statistical learning

Broader cognitive development

Development shapes the input to learning

Development shapes representations

Summary and future directions

Introduction

Beginning in infancy, humans rapidly learn a staggering amount about the structure of their environments. This robust learning of environmental regularities, known as ‘statistical learning’, is thought to explain how children learn properties of language^{1–5}, category membership^{6,7} and predictable characteristics of the visual world^{8–11}. For example, children learn which sounds frequently co-occur together to form words in their language (Fig. 1). Statistical learning is widely considered to be available to learners of all ages^{12–14}. However, this availability does not necessarily mean that statistical learning is age-invariant¹⁵. Indeed, statistical learning likely changes in quality with age, operating over different sources of data and producing different representations. Discovering whether and how statistical learning changes with development is likely to inform long-standing questions about why children excel relative to adults in some areas, such as learning language¹⁶ and solving problems that require unconventional solutions^{17,18}.

In this Review, we consider statistical learning across a broad range of development, including infants (0–24 months old), young children (2–7 years old), older children (7–10 years old), and young adults (typically 18–35 years old). We first review existing empirical work on the development of statistical learning, focusing on studies that directly address age-related changes in statistical learning. We define statistical learning as the learning of structure – for example, frequency, probabilities, or mutual information – from repeated exposure to environmental statistics^{19–21} that occurs with no explicit instruction or external feedback (Box 1). Next, we contextualize changes in statistical learning

within brain and cognitive development, considering what inputs are selected for learning and how learned information is represented. Then we consider how these changes might help explain why children surpass adults in certain learning contexts²². Finally, we propose directions for future research to explore changes in the quality of statistical learning across development.

Development of statistical learning

Statistical learning studies typically start with an exposure phase in which learners are presented with a repeating pattern (for example, with nine shapes presented multiple times, occurring in three different triplets) for a short period of time (often just minutes). Importantly, the transitions between all of the elements in the repeating pattern can provide an important cue for segmentation, with dips in transitional probabilities (how often the elements are adjacent) especially important for indicating which items should be grouped together (such as syllables in a word or shapes in a triplet).

Learners are then presented with stimuli that conform to the exposed pattern (for example, a sequence of three shapes in the previously presented order) or ‘foils’ that do not conform to the exposed pattern (for example, the same shapes in a rearranged order). Foils can take various forms, including so-called position-matched foils in which each element (shape or syllable) appears in the same position as it did during exposure (Fig. 2a) and so-called part-word foils (Fig. 2b) in which part of the original triplet/word remains intact. Typically, auditory experiments use part-word foils and most visual experiments

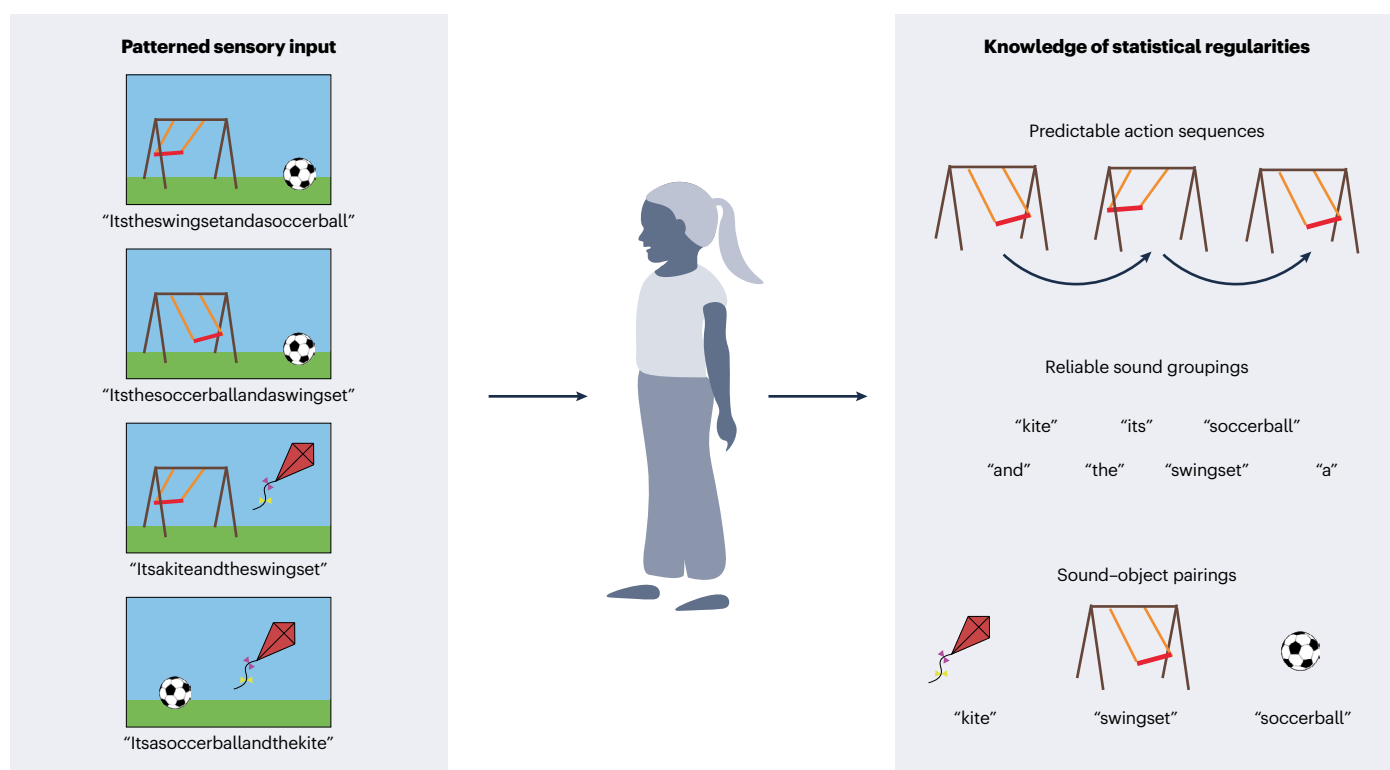


Fig. 1 | Statistical learning in the real world. Learners experience many regularities in different sensory modalities over time (left). For example, playground swings move in a particular sequence (backward–forward–backward–forward), sounds within words frequently appear together (the syllables ‘so’ and ‘cer’ appear

consecutively more often than ‘the’ and ‘so’) and objects co-occur with the words used to describe them (the word ‘kite’ is usually heard in the presence of kites). Even with no explicit instruction to learn, no reinforcement, and no external feedback, a learner can translate this input into knowledge of regularities (right panel).

use position-matched foils. A difference in response between trials of these two types of stimuli indicates learning of the pattern. Humans are able to learn various structures that operate over many different kinds of stimuli, across diverse modalities in very short periods of time^{7,23,24}.

Infants – even neonates¹³ – are adept statistical learners^{15,19,25–27}. In experimental contexts, infants can learn many different regularities across various different stimuli and both visual and auditory modalities^{1,3,28–30} (Fig. 2). Evidence for developmental change in statistical learning ability is mixed. An important factor that gives rise to this variability is how learning is measured: directly or indirectly (Table 1). Learning on direct measures is evaluated by asking participants to make assessments such as old/new memory judgements, or to choose between an exposure-conforming or non-conforming sequence in a two-alternative forced-choice (2AFC) test^{8,24,31,32} (Fig. 2a). Learning on indirect measures is inferred without overt report from the participants³³ (Fig. 2b) and can be used to measure learning during the learning process itself. Importantly, indirect tests can be used in infants through adults, whereas direct tests can only be used in participants who can be asked to overtly make assessments, typically starting around age 2 years.

Indirect measures

Indirect measures suggest that statistical learning changes little or not at all with development. These measures include reaction time, saccade latencies, looking times, and neural responses recorded while participants engage in learning or a post-learning task. Performance on several indirect measures – including reaction times to elements that follow a predictable versus unpredictable transition from exposure as measured in a post-test^{10,32,34}, or from participants pacing their own exposure^{35,36} – have not been compared across age groups. However, other measures have been used in infants, children, and adults, directly comparing their performance. In one study, eye movements of infants (7–11 months old), children, and adults (6–30 years old) were recorded while they were visually presented with a central stimulus (a cartoon octopus or turtle) that predicted both the identity and the spatial location of subsequent targets. Infants' saccade latencies revealed sensitivity to these statistical relationships in a manner that was on par with the children and adults, suggesting no age-related change in learning³⁷.

In another indirect measure known as 'neural entrainment', large-scale electrical oscillations in the brain are measured to determine their synchrony to the rhythm of structured stimuli. Notably, electrical oscillations in infants^{14,38} (even sleeping neonates³⁹) and adults⁴⁰ align with the rhythm of regularly grouped syllables, tones³⁰ and visual stimuli⁴¹ after very little exposure to the regularity. Further work has directly compared neural entrainment rates in infants and adults, and found that infants' (6-month) neural entrainment to reliable syllable sequences increased at the same rate as that of adults¹⁴. Functional near-infrared spectroscopy – a neural measure of differences in blood oxygenation – also revealed that neonates have a different brain response for consistent versus inconsistent syllable sequences after exposure to a structured stream of syllables⁴². These studies, particularly the saccade latency and entrainment findings that compare age groups directly, suggest that statistical learning does not change with development.

Direct measures

By contrast, direct measures suggest developmental change in statistical learning ability. In two studies using visual stimuli, performance on a 2AFC post-test improved across childhood: in one study between ages 5 and 12 years⁴³, and in another across three age groups³¹ (6–11 years, 12–17 years, 18–30 years). A very similar developmental trajectory was

Box 1

Definition of statistical learning

In this Review we define statistical learning as the learning of structure from repeated exposure to environmental statistics^{19–21}. Similar to others¹⁹, we further require that learning occurs without reinforcement or external feedback. This feature is central to the potential of statistical learning to answer puzzles such as how humans learn language with little, if any, feedback¹⁸³. We exclude motor-based skill learning (such as serial reaction time tasks) from statistical learning, because it is typically assessed in tasks that provide external feedback^{55,184}, generally requiring correct performance to progress through the task. However, we consider the motor learning literature to contextualize the development of statistical learning.

We consider statistical learning to be uninstructed and incidental to learners' initial goals, but the resulting statistical knowledge does not need to be unconscious. Although some definitions conceptualize statistical learning as a singular and inherently implicit learning system¹⁸⁵, we favour the conceptualization of statistical learning as supported by a set of learning processes, some of which can result in conscious access to learning^{20,21,34}. This conceptualization accords well with the mounting evidence that learners often have meta-awareness of structure in classic statistical learning paradigms^{32,34} and a broader recognition that the same underlying neural mechanisms can support both conscious and unconscious learning¹⁸⁶.

also found using non-linguistic auditory stimuli (recognizable sounds such as a door opening or bell ringing): 2AFC performance improved in children aged from 5 to 12 years⁴⁴. However, children and adults performed equally well on a 2AFC post-test when the learning stimuli were aurally presented syllables⁴⁴. A comparison of the developmental trajectories of visual and aural-syllable statistical learning in separate groups of children aged from 5 to 12 years found age-related improvement in visual but not aural-syllable statistical learning performance²³. Other studies have similarly shown no age-related improvement in aural-syllable statistical learning between 5 and 8 years old⁴⁵, and when comparing 6–7 year olds with adults²⁴. Considering everything other than aural syllables, there is improvement in performance on direct measures across development.

The lack of developmental change in aural-syllable statistical learning is a puzzle. This pattern could be to do with young children's already extensive experience with syllables relative to other stimuli. Another explanation could be a possible innate (or evolutionarily driven) predisposition to process linguistic information^{16,46}. Ultimately, much future work – including both direct and indirect measures of statistical learning for other stimuli that infants are predisposed to (such as faces) – is needed to understand these stimulus-specific differences.

