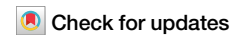




Addressing the theory crisis in statistical learning research



Christopher M. Conway¹ ✉, Holly E. Jenkins², Alice E. Milne^{3,4}, Sonia Singh⁵ & Benjamin Wilson^{6,7}

Research into statistical learning, the ability to learn structured patterns in the environment, faces a theory crisis. Specifically, three challenges must be addressed: a lack of robust phenomena to constrain theories, issues with construct validity, and challenges with establishing causality. Here, we describe and discuss each issue in relation to several prominent statistical learning phenomena. We then offer recommendations to help address the theory crisis and move the field forward.

The capacity to learn from experience is critical for survival and therefore ubiquitous across most organisms. An important type of learning, called statistical learning, involves sensitivity to structured patterns in the environment^{1–4}. Patterns come in many varieties, such as how certain objects typically appear (e.g., animals, trees, and human-made objects each have a particular form and shape) to the regularities found in temporal input streams such as music and language (e.g., hearing random notes on a piano would not be considered music, just as a random concatenation of speech syllables would sound meaningless).

Decades of research on statistical learning, and its related cousin, implicit learning^{5,6}, has demonstrated that humans (and other organisms) are sensitive to spatial and temporal patterns across many perceptual, motor, and cognitive domains^{7,8}. This type of learning appears to occur relatively automatically, without intention to learn, and results in behavioral, cognitive, and/or perceptual facilitation that is often (though not always) accompanied by a lack of conscious awareness of the learned patterns. Because of the apparent reach of statistical learning to explain functioning across a myriad of domains and situations, it is perhaps not surprising that this area of research has been exploding in recent years^{9,10}.

However, recent critiques have highlighted challenges facing statistical learning research. Namely, the construct of statistical learning itself is underspecified in terms of its underlying neurocognitive mechanisms, how these mechanisms relate to various laboratory tasks commonly used to measure statistical learning, and what is the relationship between statistical learning and other forms of learning, memory, and cognition^{9,11–14}. While some theoretical frameworks do exist, including the extraction-integration framework^{15–18}, chunking approaches^{9,19,20}, frameworks favoring the role of brain plasticity and modality constraints^{2,7,21}, and other multi-component models^{22–25}, there is no universally-agreed upon theory that can adequately address these issues.

We contend that these challenges are not unique to statistical learning research but in fact are difficulties confronting the psychological sciences

more broadly. It has been argued that the field of psychology faces a “theory crisis”; that is, psychological theories are often poor at explaining the relevant phenomena of interest and no amount of improved statistical techniques or replication studies will improve the situation^{26–28}. Whereas efforts have been made to encourage psychologists to create more formal or precise theories²⁹, Eronen and Bringmann³⁰ have recently argued that “the core of the problem is that developing good psychological theories is extremely difficult” (p. 780). Therefore, understanding the obstacles facing theory building in psychology is the first necessary step for making progress in addressing the theory crisis.

In this paper, we use Eronen and Bringmann’s³⁰ and others’ insights about the theory crisis as a foundation for moving the field of statistical learning forward, one that has seen an explosion of growth in recent years⁹. Statistical learning, and even implicit learning, have historically been challenging to define, as different tasks may elicit different cognitive processes and engage distinct neuroanatomical systems (for a review, see ref. 14). Therefore, for the purposes of this paper, we define statistical learning prototypically, as encompassing situations that tend to have a specific set of characteristics, in particular, those that involve the learning of patterns in the environment, through multiple exposures, and under incidental conditions⁷. Note that this definition can be thought of as a higher-level description about what problem or goal the system has evolved to address—i.e., to learn environmental patterns over multiple exposures under incidental conditions—rather than a definition predicated on identifying the learning processes or neural mechanisms involved. This higher-level description is similar to Marr’s³¹ top level of analysis (what he calls the “computational” level), whereas descriptions of the cognitive and neural mechanisms fall under Marr’s 2nd (“algorithmic”) and 3rd (“implementational”) levels, respectively (see Conway et al.³² for elaboration of Marr’s framework in relation to statistical learning). The advantage of starting with such a higher-level definition is that it is based on the characteristics of the *learning situation*, rather than on the characteristics of the

¹Department of Psychology, Grinnell College, Grinnell, IA, USA. ²Department of Education, University of Oxford, Oxford, UK. ³Department of Psychology, Lancaster University, Lancaster, UK. ⁴Ear Institute, University College London, London, UK. ⁵Department of Language and Cognition, University College London, London, UK. ⁶Department of Psychology, Emory University, Atlanta, GA, USA. ⁷Emory National Primate Research Center, Atlanta, GA, USA.