Neural bases of statistical learning

The neural regions that have been associated with statistical learning include the prefrontal cortex (including the medial prefrontal cortex and inferior frontal gyrus), sensory cortices, the hippocampus, and

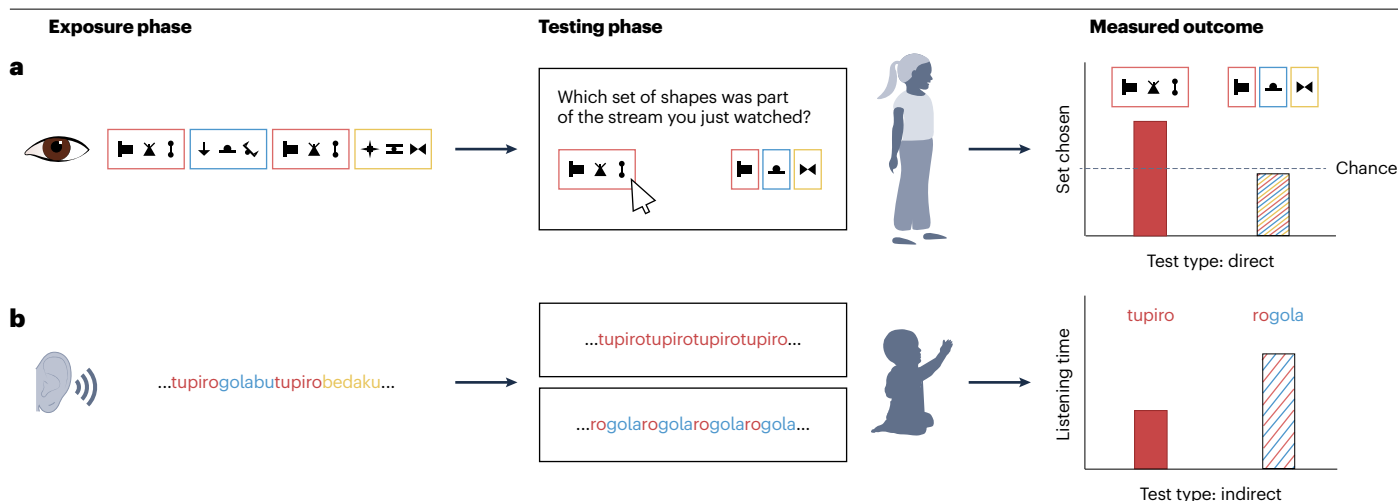


Fig. 2 | Statistical learning paradigms. **a**, In a typical visual paradigm, shapes are grouped into recurring triplets (outlined in unique colours for visualization purposes) and presented to a learner for at least a few minutes. These shapes have been used in many studies (including with infants²⁸) but the illustrated paradigm is most similar to a study with adults¹⁸². In the test phase, learners are presented with an exposed triplet and a rearranged group of shapes – a position-matched foil here – and asked to choose which they believe was part of the stream they just watched. How often the triplet is correctly selected over the foil is a direct measure

of statistical learning. **b**, In a typical auditory paradigm, syllables (for example ‘tu’, ‘pi’ and ‘ro’) are grouped into nonsense words (‘tupiro’, in red) and presented sequentially (with each word following another with equal probability) for at least a few minutes. These nonsense words have been used in many studies (including with adults²⁴) but the depicted paradigm is closest to an early study with infants¹. In the test phase, one exposed word or a reorganized set of syllables (a part-word foil in the figure) are presented on a loop. Differences in listening times to the word loop versus the foil loop are an indirect indication of learning in infants.

the basal ganglia (Fig. 3). We review the evidence for these region by region, particularly noting developmental change.

In the prefrontal cortex, the inferior frontal gyrus is more active during exposure to structured than random sequences^{47–49} in both adults and children⁵⁰. In adults, the extent to which the inferior frontal gyrus is more active during structured sequences correlates with later statistical learning performance^{6,47}. Although the involvement of this region in statistical learning is most frequently discussed in relation to its demonstrated role in language (the left inferior frontal gyrus is also known as ‘Broca’s area’)^{48,49}, it is also involved in statistical learning of visual information^{6,51}. Engagement of the inferior frontal gyrus might reflect sensitivity to specific relationships within a structured stream: its engagement increases with more exposure to same-group items and its patterns of responses reflect potential transitions between groups in visual statistical learning⁶.

The medial prefrontal cortex is also implicated in statistical learning in both infants⁵² and adults. In contrast to the inferior frontal gyrus, in adults the medial prefrontal cortex shows consistently high activation for all transitions within a group, with deactivations at group boundaries⁶. Thus, whereas the inferior frontal gyrus records specific, experienced transitions, the medial prefrontal cortex might play a more general role by signalling meaningful statistical boundaries⁶. Such a dissociation would be consistent with past memory research suggesting that the inferior frontal gyrus represents specifics⁵³, whereas the medial prefrontal cortex preferentially extracts higher-order structure of experiences^{53–57}.

Sensory cortices (primary visual and auditory regions⁵⁸ that are regularly involved in priming^{20,58–60}) are likewise engaged more for structured than random information across visual and auditory statistical learning^{48,61}. One possibility for the engagement of sensory cortices in statistical learning is that they are involved in processing the stimuli

themselves. However, the possibility has also been discussed²⁰ that these regions give rise to statistical learning directly⁴⁷. Indeed, learning in these sensory regions could take place via cortical plasticity⁶²: cortical networks that repeatedly experience structured input will gradually become tuned to it, therefore reflecting structure learning directly^{20,63}.

Activation of the basal ganglia^{58,62} – which are regularly implicated in skill learning^{64,65} and probabilistic reinforcement learning⁶⁶ – has also been observed in multiple studies of statistical learning in adults, including auditory linguistic⁴⁸ and visual^{6,47} tasks. Notably, the involvement of the basal ganglia in statistical learning is not always observed in adults⁴⁵ and has been observed in only one⁵⁰ of four statistical learning studies with infants or children^{31,45,52}. When seen, their involvement could be related to the role of the basal ganglia–dopamine system in responding to reinforcement signals^{67,68} and/or internally generated feedback⁶⁶. The basal ganglia are also involved in gating or prioritizing certain information (such as the structured information of repeating triplets) for subsequent processing in working memory^{69–72}.

Finally, the medial temporal lobes, particularly the hippocampus⁷³, are involved in adult^{74–76} and infant⁵² statistical learning. In adults, items that reliably occur together (such as a pair of shapes) come to be represented as more similar by the hippocampus, perirhinal cortex, and parahippocampal cortex^{6,51,75} following learning. Importantly, despite the involvement of the hippocampus in statistical learning, the resulting memories are not necessarily consciously accessible^{51,75,77}, suggesting that the role of the hippocampus in statistical learning is linked to its capacity to rapidly bind associations across time^{77–79}, and not its role in producing memories that are consciously accessible (Box 2).

Developmental neurobiological investigations of statistical learning are quite rare. There is just one functional neuroimaging study in infants, which implicates the hippocampus and medial prefrontal cortex in visual statistical learning³². There is also only one functional

neuroimaging study in children (aged 9–10 years), implicating the inferior frontal gyrus⁵⁰ in auditory statistical learning. There are also two structural studies that are quite consistent with each other, showing that the size of young children’s inferior frontal gyri⁴⁵ and hippocampi^{31,45} relate to their statistical learning abilities even after controlling for age.

Conspicuously missing from the list of regions underpinning statistical learning during infancy and childhood are sensory cortices and, often, the basal ganglia. This absence is striking given the prominence of these regions in priming and skill learning⁵⁸ and the relatively early maturation of sensory cortical regions⁸⁰. However, a confident appraisal of potential developmental shifts in the neural structures that are involved in statistical learning cannot be made from just two studies that have measured the function of specific brain regions.

Later in this Review, we focus on regions that are involved in infant and child statistical learning – the hippocampus, inferior frontal gyrus, and medial prefrontal cortex – to draw on insights from the development of these regions for understanding developmental shifts in what is represented in memory after statistical learning. Of note, immaturity does not mean inactivity; immature function might alter how information learned via statistical learning is represented. To ground this discussion, we first discuss the development of related cognitive abilities that influence statistical learning.

Broader cognitive development

Although our definition of statistical learning does not include attention or classically ‘implicit’ learning processes (such as skill learning due to external feedback; Box 1), these fields are highly relevant for understanding statistical learning. Attention likely plays a role in statistical learning^{10,81–83} and there is large conceptual overlap with other implicit learning processes²⁰. In particular, a consideration of attention informs possible changes in what gets selected for learning and the

consideration of implicit learning processes informs possible changes in the pace of statistical learning and the durability of learned information. The operation of multiple memory systems for statistical learning is also relevant (Box 2).

Human attentional capabilities change greatly across development^{84,85}, extending over long periods of time from infancy into adolescence^{84,86,87}. Across early and late childhood especially, children significantly improve in their ability to selectively attend to one feature of their input⁸⁶ while ignoring distracting information^{87,88} (Fig. 4a).

This relatively slow development of attention (Fig. 4b, yellow line) likely matters quite a lot for statistical learning. In adults, attention can constrain what is learned via statistical learning and some researchers claim that attention to structured information is necessary for statistical learning to take place^{10,81}. Given the ongoing development of attention^{84,86}, the input to statistical learning is therefore likely to be less constrained in children.

Classically implicit learning processes (including priming and skill learning) develop very early in life, with observations of mature perceptual priming in children as young as 3 years old^{89,90}. Skill learning (measured by serial and probabilistic serial reaction time tasks) is typically mature by the age of 4 years⁹¹ (although some studies show more protracted development⁹²). Indeed, one study showed that children aged 4–12 years had better skill learning of probabilistic sequences than learners ranging in age from 13 to 80 years⁹³. This pattern of early maturity is notable given that other aspects of human memory, such as the ability to remember specific events, show marked and continuous improvements from infancy⁹⁴ and young childhood^{95,96} all the way through adolescence⁹⁷.

A consideration of the development of implicit learning processes adds context to understanding developmental change in statistical learning. Indeed, these implicit learning processes are conceptually quite similar to statistical learning, sharing many definitional features – including

Table 1 | Developmental patterns in statistical learning

		Type of input		
		Auditory linguistic	Auditory non-linguistic	Visual
Indirect measures	Evidence	Neonates show differential brain responses to consistent versus inconsistent statistical information ³⁹ 8-month-old infants ³⁸ , 6-month-old infants and adults show similar neural entrainment to syllables ¹⁴ Children ⁵⁰ activate similar brain regions as adults ^{48,49} when listening to artificial languages	Neonates segment tone sequences into triplets as adults do, measured by electroencephalography ³⁰	Infants activate some of the same brain regions when viewing structured sequences ⁵² as adults Infants, children and adults show similar saccade latencies to predictable locations ³⁷
	Implications	Some overlap in neural regions or processes involved in statistical learning over development	Some overlap in neural processes involved in statistical learning over development	Some overlap in neural regions involved in statistical learning over development
	Limitations	Similar neural activation does not necessarily imply equivalent learning or neural functionality		
Direct measures	Evidence	Similar performance across young children (5 years old ³³), older children and adults ³⁴ in 2AFC tests	Improvement on 2AFC tests from young children to adults ⁴⁷	Improvement on 2AFC tests from young children to adults ⁴⁶ Improvement with age on old–new tests of visual statistical structure ¹⁷³ and triplet completion ⁷
	Implications	No developmental improvement in auditory linguistic statistical learning	Quality of non-linguistic statistical learning changes with age Potential involvement of increasing metacognitive abilities	Visual statistical learning improves with age Potential involvement of increasing metacognitive abilities
	Limitations	Cannot measure qualitative changes by examining improvement alone		

Note: Indirect tests are applicable for use in all age ranges (infants, young children, older children and adults) whereas direct tests are most applicable to learners after infancy. 2AFC, two-alternative forced-choice test.

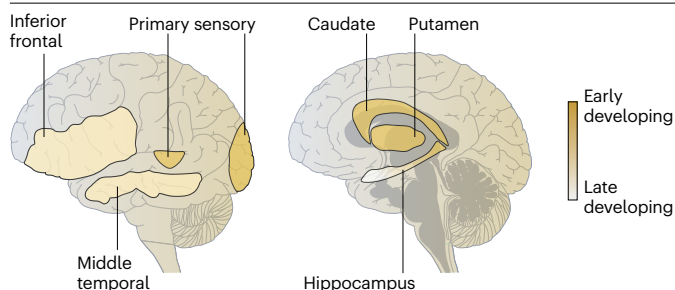


Fig. 3 | Neural underpinnings of statistical learning. Left: Inferior frontal cortex, middle temporal cortex, and primary sensory cortices (auditory and visual). Right: Basal ganglia (caudate and putamen) and hippocampus. Regions are shaded by their rate of development (with earlier developing areas in darker yellow). Hippocampal development is likely to also be relevant for statistical learning, indicated here with posterior hippocampus (dark yellow) developing earlier than anterior hippocampus (light yellow).

learning from repeated exposure without explicit instruction to learn. To the extent that these processes overlap with statistical learning and that statistical learning results in the formation of implicit memories, the learned representations that emerge from statistical learning will likely align closely with known properties of implicit memories. Implicit memories are learned slowly⁶², are context specific^{98,99}, and are durable over time^{100–102}. Due to the early development of implicit learning, these properties should be similarly true in statistical learning across ages – in infants through adults (Fig. 4b, blue line).

With this background about the development of statistical learning and relevant cognitive and neural changes, we now integrate data that bear on potential qualitative changes to features of statistical learning input and representations. Although this discussion builds on previous observations about developmental changes in statistical learning, it is largely speculative and remains to be tested. As we summarize in the final section, future work should seek to test these principles with indirect tests (if testing learners from infants through adults) and direct tests (if testing only children through adult learners).

Development shapes the input to learning

It has been noted elsewhere that the curriculum for statistical learning changes vastly with development¹⁰³. That is, infants' and children's sensory input is not the same as adults' sensory input, even when they are in the same environment. For example, infants who do not yet crawl see mainly faces, whereas crawlers see different objects from walkers, whose viewpoint differs from that of adults¹⁰⁴. Correspondingly, different information is available to learners with these changing physical statures and skills, directly shaping what sensory input is even eligible for statistical learning across development. Beyond the input itself, attention might impact what input is available to statistical learning, and developmental changes in attention might influence the development of statistical learning.

Attentional focus

Ongoing development of attention should give infants and children broader access to the statistical structure in their environments⁸⁷. Although immature attention might be a hindrance to children when trying to focus on one aspect of their environment, immature attention

might have the unexpected benefit of leading children to have greater knowledge of the non-goal-directed aspects of their experiences. For example, when asked to attend to one of two overlaid shapes, 5-year-old children, but not adults, were able to detect a change in the unattended shape⁸⁷. Similarly, 4-year-old children recalled more items than adults when asked to think about the category of the items during encoding⁸⁸. As attention matures and narrows, children lose the ability to notice changes in unattended items and use broad information to aid memory. Because children cannot use attention to filter their experience as well as adults, they can potentially learn more about the parts of experience that adults exclude for being outside their narrower focus of attention.

In line with this evidence from outside statistical learning, having better attention should create a narrower input to statistical learning for adults, relative to infants and children. Attention can gate^{10,81} statistical learning in adults such that unattended aspects of an environment are not always learned. Consequently, adults might have less to learn from at any given moment in time. By contrast, infants and children could have more options for what to learn at any given moment in time, potentially meaning they are less likely to hone in on one source of structure, learn from it, and move on to the next. In a complex environment with multiple structures, infants and children might therefore spend more time learning from multiple sources at once, slowing their learning of any given structure, but also potentially bolstering their knowledge of possible relationships among different sources of structure.

Attention gates statistical learning in both auditory and visual modalities in adults^{10,81,82}. However, the role that developing attention plays in shaping the input to statistical learning could play out differently in the two modalities. In audition, information can be fleeting and sources of structure might be difficult to discretize (such as by using space). By contrast, in vision, information often persists for longer and can be separated in space. These differences could manifest in modality-specific developmental timelines in how attention shapes learning. That said, much work is needed to further understand this speculative link between attention and modality-specific differences in statistical learning across development (Table 2).

Beyond attention, other aspects of development – such as physical development – constrain and pace the content of input to all learning systems, including statistical learning¹⁵. Attention and other aspects of development likely interact to produce learning in infants and children that is truly unique, constrained by their physicality and broadly spread across the input they do receive. Future work is greatly needed to explore these interactions.

Prior knowledge

Beyond interactions with attention, differences in knowledge of and experience with particular stimuli might also impact statistical learning directly, and indirectly via its ability to shift what a learner attends to^{105,106}.

In adults, knowledge seems to impact statistical learning: adults fail to learn structures that conflict with their previous knowledge. For instance, adult learners exposed to syllables comprising sound combinations that never occur in their native language did not learn them^{107,108}. These findings suggest that it might be difficult to learn regularities that conflict with one's previous experience. As yet, cross-linguistic statistical learning studies in infants (who have less experience with language) have not been performed, but less knowledge of their native language (coupled with greater sensitivity to a broad range of phonemes²⁹) would likely permit learning to occur.

Given that speech-based statistical learning shows a unique developmental pattern (in direct measures), future work outside the linguistic domain is needed to fully understand the role of prior knowledge on statistical learning across ages.

Knowledge can also impact learning via its effect on what learners attend to. Indeed, adults typically attended to surprising¹⁰⁹ or novel¹¹⁰ information over what is already familiar and known^{109,111–115}. Consequently, when presented with multiple statistical regularities to learn, learning and experience should shift one's attention as the yet-to-be-learned information will be most novel. Because adults' attention is more developed than children's, these experience-driven or learning-driven shifts in attention are likely to constrain what adults learn, more so than for children.

Some evidence that knowledge and attention interact during learning in adults comes from work on the statistical learning of aurally presented syllables. When statistical structure changes (without a cue) halfway through exposure, adults fail to learn the statistics present in the second half¹¹⁶. Importantly, because both the first-half and the second-half structures contain the same elements (syllables), the halted learning of the second structure could reflect the waning of attention (due to the lack of novelty) following the initial learning of these syllables when they were part of the first structure. Learners of all ages tend to be curious about things that they do not already know^{113–115}. With their attention in search of new things, they might stop attending to the familiar syllables and, therefore, miss the new structure hidden in the same syllables. In support of this learning-based explanation,

a follow-up study tested learning of one structure every minute, and switched the structure immediately after evidence of learning. This paradigm permitted learning of both structures¹¹⁷, potentially by preventing the waning of attention following successful learning. However, additional work is needed to directly measure whether attention wanes with knowledge and whether shifts in attention directly predict learning. Interestingly, when participants were exposed to random visual information prior to structured, their learning of the structured information was not hindered¹¹⁸, suggesting that learning (or having something to learn) might be key for directing attentional shifts.

In further support of the idea that attentional processes interact with learning or structured experience, another study shows that as adults gain experience in an environment, they shift their attention from less complex streams of information to more complex streams¹¹¹. One explanation for this behaviour has to do with learning: as knowledge is gained, adult learners shift their attention to as-yet-unlearned stimuli. Although the role of experience in shifting infants' and children's attention during statistical learning has yet to be explored, infants^{119,120} and children¹²¹, similar to adults, prefer to look at information of 'medium' complexity (usually defined as neither the highest nor lowest entropy source of information available). Infants' attention (how long they look at a statistical event) is also predicted by how much they can learn from that event¹²². These findings strongly suggest that infants through adults are equipped to monitor how useful a particular structure is. Useful structures might contain just enough novel, unlearned, information.

Box 2

Lessons from multiple memory systems

Work in adults informs how implicit and explicit memory might interact during statistical learning. Adults can demonstrate statistical learning on direct and indirect assessments (thought to reflect explicit and implicit memory, respectively^{10,34}) for the same material^{34,187}. Although it could be argued that implicit and explicit processes compete during statistical learning¹⁸⁸, performance on explicit and implicit assessments tends not to be correlated across participants³⁴, suggesting instead that both systems operate in parallel during statistical learning, without trading off. However, test items classified as poorly learned on direct assessments elicit stronger evidence of learning on indirect assessments¹⁸⁷, perhaps reflecting competition between underlying learning systems. In other cases, the systems might cooperate, with early implicit learning driving attention towards statistical regularities^{189,190}, and potentially supporting conscious awareness of regularities when sufficiently strong²⁰.

These complex interactions are mirrored in the multiple memory system literature. Evidence for double dissociations between implicit (basal ganglia-based) and explicit (hippocampal-based) learning systems — including from the landmark case of patient H.M.⁷³ — paint a picture of independence^{191–194}. However, in rodents, lesioning one system can improve performance on tasks relying on the other^{195–199} — strong evidence that hippocampal and striatal memory systems compete (for a review see²⁰⁰). Two studies in humans also suggest

a trade-off between hippocampal and striatal systems^{188,201}. Other research is consistent with cooperative interaction. Hippocampal activity has been positively related to feedback-based learning when the task is fully attended to²⁰², when feedback is delayed by more than a few seconds²⁰³, when the stimuli are subject to high interference²⁰⁴, when learning depends on binding together multiple stimuli^{205,206}, or when drawing inferences across separately learned relationships^{207,208}. These varied circumstances align with prominent views of hippocampal function that emphasize its capacity to bind associations across space and time⁷⁹. Likewise, the striatum has been implicated in declarative memory retrieval (for a review see²⁰⁹). Much of this research finds task-related functional connectivity between the hippocampus and striatum, further arguing for the possibility of cooperative interactions.

There are two important lessons for statistical learning from this literature. First, interactions between memory systems are common but the nature of the interaction depends on the content, structure, and timing of what is being learned. Second, the mappings of brain region to function are not entirely consistent across studies. Rather than focusing on conscious accessibility, these systems should be characterized based on the specific processes they contribute to, even beyond memory⁷⁷. Our Review implements these lessons by highlighting stimulus-level differences and turning to the brain for mechanistic insights about processes and representation.

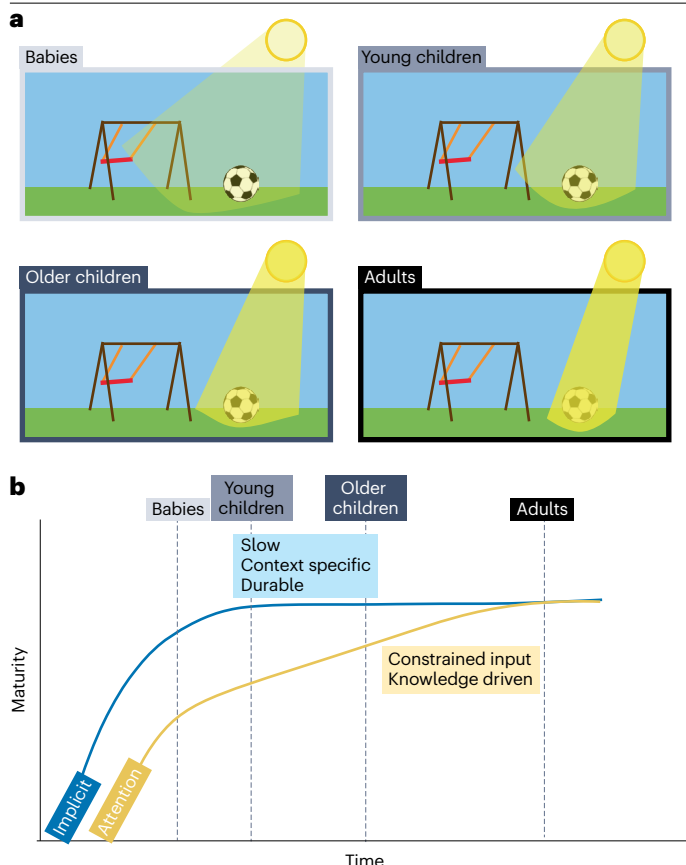


Fig. 4 | The development of attention and implicit learning processes. a, Shifts in how attention changes with development, going from more broad to narrow. **b,** A theoretical trajectory for each system, with implicit learning processes (blue) maturing very early in life and attention (yellow) maturing much later, into adulthood.