✉ e-mail: conwaych@grinnell.edu

learner, which are “unseen” cognitive and neural processes that can be difficult to identify¹⁴. Thus, this definition provides (what we believe is) a relatively uncontroversial description of the types of situations that statistical learning typically encompasses and that clearly specifies to what extent any given task or situation is relevant to this area of research.

With this definition in place, the ultimate aim should be to derive a theoretical framework that provides meaningful and valid conceptualizations of the various phenomena that statistical learning touches upon. To work towards this goal, in the following sections we first describe three key issues as laid out by Eronen and Bringmann: lack of robust phenomena, issues with construct validity, and issues with establishing causality. Next, we discuss each issue in terms of how it relates to statistical learning by focusing on several research phenomena that have been examined in some detail. In the final section of this paper, we offer suggestions for ways to address the theory crisis in statistical learning research to move the field forward in fruitful directions.

Challenges facing SL research: lack of robust phenomena

Eronen and Bringmann³⁰ contended that there is a lack of robust phenomena that can be used to constrain and formulate psychological theories. To understand this problem fully, it is necessary to describe the difference between data, phenomena, and theory^{29,30} (see Fig. 1). Data are the raw measurements collected in psychology experiments (i.e., measures of behavior or neurophysiology). Phenomena are a description of the effects compiled across the observed data; an example in psychology might be the Stroop effect^{33,34} or implicit bias/stereotype effect³⁵, which are descriptions of consistently occurring patterns within data. Finally, theories are created to explain the phenomena; and yet, phenomena also constrain theories³⁰. It is this last point that Eronen and Bringmann³⁰ claimed has received little attention in psychology. Without the existence of robust—i.e., consistent and stable—phenomena, then theories are on shaky ground.

Is there a lack of robust phenomena in statistical learning research? What are the relevant phenomena in the first place, and what is the metric

The Theory of Evolution as a model for good theory building

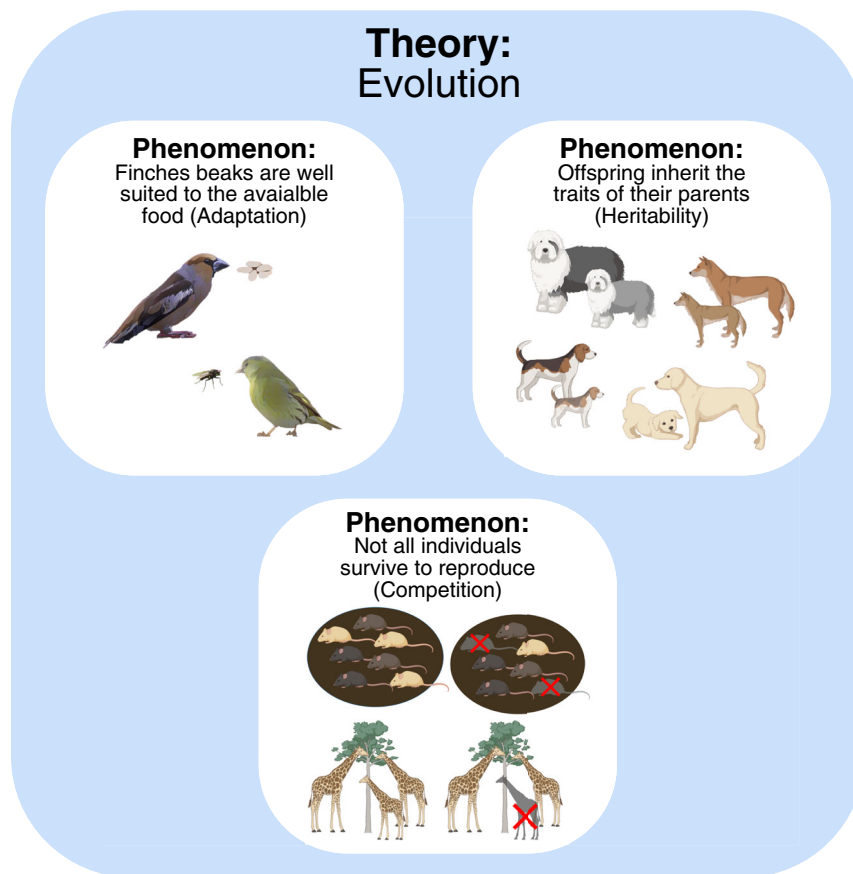


Fig. 1 | The theory of evolution as a model for good theory building. Eronen and Bringmann³⁰ emphasize the distinction between *data*, *phenomena*, and *theories*. Given the lack of universally accepted theories within Psychology, we turn instead to the theory of evolution to provide concrete examples of the interplay between data, phenomena, and theory. Charles Darwin did not set out with the goal of developing the theory of evolution. Rather, he initially made a wide range of different, at the time disparate, observations, including, for example, that the large ground finch primarily ate seeds (which happened to have a large, powerful beak), while the green warbler finch (with a narrower, pointed beak) fed upon insects. These observations represent *data* (individual images in the figure). Based on many pieces of data, more general *phenomena* emerged (e.g., that animals are well adapted to their environments; that offspring inherit the traits of their parents; and that differential survival rates leads to changes in how a given phenotype is represented in a population). Finally, Darwin advanced the *Theory* of evolution by natural selection¹⁵¹, which