However, infants' and children's ongoing attentional development could mean that they both attend to more at once and shift their (larger) attentional focus more slowly than adults when there are multiple structures to learn. Not only is their attention more distributed – meaning that they are less likely to home in on one source of structure at a time – but having less knowledge means that more of what they are exposed to is likely to be novel or surprising, which could alter the pace of attentional shifts that do occur. Children's broader focus and reduced knowledge therefore means that if and when attention constrains their input, it might be guided less by experience, making the factors that determine what children will attend to and learn from potentially less predictable than in adults and more likely to be driven by the environment.

These intriguing predictions provide testable ways of expanding knowledge of statistical learning by linking existing research on child development with existing research on the known cognitive mechanisms that interact with statistical learning. Some of these predictions differ slightly by input modality or test method (Table 2). Ultimately, much future work is needed to understand the role of attention in shaping statistical learning and how knowledge interacts with attention to gate the input to statistical learning across development.

Development shapes representations

Beyond shaping the input, developmental change is also likely to alter the representations that are formed following statistical learning. In particular, the balance of general versus specific representations is likely to shift with age. Evidence from literature outside statistical learning suggests that highly specific memory representations seem to emerge around age 2 years and become more precise across childhood^{123–129}. There are at least two potential routes to general representations that – although behaviourally similar – could emerge at different points in development and for different reasons. The first type of general representation might emerge in infancy and could be due to neural maturation or forgetting¹³⁰; we refer to the resulting representations as fuzzy. The second type of general representation might emerge later in childhood^{123,131,132}, reflect commonalities across events and experiences, and exist alongside specific information; we refer to the resulting representations as broad.

Fuzzy and broad representations might be behaviourally indistinguishable, but they seem to have distinct developmental time courses. Considering these representation types alongside the development of specific representations seems to predict non-linearities in which representations are likely to be predominant and guide behaviour across development.

Fuzzy representations

Representations of structure supporting above chance statistical learning in infancy might predominantly be fuzzy, an idea which has some preliminary empirical support¹³³, for two potential reasons: infants' fast pace of forgetting and/or the immaturity of their neural circuitry.

Outside statistical learning, models of learning and memory have highlighted that forgetting can produce fuzzy, general memories^{134,135}. According to fuzzy trace theory, one route to forming fuzzy representations is through forgetting – as unique details of particular experiences fade, learners retain only parts of their memory for that experience¹³⁰. Over many similar experiences, the remaining memories are likely to reflect the most common elements of those experiences, and therefore can be thought of as general¹³⁴. Related work has shown that the 'spacing effect' can have a similar effect on representations: spacing the presentation of different items from the same category out over time – which leads to greater forgetting between presentations – boosts learners' general knowledge of the presented information in adults¹³⁶ and children¹³⁷. Thus, faster forgetting – and greater forgetting between experiences – can produce fuzzy, general representations (Fig. 5).

Forgetting impacts all learners' representations, from infants to adults. But, crucially, infants^{138,139} forget much more quickly than young children, older children and adults, an effect that is more pronounced for sequential information (information that unfolds over time, as in many studies of statistical learning)¹⁴⁰. As a consequence, infants might demonstrate fuzzy knowledge of their statistical experiences very early after learning, with these fuzzy representations possibly being the only thing that is observed in an immediate post-test.

Patterns of brain development are a second reason that the representations that infants learn are likely to be fuzzy. In particular, the complementary learning systems framework details that cortical learning systems support slowly learned general representations and hippocampal learning systems support quickly learned specifics¹⁴¹. Cortical learning systems are available to infants, and are thought to support early memories⁹⁴. Modelling work on cortical learning¹⁴¹ notes that the cortex requires many similar experiences to learn,

without storing the specifics of those experiences, which leads to fuzzy representations after extensive exposure (Fig. 6, yellow line).

In addition, hippocampal-mediated learning in infancy⁵² might result in fuzzy representations. Although the hippocampus is known for supporting fast-specific learning, modelling work suggests it can also support slow, general learning that is very similar to cortical learning⁷⁸. Subfields of the hippocampus perform different functions¹⁴² and give rise to different representations through two pathways that connect them^{94,135}. The less complex monosynaptic pathway develops from infancy and the more complex trisynaptic pathway develops later (during toddlerhood, thought to be at around 18–24 months in humans based on monkey work)^{143,144}. The early-developing monosynaptic pathway is thought to learn regularities from many similar experiences, yielding fuzzy representations in humans and monkeys alike⁷⁸.

Much of the data on these exact timelines come from histological studies in animals and behavioural human memory development research, so these possibilities for statistical learning require confirmation in human infants. However, the high rate of forgetting, the availability of cortical learning systems and the early development of the monosynaptic hippocampal pathway converge to suggest the possibility that the representations infants learn from statistical regularities might be fuzzy (Fig. 6, yellow line).

Specific representations

After infancy, as forgetting slows and hippocampal learning systems mature, the outcome of statistical learning might shift, resulting in specific and broad representations. The trisynaptic pathway of the hippocampus – which yields rapid encoding of specific memories – is not thought to emerge in humans until around 18–24 months of age, in the shift from infancy to young childhood^{143,145}. The trisynaptic pathway continues to undergo refinement for years thereafter^{129,146,147} (Fig. 6, red line). A great deal of behavioural data on human memory development also suggest that specific-episode memories likely emerge around age 2 years, although their prevalence increases into adulthood^{95,96,148}. Thus, learners from young childhood onward should be able

to represent and retain specific information. These specific representations are likely underpinned by hippocampal maturity in the tri-synaptic pathway^{149,150}, but other relevant changes in memory are also at play¹⁵¹.

With an ability to represent specific information thanks to the functionality of the trisynaptic pathway (Fig. 6, red line) and, possibly, slightly reduced fuzzy representations relative to infants (Fig. 6, yellow line) due to reduced forgetting, it is possible that early childhood might be a unique period in which specific representations are the most likely to guide behaviour.

Broad representations

As learners enter late childhood, additional factors should permit the representation of general information alongside specific¹³¹. Other aspects of hippocampal development can support the formation of broad representations that emerge in older children (roughly 7 years old, although the likelihood of forming a broad memory increases into adolescence^{96,97}; Fig. 6, blue line). Indeed, across species there is heterogeneity of development along the hippocampal long axis (anterior to posterior) throughout childhood and adolescence^{31,152–155} – an important dimension that aligns with how much detail adults represent in their memories^{53,156,157}. The posterior hippocampus (thought to specialize in fine-grained, specific, representations of individual experiences) appears to functionally develop several years earlier¹⁴⁹ than the anterior hippocampus^{155,158}. This later anterior development could be important for the formation of broad representations because of its later-developing connections with the medial prefrontal cortex¹⁵⁹, a region thought to house broad or schematic knowledge^{56,112}. Together, this development might afford learners increasingly flexible access to broad memories, such as commonalities across events, alongside specific memories (Fig. 6, blue line) as they progress through adolescence^{156,160,161}. Also notable for this timetable is that hippocampal coupling with the inferior frontal gyrus precedes that of the medial prefrontal cortex¹⁵⁸; this coupling might bias younger children towards first recording the specifics of their experiences and, then increasingly accentuating the broader properties of an experience across development^{6,51,53,55,57,162}. We therefore

Table 2 | Qualitative changes in statistical learning

Type of input	Input to learning	Learned representations
Auditory linguistic	The input to learning might shift with short-term and accumulated long-term experience, as a function of environmental complexity or developing attention Prior experience with language is likely to heavily bias representations formed (and therefore future input) more than for other types of input	Memory representations might change with age, potentially with a bias towards fuzzy representations even in adults because of the fleeting nature of auditory stimuli Because representations might take many experiences to accumulate, they will likely be quite durable
Auditory non-linguistic	The input to learning might shift with short-term and accumulated long-term experience, as a function of environmental complexity, or developing attention In simple situations, all ages might attend to similar input In complex situations where a learner can choose what to focus on, adults might show more focused learning than children and infants	The representations formed might be quite fuzzy because of the fleeting nature of auditory stimuli, so the time course of representations is key to study Because fuzzy representations are built over long stretches of time, they are likely to be quite durable
Visual	The input to learning might shift with short-term and accumulated long-term experience, as a function of environmental complexity, or developing attention, especially when selective attention is required Past experience likely shapes the information selected for learning in all ages The pace of learning might differ across age groups due to changes in the use of top-down learning, with adults showing fast learning of one thing at a time and infants and children learning more information slowly	The output of visual statistical learning might differ across ages such that infants represent different structures than adults and children, but much research is still needed The representations formed for spatial visual statistical learning might be influenced by the narrowness of the learner's attention Specific representations might not be very durable



Fig. 5 | Greater forgetting can lead to more general representations. The forgetting rate shapes the memories a learner forms for category information. For any learner, time since experiencing a unique bike determines how well they remember that bike: the longer it has been since that experience, the less they will remember it (illustrated by fading). Their forgetting rate determines their overall memory for the concept ‘bike’ when they encounter another bike. **a**, Learners with fast forgetting rates (especially true for infants) would accumulate a representation of concepts that emphasizes the commonalities

across experiences. For instance, memory for the red bike will be worse when they experience the yellow bike, and over time only commonalities (handle bars, wheels, frame) will remain in memory. **b**, Learners with slow forgetting rates would accumulate a representation of concepts that maintain details of unique experiences. For instance, memory for the red bike will be strong when they encounter the yellow bike, and memory for the concept ‘bike’ will retain the idiosyncratic parts of each bike experience (cargo rack, dropped handle bars).

predict that representations might shift again – from being primarily specific in early childhood, to being both specific and broad in older children and adults.

Learning outcomes across development

Although these predictions about possible shifts in representations underlying statistical learning – from infancy through late childhood and adulthood – need to be tested, there is some empirical support for these dynamics. For instance, 8-month-old infants represent their statistical experiences in ‘chunks’ (general groupings), instead of as specific transitional probabilities between items¹³³ as adults have been shown to do¹⁶³. This finding provides behavioural support for the idea that infants have fuzzy representations^{94,135}. Infants can generalize to a new instance of a pattern in statistical rule learning⁵, and after exposure to statistical structure¹⁶⁴. Because fuzzy representations reflect only the most common and general elements of human statistical experiences^{136,137}, infants’ fuzzy representations could support their ability to generalize by allowing them to treat novel information similarly to old, so long as it possesses the same general elements.

In another study, 5–9 year olds and young adults were presented with a visual statistical learning stream made up of repeating shape triplets¹⁶⁵. The main focus of this work was on responses to test items

that matched the statistics of exposure in terms of general groups but mismatched in specific transitions, and vice versa. Children 5–7 years old showed memory for the specific transitions but not the broad groups, whereas 8–9 year olds and adults remembered both broad and specific information¹⁶⁵. This finding in adults replicates previous work¹⁶³ and is supported by intracranial electroencephalography data showing that the human adult brain can represent both specific (transitional probability) and broad (group level) information⁴¹.

Further evidence that young children’s representations might be especially specific comes from work outside statistical learning. Young children (4–5 years old) retain more specific information than adults about exemplars in category learning studies¹⁶⁶, and are less likely than adults to falsely remember a semantically related lure when presented with a list of related words^{167,168}. These findings seem to suggest that young children do not represent the broad semantic gist of word lists, or do so to a lesser extent than older children and adults¹³².

Taken together, these data provide preliminary support for the idea that representations that are formed after statistical learning shift with development: being fuzzy in infancy, highly specific in early childhood, and then increasingly broad in later childhood and adulthood. Although these predictions are consistent with the existing evidence that statistical learning might improve in childhood, they also highlight

the need to consider more nuanced changes in statistical learning than simply performance levels (Table 2).

Summary and future directions

Our review of the development of statistical learning revealed different developmental trajectories in linguistic and non-linguistic domains, with non-linguistic auditory and visual statistical learning improving across childhood whereas linguistic (specifically aurally presented syllables) learning does not change. We also note that these changes were only observed on direct measures of statistical learning, with indirect measures suggesting less change (see Table 1). Critically, understanding the development of statistical learning can be enriched by considering research on memory, learning, and attention. This perspective embraces the consideration of how multiple learning mechanisms might underlie statistical learning^{20,34}. Reviewing related behavioural and neural evidence, we posited that infants and young children might have broader and less-filtered access to the statistical structure present in their learning environments and that what they represent after statistical learning might shift with age.

These ideas could help explain two developmental reversals in learning: why children learn language better^{169,170} and why children are better able to solve problems that require unconventional solutions^{171,18}. Regarding language, we join previous work that suggests that children's prowess in language learning has to do with their ongoing cognitive development^{16,171,172}. Children, especially before the age of 3 years (although later depending on the complexity of the task¹⁷³), tend to generalize across variability in their input^{174,175}. This generalization leads to language errors such as over-regularizing irregular verbs (for example, young children might say 'goed' instead of 'went', adding the standard '-ed' suffix). There are also multiple demonstrations of young children making irregular input – either non-native input from parents¹⁷⁶ or irregular input in an experimental setting¹⁷³ – more regular in their own productions^{177,178}. These observations line up perfectly with the suggestion that very young learners (around the age of 2 years) might be representing commonalities across many events due to forgetting of specific examples.

Many of the other developmental changes outlined here also come to bear on children's unique advantages in learning language. In particular, immediately representing widely sampled specifics could permit a more informed but slower analysis of the data, which could be particularly helpful for learning the grammar of a language, in which there are many exceptions to general rules¹⁷¹. This learning of exceptions would be expectedly slow, in line with the observation that young children surpass their parents when learning in immersive contexts, but only after years (parents do better in the shorter term)¹⁷⁹. Although infants get much credit for their impressive learning feats¹⁸⁰, young children do the heavy lifting when it comes to learning about complex grammatical structure¹⁸¹, in line with a shift from fuzzy to specific representations.

Understanding the input to and output of statistical learning can also aid the understanding of children's unique abilities to solve problems that many adults struggle with. Children are better able to 'think outside the box' to use familiar objects for novel purposes¹⁷ and reason about unconventional causes¹⁸. By sampling their input more broadly, young children are likely better able to see more relationships in their environments than adults can. In addition, representing specifics (and not generalities) might permit children to see things that adults miss, especially if these specifics are less consistent with children's previous experiences and knowledge.

Research is greatly needed to explore how attention shapes statistical learning in complex and naturalistic learning environments for different-aged learners. Both direct (where possible) and indirect data from young children (<5 years old) and infants are greatly needed to understand qualitative changes in statistical learning. Different dependent measures can reveal different aspects of developmental change. For example, eye-tracking measures could be used to understand which parts of their input younger versus older learners are attending to. Indirect measurements might be most sensitive to changes in the input a learner receives over development, whereas direct measures can additionally speak to the representations formed.

Central to this goal is examining what information is prioritized for statistical learning across development and how it relates to the development of attention. Also needed is work considering how breadth of learning might trade off with depth of learning across development. For example, testing whether children are able to learn more statistical information than adults at one time, while also measuring how well each age group learns about the patterns present in different sources of information, could help inform the potential consequences of less focused attention in childhood. Also greatly needed is work examining the representations that are ultimately formed after statistical learning, testing for the presence of fuzzy, specific, and broad memory. In particular, future work should probe the specificity of the information that learners retain after statistical learning and how specificity interacts with forgetting. For example, measuring participants' memory representations for statistical structure (either via a behavioural test or functional neuroimaging) at different points in time after learning could clarify whether statistical memories become fuzzier over time in all age groups. Finally, much work is needed to understand how statistical learning is supported in the developing human brain.

To conclude, it seems overly simplistic to state that statistical learning gets either better or worse with development. Rather, our review of related domains leads us to suggest that, across development, what serves as input to statistical learning likely changes just as much as its output. These insights matter for understanding how

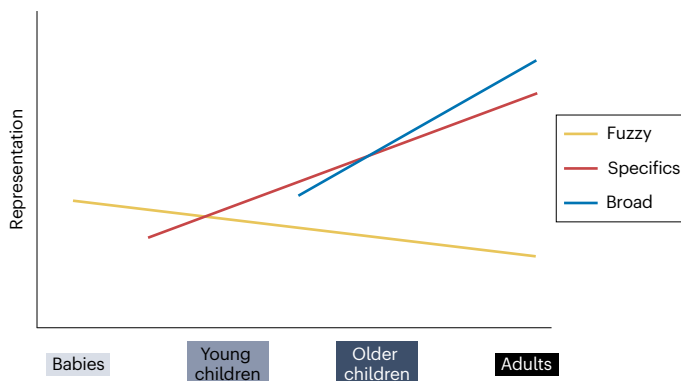


Fig. 6 | How statistical representations can change across development. Representations that might emerge at different points in development. Values on the y-axis indicate the degree to which representations should be fuzzy (yellow), specific (red) and/or broad (blue). Fuzzy memories should emerge already for infants, whereas specific representations should emerge for young children. Older children and adults should be able to represent broad aspects of structure (blue) alongside specifics (red). The exact time in development at which point each representation emerges and the rate at which it matures should be refined in future work.

developmental changes in statistical learning could fundamentally alter the way in which infants and children interact with, learn about, and remember their experiences.