explained and unified these phenomena. Critically, a good theory, like evolution, must be able to explain all the relevant phenomena, and remains open to falsification if counter examples are found. In this way, not only do theories explain phenomena, but phenomena tightly constrain the space of possible theories. Note, the theory of evolution was selected as an example as it is extremely well developed and supported by data. The development of any new theory, including a theory of statistical learning, is likely to be less clear and much noisier. For example, it is yet to be determined whether all of the phenomena outlined in this article (or other phenomena, not discussed here), are sufficiently robust to be included in an eventual theory. Moreover, it is possible that a theory may arise that explains some, but not all of these phenomena. These difficulties represent some of the key challenges of theory building, but we hope that the proposals made in this article will help advance such theories in the future. The images of animals within the figure were created using BioRender.com. Wilson, B. (2025) <https://BioRender.com/5iaoy0f>.

for evaluating whether they are robust? Robustness of a phenomenon is at least partly related to the notion of replicability; if a particular pattern of data is not observed consistently across studies, then it is not a robust finding. However, as pointed out by Eronen and Bringmann, it is just as important that for a phenomenon to be robust, it should be “verifiable and detectable in several independent ways and not dependent on a specific theoretical framework or observation method”^{30,36,37} (p. 780).

Based on this conceptualization of robustness, we suggest that the overall phenomenon of statistical learning itself, that is, the general finding that people are sensitive to patterns in the environment under incidental conditions, is robust, demonstrated across many studies in a multitude of situations using various laboratory tasks and stimuli³⁸. While this general phenomenon appears robust, it is also important to delineate a number of additional sub-phenomena that have been examined in the statistical learning literature, including but not limited to: modality effects³⁹, effects of input complexity⁴⁰, age effects⁴¹, species differences⁴², role of attention in learning⁴³, competition between executive functions and statistical learning⁴⁴, the contribution of statistical learning to language acquisition¹³, and atypical learning in developmental disorders⁴⁵ (see also⁷, for a review of some of these phenomena). In subsequent sections we provide an evaluation of the robustness of several of these key phenomena.

The point is that a good theoretical framework of statistical learning should not only encompass the general phenomenon of statistical learning itself, but also these additional sub-phenomena (e.g., the presence of modality effects)—but only those that are considered robust. If a phenomenon is not a robust one, then it may be counter-productive to devise theoretical frameworks to explain it. Attempting to devise theories to explain phenomena that are not robust could lead to the unfortunate situation where “correct” theories are rejected because they fail to account for these (unrobust) phenomena. We return to this point at the end of the paper when discussing recommendations. It is therefore imperative to identify which phenomena appear to be robust and which ones need additional empirical support or testing. We discuss the robustness of several such phenomena, in Section “Evaluating Statistical Learning Phenomena”, below.

Challenges facing SL research: issues with construct validity

The second challenge facing the psychological sciences, highlighted by Eronen and Bringmann³⁰ concerns issues with construct validity. A simple definition of construct validity is that a measurement measures what it is intended to measure⁴⁶. Whereas reliability can be easily quantified and tested, construct validity is more nebulous. As such, psychological scientists typically report little evidence for validity, instead focusing on more concrete psychometric properties such as reliability^{47,48}. For instance, statistical learning researchers have made strides to improve the reliability of the measurements used^{49–52}. This is an important and necessary step. However, while issues of validity have been mentioned in the statistical learning literature, including the use of vague or circular definitions for statistical learning and/or the use of different terms and a multitude of ways to measure it^{9,11,14,53}, no formal attempts at establishing construct validity have been made.

A further challenge to assessing the validity of measures in statistical learning research is that different types of methods have been used to assess statistical learning (see refs. 11,53), including the following tasks: artificial grammar learning (AGL)⁵, serial reaction time (SRT)⁵⁴, Hebb repetition⁵⁵, triplet segmentation or embedded pattern tasks⁴, cross-situational learning⁵⁶, and contextual cueing⁵⁷. Not only this, but even the same paradigm (e.g., AGL) can vary in terms of the type of input used or how learning is measured. The challenge with using such an array of task methodologies is that, while all of them on the face of it appear to measure statistical learning, each one may emphasize a different type or aspect of such learning, and potentially be supported by different underlying neurocognitive mechanisms.