Published online: 2 March 2023

References

- Saffran, J. R., Aslin, R. N. & Newport, E. L. Statistical learning by 8-month-old infants. *Science* **274**, 1926–1928 (1996).
- Maye, J., Werker, J. F. & Gerken, L. A. Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition* **82**, 101–111 (2002).
- Gómez, R. L. Variability and detection of invariant structure. *Psychol. Sci.* **13**, 431–436 (2002).
- Pelucchi, B., Hay, J. F. & Saffran, J. R. Statistical learning in a natural language by 8-month-old infants. *Child Dev.* **80**, 674–685 (2009).
- Marcus, G. F. et al. Rule learning by seven-month-old infants. *Science* **283**, 77–81 (1999).
- Schapiro, A. C., Rogers, T. T., Cordova, N. I., Turk-Browne, N. B. & Botvinick, M. M. Neural representations of events arise from temporal community structure. *Nat. Neurosci.* **16**, 486–492 (2013).
- Jung, Y., Walther, D. B. & Finn, A. S. Children automatically abstract categorical regularities during statistical learning. *Dev. Sci.* **24**, 1–12 (2020).
- Fiser, J. & Aslin, R. N. Unsupervised statistical learning of higher-order spatial structures from visual scenes. *Psychol. Sci.* **12**, 499–504 (2001).
- Fiser, J. & Aslin, R. N. Statistical learning of new visual feature combinations by infants. *Proc. Natl Acad. Sci. USA* **99**, 15822–15826 (2002).
- Turk-Browne, N. B., Jungé, J. & Scholl, B. J. The automaticity of visual statistical learning. *J. Exp. Psychol. Gen.* **134**, 552–564 (2005).
- Turk-Browne, N. B., Isola, P. J., Scholl, B. J. & Treat, T. A. Multidimensional visual statistical learning. *J. Exp. Psychol. Learn. Mem. Cogn.* **34**, 399–407 (2008).
- Campbell, K. L., Zimmerman, S., Healey, M., Lee, M. & Hasher, L. Age differences in visual statistical learning. *Psychol. Aging* **27**, 650–656 (2012).
- Teinonen, T., Fellman, V., Näätänen, R., Alku, P. & Huotilainen, M. Statistical language learning in neonates revealed by event-related brain potentials. *BMC Neurosci.* **10**, 21 (2009).
- Choi, D., Batterink, L. J., Black, A. K., Paller, K. A. & Werker, J. F. Preverbal infants discover statistical word patterns at similar rates as adults: evidence from neural entrainment. *Psychol. Sci.* **31**, 1161–1173 (2020).
- Smith, L. B., Jayaraman, S., Clerkin, E. & Yu, C. The developing infant creates a curriculum for statistical learning. *Trends Cogn. Sci.* <https://doi.org/10.1016/j.tics.2018.02.004> (2018).
- Newport, E. L. Maturation constraints on language learning. *Cogn. Sci.* **14**, 11–28 (1990).
- Adamson, R. E. Functional fixedness as related to problem solving: a repetition of three experiments. *J. Exp. Psychol.* **44**, 288–291 (1952).
- Lucas, C. G., Bridgers, S., Griffiths, T. L. & Gopnik, A. When children are better (or at least more open-minded) learners than adults: developmental differences in learning the forms of causal relationships. *Cognition* **131**, 284–299 (2014).
- Aslin, R. N. Statistical learning: a powerful mechanism that operates by mere exposure. *Wiley Interdiscip. Rev. Cogn. Sci.* **8**, 1373 (2017).
- Conway, C. M. How does the brain learn environmental structure? Ten core principles for understanding the neurocognitive mechanisms of statistical learning. *Neurosci. Biobehav. Rev.* **112**, 279–299 (2020).
- Arciuli, J. The multi-component nature of statistical learning. *Philos. Trans. R. Soc. B Biol. Sci.* **372**, 20160058 (2017).
- Gualtieri, S. & Finn, A. S. The sweet spot: when children's developing abilities, brains, and knowledge make them better learners than adults. *Perspect. Psychol. Sci.* **17**, 1322–1338 (2021).
- Raviv, L. & Arnon, I. The developmental trajectory of children's auditory and visual statistical learning abilities: modality-based differences in the effect of age. *Dev. Sci.* **21**, 12593 (2018).
- Saffran, J. R., Newport, E. L., Aslin, R. N., Tunick, R. A. & Barrueco, S. Incidental language learning: listening (and learning) out of the corner of your ear. *Psychol. Sci.* **8**, 101–105 (1997).
- Krogh, L., Vlach, H. A. & Johnson, S. P. Statistical learning across development: flexible yet constrained. *Front. Psychol.* **3**, 1–11 (2013).
- Saffran, J. R. & Kirkham, N. Z. Infant statistical learning. *Annu. Rev. Psychol.* **69**, 181–203 (2018).
- Saffran, J. R. Statistical language learning in infancy. *Child. Dev. Perspect.* **14**, 49–54 (2020).
- Kirkham, N. Z., Slemmer, J. A. & Johnson, S. P. Visual statistical learning in infancy: evidence for a domain general learning mechanism. *Cognition* **83**, 4–5 (2002).
- Werker, J. F. & Tees, R. C. Cross-language speech perception: evidence for perceptual reorganization during the first year of life. *Infant. Behav. Dev.* **7**, 49–63 (1984).
- Kudo, N., Nonaka, Y., Mizuno, N., Mizuno, K. & Okanoya, K. On-line statistical segmentation of a non-speech auditory stream in neonates as demonstrated by event-related brain potentials. *Dev. Sci.* **14**, 1100–1106 (2011).
- Schlichting, M. L., Guarino, K. F., Schapiro, A. C., Turk-Browne, N. B. & Preston, A. R. Hippocampal structure predicts statistical learning and associative inference abilities during development. *J. Cogn. Neurosci.* **29**, 37–51 (2016).
- Forest, T. A., Lichtenfeld, A., Alvarez, B. & Finn, A. S. Superior learning in synesthetes: consistent grapheme-color associations facilitate statistical learning. *Cognition* **186**, 72–81 (2019).
- Jacoby, L. L. A process dissociation framework: separating automatic from intentional uses of memory. *J. Mem. Lang.* **30**, 513–541 (1991).
- Batterink, L. J., Reber, P. J., Neville, H. J. & Paller, K. A. Implicit and explicit contributions to statistical learning. *J. Mem. Lang.* **83**, 62–78 (2015).
- Siegelman, N., Bogaerts, L., Armstrong, B. C. & Frost, R. What exactly is learned in visual statistical learning? Insights from Bayesian modeling. *Cognition* **192**, 104002 (2019).
- van Wittenloostuijn, M., Lammertink, I., Boersma, P., Wijnen, F. & Rispens, J. Assessing visual statistical learning in early-school-aged children: the usefulness of an online reaction time measure. *Front. Psychol.* **10**, 1–16 (2019).
- Amso, D. & Davidow, J. The development of implicit learning from infancy to adulthood: item frequencies, relations, and cognitive flexibility. *Dev. Psychobiol.* **54**, 664–673 (2012).
- Kabdebon, C., Pena, M., Buiatti, M. & Dehaene-Lambertz, G. Electrophysiological evidence of statistical learning of long-distance dependencies in 8-month-old preterm and full-term infants. *Brain Lang.* **148**, 25–36 (2015).
- Fló, A., Benjamin, L., Palu, M. & Dehaene-Lambertz, G. Sleeping neonates track transitional probabilities in speech but only retain the first syllable of words. *Sci. Rep.* <https://doi.org/10.1101/2021.07.16.452631> (2022).
- Batterink, L. J. & Paller, K. A. Online neural monitoring of statistical learning. *Cortex* **90**, 31–45 (2017).
- Henin, S. et al. Learning hierarchical sequence representations across human cortex and hippocampus. *Sci. Adv.* **7**, 1–13 (2021).
- Fló, A. et al. Newborns are sensitive to multiple cues for word segmentation in continuous speech. *Dev. Sci.* **22**, e12802 (2019).
- Arciuli, J. & Simpson, I. C. Statistical learning in typically developing children: the role of age and speed of stimulus presentation. *Dev. Sci.* **3**, 464–473 (2011).
- Shufaniya, A. & Arnon, I. Statistical learning is not age-invariant during childhood: performance improves with age across modality. *Cogn. Sci.* **42**, 3100–3115 (2018).
- Finn, A. S., Kharitonova, M., Holtby, N. & Sheridan, M. A. Prefrontal and hippocampal structure predict statistical learning ability in early childhood. *J. Cogn. Neurosci.* **31**, 126–137 (2019).
- Ramsar, M. & Gitcho, N. Developmental change and the nature of learning in childhood. *Trends Cogn. Sci.* **11**, 274–279 (2007).
- Turk-Browne, N. B., Scholl, B. J., Chun, M. M. & Johnson, M. K. Neural evidence of statistical learning: efficient detection of visual regularities without awareness. *J. Cogn. Neurosci.* **21**, 1934–1945 (2009).
- Karuza, E. A. et al. The neural correlates of statistical learning in a word segmentation task: an fMRI study. *Brain Lang.* **127**, 46–54 (2013).
- McNealy, K., Mazziotta, J. C. & Dapretto, M. Cracking the language code: neural mechanisms underlying speech parsing. *J. Neurosci.* **26**, 7629–7639 (2006).
- McNealy, K., Mazziotta, J. C. & Dapretto, M. The neural basis of speech parsing in children and adults. *Dev. Sci.* **13**, 385–406 (2010).
- Schapiro, A. C., Turk-Browne, N. B., Norman, K. A. & Botvinick, M. M. Statistical learning of temporal community structure in the hippocampus. *Hippocampus* **26**, 3–8 (2016).
- Ellis, C. T. et al. Evidence of hippocampal learning in human infants. *Curr. Biol.* **31**, 3358–3364.e4 (2021).
- Schlichting, M. L., Mumford, J. A. & Preston, A. R. Learning-related representational changes reveal dissociable integration and separation signatures in the hippocampus and prefrontal cortex. *Nat. Commun.* **6**, 1–10 (2015).
- Mack, M. L., Love, B. C. & Preston, A. R. Dynamic updating of hippocampal object representations reflects new conceptual knowledge. *Proc. Natl Acad. Sci. USA* **113**, 13203–13208 (2016).
- Bowman, C. R. & Zeithamova, D. Abstract memory representations in the ventromedial prefrontal cortex and hippocampus support concept generalization. *J. Neurosci.* **38**, 2605–2614 (2018).
- Gilboa, A. & Marlatte, H. Neurobiology of schemas and schema-mediated memory. *Trends Cogn. Sci.* **21**, 618–631 (2017).
- Preston, A. R. & Eichenbaum, H. Interplay of hippocampus and prefrontal cortex in memory. *Curr. Biol.* **23**, R764–R773 (2013).
- Gabrieli, J. D. E. Cognitive neuroscience of human memory. *Annu. Rev. Psychol.* **49**, 87–115 (1998).
- Keane, M. M., Gabrieli, J. D. E., Mapstone, H. C., Johnson, K. A. & Corkin, S. Double dissociation of memory capacities after bilateral occipital-lobe or medial temporal-lobe lesions. *Brain* **118**, 1129–1148 (1995).
- Fleischman, D. A. et al. Conceptual priming in perceptual identification for patients with Alzheimer's disease and a patient with right occipital lobectomy. *Neuropsychology* **9**, 187–197 (1995).
- Turk-Browne, N. B., Scholl, B. J., Johnson, M. K. & Chun, M. M. Implicit perceptual anticipation triggered by statistical learning. *J. Neurosci.* **30**, 11177–11187 (2010).
- Reber, P. J. The neural basis of implicit learning and memory: a review of neuropsychological and neuroimaging research. *Neuropsychologia* **51**, 2026–2042 (2013).
- Hasson, U., Chen, J. & Honey, C. J. Hierarchical process memory: memory as an integral component of information processing. *Trends Cogn. Sci.* **19**, 304–313 (2015).
- Willingham, D. B., Salidis, J. & Gabrieli, J. D. E. Direct comparison of neural systems mediating conscious and unconscious skill learning. *J. Neurophysiol.* **88**, 1451–1460 (2002).

65. Keele, S. W., Mayr, U., Ivry, R., Hazeltine, E. & Heuer, H. The cognitive and neural architecture of sequence representation. *Psychol. Rev.* **110**, 316–339 (2003).
66. Foerde, K. & Shohamy, D. The role of the basal ganglia in learning and memory: insight from Parkinson's disease. *Neurobiol. Learn. Mem.* **96**, 624–636 (2013).
67. Frank, M. J., Seeberger, L. C. & Reilly, R. C. O. By carrot or by stick: cognitive reinforcement learning in parkinsonism. *Science* **306**, 1940–1944 (2004).
68. Frank, M. J., O'Reilly, R. C. & Curran, T. When memory fails, intuition reigns: midazolam enhances implicit inference in humans. *Psychol. Sci.* **17**, 700–707 (2006).
69. Chatham, C. H., Frank, M. J. & Badre, D. Corticostriatal output gating during selection from working memory. *Neuron* **81**, 930–942 (2014).
70. Awh, E. & Vogel, E. K. The bouncer in the brain. *Nat. Neurosci.* **11**, 5–6 (2008).
71. McNab, F. & Klingberg, T. Prefrontal cortex and basal ganglia control access to working memory. *Nat. Neurosci.* **11**, 103–107 (2008).
72. Darki, F. & Klingberg, T. The role of fronto-parietal and fronto-striatal networks in the development of working memory: a longitudinal study. *Cereb. Cortex* **25**, 1587–1595 (2015).
73. Scoville, W. B. & Milner, B. Loss of recent memory after bilateral hippocampal lesions. *J. Neuropsychiatry Clin. Neurosci.* **20**, 11–21 (1957).
74. Schapiro, A. C., Gregory, E., Landau, B., McCloskey, M. & Turk-Browne, N. B. The necessity of the medial temporal lobe for statistical learning. *J. Cogn. Neurosci.* **26**, 1736–1747 (2014).
75. Schapiro, A. C., Kustner, L. V. & Turk-Browne, N. B. Shaping of object representations in the human medial temporal lobe based on temporal regularities. *Curr. Biol.* **22**, 1622–1627 (2012).
76. Covington, N. V., Brown-Schmidt, S. & Duff, M. C. The necessity of the hippocampus for statistical learning. *J. Cogn. Neurosci.* **30**, 1–19 (2018).
77. Shohamy, D. & Turk-Browne, N. B. Mechanisms for widespread hippocampal involvement in cognition. *J. Exp. Psychol. Gen.* **142**, 1159–1170 (2013).
78. Schapiro, A. C., Turk-Browne, N. B., Botvinick, M. M. & Norman, K. A. Complementary learning systems within the hippocampus: a neural network modelling approach to reconciling episodic memory with statistical learning. *Philos. Trans. R. Soc. B Biol. Sci.* **372**, 20160049 (2017).
79. Eichenbaum, H. & Cohen, N. J. *Conditioning to Conscious Recollection: Memory Systems of the Brain* (Oxford Academic, 2001).
80. Lenroot, R. K. & Giedd, J. N. Brain development in children and adolescents: insights from anatomical magnetic resonance imaging. *Neurosci. Biobehav. Rev.* **30**, 718–729 (2006).
81. Toro, J. M., Sinnett, S. & Soto-Faraco, S. Speech segmentation by statistical learning depends on attention. *Cognition* **97**, 25–34 (2005).
82. Forest, T. A. & Finn, A. S. Attention selectively boosts learning of statistical structure. *Mind Model.* 1674–1679 (2018).
83. Batterink, L. J. & Paller, K. A. Statistical learning of speech regularities can occur outside the focus of attention. *Cortex* **115**, 56–71 (2019).
84. Plude, D. J., Enns, J. T. & Brodeur, D. The development of selective attention: a life-span overview. *Acta Psychol.* **86**, 227–272 (1994).
85. Fortenbaugh, F. C. et al. Sustained attention across the life span in a sample of 10,000: dissociating ability and strategy. *Psychol. Sci.* **26**, 1497–1510 (2015).
86. Hanania, R. & Smith, L. B. Selective attention and attention switching: toward a unified developmental approach. *Dev. Sci.* **13**, 622–635 (2010).
87. Plebanek, D. J. & Sloutsky, V. M. Costs of selective attention: when children notice what adults miss. *Psychol. Sci.* **28**, 723–732 (2017).
88. Deng, W. S. & Sloutsky, V. M. Selective attention, diffused attention, and the development of categorization. *Cogn. Psychol.* **91**, 24–62 (2016).
89. Drummey, A. B. & Newcombe, N. Remembering versus knowing the past: children's explicit and implicit memories for pictures. *J. Exp. Child Psychol.* **59**, 549–565 (1995).
90. Carroll, M., Byrne, B. & Kirsner, K. Autobiographical memory and perceptual learning: a developmental study using picture recognition, naming latency, and perceptual identification. *Mem. Cognit.* **13**, 273–279 (1985).
91. Thomas, K. M. & Nelson, C. A. Serial reaction time learning in preschool- and school-age children. *J. Exp. Child Psychol.* **79**, 364–387 (2001).
92. Thomas, K. M. et al. Evidence of developmental differences in implicit sequence learning: an fMRI study of children and adults. *J. Cogn. Neurosci.* **16**, 1339–1351 (2004).
93. Janacsek, K., Fiser, J. & Nemeth, D. The best time to acquire new skills: age-related differences in implicit sequence learning across the human lifespan. *Dev. Sci.* **15**, 496–505 (2012).
94. Gómez, R. L. & Edgin, J. O. The extended trajectory of hippocampal development: implications for early memory development and disorder. *Dev. Cogn. Neurosci.* **18**, 57–69 (2016).
95. Olson, I. R. & Newcombe, N. S. in *Wiley Handbook on Child Development* (eds Bauer, P. J. & Fivush, R.) 285–308 (Wiley-Blackwell, 2014).
96. Ghetti, S. & Bunge, S. A. Neural changes underlying the development of episodic memory during middle childhood. *Dev. Cogn. Neurosci.* **4**, 1–29 (2012).
97. Ghetti, S. & Fandakova, Y. Neural development of memory and metamemory in childhood and adolescence: toward an integrative model of the development of episodic recollection. *Annu. Rev. Dev. Psychol.* **2**, 365–388 (2020).
98. Smith, M. A., Ghazizadeh, A. & Shadmehr, R. Interacting adaptive processes with different timescales underlie short-term motor learning. *PLoS Biol.* **4**, e179 (2006).
99. McDougle, S. D., Bond, K. M. & Taylor, J. A. Explicit and implicit processes constitute the fast and slow processes of sensorimotor learning. *J. Neurosci.* **35**, 9568 (2015).
100. Goshen-Gottstein, Y. & Kempinsky, H. Probing memory with conceptual cues at multiple retention intervals: a comparison of forgetting rates on implicit and explicit tests. *Psychon. Bull. Rev.* **8**, 139–146 (2001).
101. Graf, P., Squire, L. R. & Mandler, G. The information that amnesic patients do not forget. *J. Exp. Psychol. Learn. Mem. Cogn.* **10**, 164–178 (1984).
102. Rappold, V. A. & Hashtroudi, S. Does organization improve priming? *J. Exp. Psychol. Learn. Mem. Cogn.* **17**, 103–114 (1991).
103. Clerkin, E. M., Hart, E., Reh, J. M., Yu, C. & Smith, L. B. Real-world visual statistics and infants' first-learned object names. *Philos. Trans. R. Soc. B Biol. Sci.* **372**, 20160055 (2017).
104. Smith, L. B., Yu, C., Yoshida, H. & Fausey, C. M. Contributions of head-mounted cameras to studying the visual environments of infants and young children. *J. Cogn. Dev.* **16**, 407 (2015).
105. Zhao, J. & Luo, Y. Statistical regularities guide the spatial scale of attention. *Atten. Percept. Psychophys.* **79**, 24–30 (2016).
106. Wang, B., Samara, I. & Theeuwes, J. Statistical regularities bias overt attention. *Atten. Percept. Psychophys.* **81**, 1812–1821 (2019).
107. Finn, A. S. & Hudson Kam, C. L. The curse of knowledge: first language knowledge impairs adult learners' use of novel statistics for word segmentation. *Cognition* **108**, 477–499 (2008).
108. Toro, J. M., Pons, F., Bion, R. A. H. & Sebastián-Gallés, N. The contribution of language-specific knowledge in the selection of statistically-coherent word candidates. *J. Mem. Lang.* **64**, 171–180 (2011).
109. Itti, L. & Baldi, P. Bayesian surprise attracts human attention. *Vis. Res.* **49**, 1295–1306 (2009).
110. Ranganath, C. & Rainer, G. Neural mechanisms for detecting and remembering novel events. *Nat. Rev. Neurosci.* **4**, 193–202 (2003).
111. Forest, T. A., Siegelman, N. & Finn, A. S. Attention shifts to more complex locations with experience. *Psychol. Sci.* **33**, 2059–2072 (2022).
112. Van Kesteren, M. T. R., Ruiter, D. J., Fernández, G. & Henson, R. N. How schema and novelty augment memory formation. *Trends Neurosci.* **35**, 211–219 (2012).
113. Zettersten, M. & Saffran, J. R. Sampling to learn words: adults and children sample words that reduce referential ambiguity. *Dev. Sci.* <https://doi.org/10.1111/desc.13064> (2020).
114. Wade, S. & Kidd, C. The role of prior knowledge and curiosity in learning. *Psychon. Bull. Rev.* **26**, 1377–1387 (2019).
115. Loewenstein, G. The psychology of curiosity: a review and reinterpretation. *Psychol. Bull.* **116**, 75–94 (1994).
116. Gebhart, A. L., Aslin, R. N. & Newport, E. L. Changing structures in midstream: learning along the statistical garden path. *Cogn. Sci.* **33**, 1087–1116 (2009).
117. Bulgarelli, F. & Weiss, D. J. Anchors aweigh: the impact of overlearning on entrenchment effects in statistical learning. *J. Exp. Psychol. Learn. Mem. Cogn.* **42**, 1621–1631 (2016).
118. Kóbor, A., Horváth, K., Kardos, Z., Nemeth, D. & Janacsek, K. Perceiving structure in unstructured stimuli: implicitly acquired prior knowledge impacts the processing of unpredictable transitional probabilities. *Cognition* **205**, 104413 (2020).
119. Kidd, C., Piantadosi, S. T. & Aslin, R. N. The Goldilocks effect: human infants allocate attention to visual sequences that are neither too simple nor too complex. *PLoS ONE* **7**, e36399 (2012).
120. Kidd, C., Piantadosi, S. T. & Aslin, R. N. The Goldilocks effect in infant auditory attention. *Child. Dev.* **85**, 1795–1804 (2014).
121. Cubit, L. S., Canale, R., Handsman, R., Kidd, C. & Bennetto, L. Visual attention preference for intermediate predictability in young children. *Child. Dev.* **92**, 691–703 (2021).
122. Poli, F., Serino, G., Mars, R. B. & Hunnius, S. Infants tailor their attention to maximize learning. *Sci. Adv.* **6**, 1–9 (2020).
123. Ngo, C. T., Benear, S. L., Popal, H., Olson, I. & Newcombe, N. S. Contingency of semantic generalization on episodic specificity: variations across development. *Curr. Biol.* **31**, 1–8 (2021).
124. Drummey, A. B. & Newcombe, N. S. Developmental changes in source memory. *Dev. Sci.* **5**, 502–513 (2002).
125. Hudson, J. A., Fivush, R. & Kuebli, J. Scripts and episodes: the development of event memory. *Appl. Cogn. Psychol.* **6**, 483–505 (1992).
126. Saragosa-Harris, N. M. et al. Associative memory persistence in 3- to 5-year-olds. *Dev. Sci.* <https://doi.org/10.1111/desc.13105> (2021).
127. Lloyd, M. E., Doydum, A. O. & Newcombe, N. S. Memory binding in early childhood: evidence for a retrieval deficit. *Child. Dev.* **80**, 1321–1328 (2009).
128. Raj, V. & Bell, M. A. Cognitive processes supporting episodic memory formation in childhood: the role of source memory, binding, and executive functioning. *Dev. Rev.* **30**, 384–402 (2010).
129. Sluzenski, J., Newcombe, N. S. & Kovacs, S. L. Binding, relational memory, and recall of naturalistic events: a developmental perspective. *J. Exp. Psychol. Learn. Mem. Cogn.* **32**, 89–100 (2006).
130. Brainerd, C. J. & Reyna, V. F. Fuzzy-trace theory and memory development. *Dev. Rev.* **24**, 396–439 (2004).
131. Barr, R. & Brito, N. From specificity to flexibility: early developmental changes in memory generalization. *Wiley Handb. Dev. Child. Mem.* <https://doi.org/10.1002/9781118597705.CH20> (2014).
132. Schlichting, M. L., Guarino, K. F., Roome, H. E. & Preston, A. R. Developmental differences in memory reactivation relate to encoding and inference in the human brain. *Nat. Hum. Behav.* <https://doi.org/10.1038/s41562-021-01206-5> (2021).

133. Slone, L. K. & Johnson, S. P. When learning goes beyond statistics: infants represent visual sequences in terms of chunks. *Cognition* **178**, 92–102 (2018).
134. Reyna, V. F. & Brainerd, C. J. Fuzzy-trace theory: an interim synthesis. *Learn. Individ. Differ.* **7**, 1–75 (1995).
135. Gómez, R. L. Do infants retain the statistics of a statistical learning experience? Insights from a developmental cognitive neuroscience perspective. *Philos. Trans. R. Soc. B Biol. Sci.* **372**, 20160054 (2017).
136. Eglington, L. G. & Kang, S. H. K. Interleaved presentation benefits science category learning. *J. Appl. Res. Mem. Cogn.* **6**, 475–485 (2017).
137. Vlach, H. A. The spacing effect in children's generalization of knowledge: allowing children time to forget promotes their ability to learn. *Child. Dev. Perspect.* **8**, 163–168 (2014).
138. Howe, M. L. & Courage, M. L. Independent paths in the development of infant learning and forgetting. *J. Exp. Child Psychol.* **67**, 131–163 (1997).
139. Bhatt, R. S. & Rovee-collier, C. Infants' forgetting of correlated attributes and object recognition. *Child. Dev.* **67**, 172–187 (1996).
140. Bauer, P. J., Wenner, J. A., Dropik, P. L., Wewerka, S. S. & Howe, M. L. Parameters of remembering and forgetting in the transition from infancy to early childhood. *Monogr. Soc. Res. Child Dev.* **65**, 1–204 (2000).
141. McClelland, J. L., McNaughton, B. L. & O'Reilly, R. C. Why there are complementary learning systems in the hippocampus and neocortex: insights from the success and failures of connectionist models of learning and memory. *Psychol. Rev.* **102**, 419–457 (1995).
142. Duncan, K. D. & Schlichting, M. L. Neurobiology of learning and memory hippocampal representations as a function of time, subregion, and brain state. *Neurobiol. Learn. Mem.* **153**, 40–56 (2018).
143. Lavenex, P. & Banta Lavenex, P. Building hippocampal circuits to learn and remember: insights into the development of human memory. *Behav. Brain Res.* **254**, 8–21 (2013).
144. Jabès, A. & Nelson, C. A. 20 years after 'the ontogeny of human memory: a cognitive neuroscience perspective', where are we? *Int. J. Behav. Dev.* **39**, 293–303 (2015).
145. Hayne, H., MacDonald, S. & Barr, R. Developmental changes in the specificity of memory over the second year of life. *Infant. Behav. Dev.* **20**, 233–245 (1997).
146. Ribordy, F., Jabès, A., Banta Lavenex, P. & Lavenex, P. Development of allocentric spatial memory abilities in children from 18 months to 5 years of age. *Cogn. Psychol.* **66**, 1–29 (2013).
147. Pathman, T., Doydum, A. & Bauer, P. J. Bringing order to life events: memory for the temporal order of autobiographical events over an extended period in school-aged children and adults. *J. Exp. Child Psychol.* **115**, 309–325 (2013).
148. Keresztes, A., Ngo, C. T., Lindenberger, U., Werkle-Bergner, M. & Newcombe, N. S. Hippocampal maturation drives memory from generalization to specificity. *Trends Cogn. Sci.* **22**, 676–686 (2018).
149. Callaghan, B. et al. Age-related increases in posterior hippocampal granularity are associated with remote detailed episodic memory in development. *J. Neurosci.* **41**, 1738–1754 (2021).
150. Keresztes, A. et al. Hippocampal maturity promotes memory distinctiveness in childhood and adolescence. *Proc. Natl Acad. Sci. USA* **114**, 9212–9217 (2017).
151. Shing, Y. L., Werkle-Bergner, M., Li, S. C. & Lindenberger, U. Associative and strategic components of episodic memory: a life-span dissociation. *J. Exp. Psychol. Gen.* **137**, 495–513 (2008).
152. Riggins, T., Blankenship, S. L., Mulligan, E., Rice, K. & Redcay, E. Developmental differences in relations between episodic memory and hippocampal subregion volume during early childhood. *Child. Dev.* **86**, 1710–1718 (2015).
153. Lee, J. K. et al. Changes in anterior and posterior hippocampus differentially predict item-space, item-time, and item-item memory improvement. *Dev. Cogn. Neurosci.* **41**, 100741 (2020).
154. Gogtay, N. et al. Dynamic mapping of normal human hippocampal development. *Hippocampus* **16**, 664–672 (2006).
155. Langnes, E. et al. Anterior and posterior hippocampus macro- and microstructure across the lifespan in relation to memory—a longitudinal study. *Hippocampus* **30**, 678–692 (2020).
156. Strange, B. A., Witter, M. P., Lein, E. S. & Moser, E. I. Functional organization of the hippocampal longitudinal axis. *Nat. Rev. Neurosci.* **15**, 655–669 (2014).
157. Schlichting, M. L. & Preston, A. R. Hippocampal-medial prefrontal circuit supports memory updating during learning and post-encoding rest. *Neurobiol. Learn. Mem.* **134**, 37–51 (2016).
158. Calabro, F. J., Murty, V. P., Jalbrzikowski, M., Tervo-Clemmens, B. & Luna, B. Development of hippocampal-prefrontal cortex interactions through adolescence. *Cereb. Cortex* **30**, 1548–1558 (2020).
159. Barbas, H. & Blatt, G. J. Topographically specific hippocampal projections target functionally distinct prefrontal areas in the rhesus monkey. *Hippocampus* **5**, 511–533 (1995).
160. DeMaster, D., Coughlin, C. & Ghetti, S. Retrieval flexibility and reinstatement in the developing hippocampus. *Hippocampus* **26**, 492–501 (2016).
161. Brunec, I. K. et al. Multiple scales of representation along the hippocampal anteroposterior axis in humans. *Curr. Biol.* **28**, 2129–2135.e6 (2018).
162. Zeithamova, D. & Preston, A. R. Flexible memories: differential roles for medial temporal lobe and prefrontal cortex in cross-episode binding. *J. Neurosci.* **30**, 14676–14684 (2010).
163. Forest, T. A., Finn, A. S. & Schlichting, M. L. General precedes specific in memory representations for structured experience. *J. Exp. Psychol. Gen.* <https://doi.org/10.1037/xge0001104> (2021).
164. Gómez, R. L., Bootzin, R. R. & Nadel, L. Naps promote abstraction in language-learning infants. *Psychol. Sci.* **17**, 670–674 (2006).
165. Forest, T. A., Abolghasem, Z., Finn, A. S. & Schlichting, M. L. Memories of structured input become increasingly distorted across development. *Child Dev.* (in the press).
166. Sloutsky, V. M. & Fisher, A. V. When development and learning decrease memory: evidence against category-based induction in children. *Psychol. Sci.* **15**, 553–558 (2004).
167. Brainerd, C. J., Reyna, V. F. & Ceci, S. J. Developmental reversals in false memory: a review of data and theory. *Psychol. Bull.* **134**, 343–382 (2008).
168. Brainerd, C. J., Reyna, V. F. & Forrest, T. J. Are young children susceptible to the false-memory illusion? *Child. Dev.* **73**, 1363–1377 (2002).
169. Johnson, J. S. & Newport, E. Critical period effects in second language learning: the influence of maturational state on the acquisition of English as a second language. *Cogn. Psychol.* **21**, 60–99 (1989).
170. Birdsong, D. (ed.) *Second Language Acquisition and the Critical Period Hypothesis* (Routledge, 1999).
171. Arnon, I. & Ramscar, M. Granularity and the acquisition of grammatical gender: how order-of-acquisition affects what gets learned. *Cognition* **122**, 292–305 (2012).
172. Thompson-Schill, S., Ramscar, M. & Chrysikou, E. G. Cognition without control. *Psychol. Sci.* **18**, 259–263 (2009).
173. Hudson, C. L. & Newport, E. L. Getting it right by getting it wrong: when learners change languages. *Cogn. Psychol.* **59**, 30–66 (2009).
174. Bever, T. G. in *Regression in Mental Development: Basic Properties and Mechanisms* (ed. Bever, T. G.) 153–188 (Lawrence Erlbaum, 1982).
175. Kessen, W. & Kessen, M. L. Behavior of young children in a two-choice guessing problem. *Child Dev.* **32**, 779–788 (1961).
176. Singleton, J. L. & Newport, E. L. When learners surpass their models: the acquisition of American Sign Language from inconsistent input. *Cogn. Psychol.* **49**, 370–407 (2004).
177. Wonnacott, E. Balancing generalization and lexical conservatism: an artificial language study with child learners. *J. Mem. Lang.* **65**, 1–14 (2011).
178. Perfors, A. When do memory limitations lead to regularization? An experimental and computational investigation. *J. Mem. Lang.* **67**, 486–506 (2012).
179. Snow, C. E. & Hoefnagel-Höhle, M. The critical period for language acquisition: evidence from second language learning. *Child. Dev.* **49**, 1114–1128 (1978).
180. Aslin, R. N., Saffran, J. R. & Newport, E. L. Computation of conditional probability statistics by 8-month-old infants. *Psychol. Sci.* **9**, 321–324 (1998).
181. Brown, R. *A First Language* (Harvard Univ. Press, 1973).
182. Fiser, J. & Aslin, R. N. Statistical learning of higher-order temporal structure from visual shape sequences. *J. Exp. Psychol. Learn. Mem. Cogn.* **28**, 458–467 (2002).
183. Chomsky, N. *Rules and Representations* (Cambridge Univ. Press, 1980).
184. Nissen, M. J. & Bullemer, P. Attentional requirements of learning: evidence from performance measures. *Cogn. Psychol.* **19**, 1–32 (1987).
185. Perruchet, P. & Pacton, S. Implicit learning and statistical learning: one phenomenon, two approaches. *Trends Cogn. Sci.* **10**, 233–238 (2006).
186. Henke, K. A model for memory systems based on processing modes rather than consciousness. *Nat. Rev. Neurosci.* **11**, 523–532 (2010).
187. Bays, B. C., Turk-Browne, N. B. & Seitz, A. R. Dissociable behavioural outcomes of visual statistical learning. *Vis. Cogn.* **23**, 1072–1097 (2016).
188. Poldrack, R. A. et al. Interactive memory systems in the human brain. *Nature* **414**, 546–550 (2001).
189. Alamia, A. & Zénon, A. Statistical regularities attract attention when task-relevant. *Front. Hum. Neurosci.* **10**, 1–10 (2016).
190. Zhao, J., Al-Aidroos, N. & Turk-Browne, N. B. Attention is spontaneously biased toward regularities. *Psychol. Sci.* **24**, 667–677 (2013).
191. Squire, L. R. & Zola-Morgan, S. Memory: brain systems and behavior. *Trends Neurosci.* **11**, 170–175 (1988).
192. Bayley, P. J., Frascino, J. C. & Squire, L. R. Robust habit learning in the absence of awareness and independent of the medial temporal lobe. *Nature* **436**, 550–553 (2005).
193. Knowlton, B. J., Mangels, J. A. & Squire, L. R. A neostriatal habit learning system in humans. *Science* **273**, 1399–1402 (1996).
194. Schacter, D. L. in *Memory Systems of the Brain: Animal and Human Cognitive Processes* (eds Weinberger, N. M., McGaugh, J. L. & Lynch, G.) 351–379 (Guildford Publications, 1985).
195. McDonald, R. J. & White, N. M. A triple dissociation of memory systems: hippocampus, amygdala, and dorsal striatum. *Behav. Neurosci.* **107**, 3–22 (1993).
196. McDonald, R. J. & White, N. M. Hippocampal and nonhippocampal contributions to place learning in rats. *Behav. Neurosci.* **109**, 579–593 (1995).
197. Packard, M. G., Hirsh, R. & White, N. M. Differential effects of fornix and caudate nucleus lesions on two radial maze tasks: evidence for multiple memory systems. *J. Neurosci.* **9**, 1465–1472 (1989).
198. Ferbinteanu, J. & McDonald, R. J. Dorsal/ventral hippocampus, fornix, and conditioned place preference. *Hippocampus* **11**, 187–200 (2001).

199. Chai, S.-C. & White, N. M. Effects of fimbria-fornix, hippocampus, and amygdala lesions on discrimination between proximal locations. *Behav. Neurosci.* **118**, 770–784 (2004).
200. Poldrack, R. A. & Packard, M. G. Competition among multiple memory systems: converging evidence from animal and human brain studies. *Neuropsychologia* **41**, 245–251 (2003).
201. Wimmer, G. E., Braun, E. K., Daw, N. D. & Shohamy, D. Episodic memory encoding interferes with reward learning and decreases striatal prediction errors. *J. Neurosci.* **34**, 14901–14912 (2014).
202. Foerde, K., Knowlton, B. J. & Poldrack, R. A. Modulation of competing memory systems by distraction. *Proc. Natl Acad. Sci. USA* **103**, 11778–11783 (2006).
203. Foerde, K. & Shohamy, D. Feedback timing modulates brain systems for learning in humans. *J. Neurosci.* **31**, 13157–13167 (2011).
204. Dickerson, K. C. & Delgado, M. R. Contributions of the hippocampus to feedback learning. *Cogn. Affect. Behav. Neurosci.* **15**, 861–877 (2015).
205. Ballard, I. C., Wagner, A. D. & McClure, S. M. Hippocampal pattern separation supports reinforcement learning. *Nat. Commun.* **10**, 1073 (2019).
206. Duncan, K., Doll, B. B., Daw, N. D. & Shohamy, D. More than the sum of its parts: a role for the hippocampus in configural reinforcement learning. *Neuron* **98**, 645–657 (2018).
207. Wimmer, G. E. & Shohamy, D. Preference by association: how memory mechanisms in the hippocampus bias decisions. *Science* **338**, 270–273 (2012).
208. Shohamy, D. & Wagner, A. D. Integrating memories in the human brain: hippocampal–midbrain encoding of overlapping events. *Neuron* **60**, 378–389 (2008).
209. Scimeca, J. M. & Badre, D. Striatal contributions to declarative memory retrieval. *Neuron* **75**, 380–392 (2012).

Acknowledgements

The authors thank all members of the Learning and Neural Development, Budding Minds, and Duncan laboratories at the University of Toronto for continued helpful discussion and support. They also thank C. Hudson Kam for feedback on an earlier version of the manuscript.

Author contributions

A.S.F. and T.A.F. researched data for the article. A.S.F. and T.A.F. contributed substantially to discussion of the content. All authors wrote the article. All authors reviewed and/or edited the manuscript before submission.

Competing interests

The authors declare no competing interests.

Additional information

Peer review information *Nature Reviews Psychology* thanks Laura Batterink, Karolina Janacek and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

© Springer Nature America, Inc. 2023