As pointed out by Eronen and Bringmann³⁰, weak construct validity impacts robustness of phenomena. Because phenomena depend upon inferences made over data, if those data are not well-understood or well-validated, then the phenomena themselves are unlikely to be robust.

Challenges facing SL research: issues with establishing causality

The final challenge raised by Eronen and Bringmann³⁰ is that determining causality among psychological variables is extremely challenging. As they note, good theories should “capture causal relationships between psychological variables” (p.783). However, doing so is extremely difficult⁵⁸. While it is well-understood how to establish causality by manipulating external, non-psychological variables using sound experimental design (e.g.,⁵⁹ psychological theories need to understand the causal relationships between *psychological* variables, which are much more difficult to directly observe and manipulate. Eronen and Bringmann³⁰ refer to this as the “fat-handed” problem: to manipulate a psychological variable requires manipulating an external variable, but the manipulated external variable might be affecting other psychological processes/variables instead of or in addition to the specific psychological variable of interest.

To take an example from statistical learning research, one ongoing question is to what extent statistical learning depends upon attention (e.g.,^{7,60}). One common way to manipulate participants’ attention is to use task instructions to direct selective attention to some stimuli but not others and to test if statistical learning still occurs for the unattended stimuli (e.g.,⁴³). While such an approach is commendable and logical, the question is whether other psychological variables apart from attention may also be affected. For instance, it is possible that variables such as motivation, confidence, and/or effort may also be affected by the task instructions, and if so, it is difficult to know to what extent attention itself (or lack thereof) causally affected learning. Note that the problem of causality is exacerbated by the fact that different statistical learning tasks may depend on different underlying cognitive or neural processes. In sum, because psychological variables can only be measured indirectly, it is hard to verify which variables have specifically been changed and manipulated in any given paradigm.

Evaluating statistical learning phenomena

As we have seen, the challenges facing statistical learning research make it difficult to formulate accurate and reliable theories. To further illustrate how these issues pertain to statistical learning research, we focus on four phenomena that have been examined to varying degrees in the research literature: modality effects; age effects; species differences; and atypical statistical learning in developmental disorders, such as dyslexia. For each we consider (1) the robustness of the phenomena, (2) potential issues with construct validity, and (3) challenges with establishing causality. With this approach, we are advocating a “bottom-up” stance for theory-building. That is, before creating theories to explain particular phenomena, it is necessary to have a solid understanding of the robustness of the phenomena themselves, which includes assessing their validity and potential causal relationships between psychological variables.

Modality effects

Statistical learning has been observed across a range of scenarios, from auditory⁶¹ and visual scene analysis¹⁰ to language processing⁶² and motor learning^{63,64}. However, comparing across modalities has indicated that the conditions that optimize statistical learning vary depending on both modality (e.g., visual vs. auditory) and domain (e.g., linguistic vs. nonlinguistic)^{39,65–69}, for a summary, also see ref. 70. These modality-specific representations may exist alongside modality-independent representations and domain-general mechanisms^{2,71}, although this remains unclear⁷².

Robustness. A key theory² of statistical learning posits that statistical learning exists across modalities but has modality-specific constraints. This is supported by consistent evidence for differences between modalities; for example, performance in the auditory modality is typically superior to the visual modality for the serial presentation of materials^{39,67}, especially at faster rates^{65,66}. On the other hand, the auditory advantage seems to disappear when the visual stimuli are simultaneously presented^{65,67}. There are also asymmetries across modalities for the same domain, for example, visual tasks show a non-linguistic advantage while

auditory tasks show a linguistic advantage⁷³. Furthermore, modalities appear to operate with relative independence, showing limited transfer⁷⁴ (though also see ref. 75 for a possible exception and Milne, Wilson & Christiansen⁷⁰ for further discussion). Additionally, there is a lack of correlation across modalities^{52,76,77} as well as different developmental trajectories for statistical learning in different modalities⁷⁸. The presence of modality differences across a broad range of tasks, stimuli and research groups speaks to the robustness of this phenomenon. However, the same theory² suggests the presence of a unitary general statistical learning mechanism, such as the hippocampus, yet the existence and nature of such a mechanism is still unclear (see Bogaerts et al.⁷² for a discussion on whether a “good statistical learner” exists).

Construct validity. Modality effects may be greatly influenced by construct validity. For example, what is the best way to measure statistical learning in each modality to provide a “fair” comparison between them? That is, how can we make methods comparable across modalities, while recognizing that there are some fundamental differences across sensory systems? At a minimum, discriminability of the individual items should be comparable, as Conway & Christiansen³⁹ ensured in a pre-training phase before exposure to the statistical input stream. Even when such methodological controls are used to equate stimuli across modalities, it is difficult to isolate statistical learning itself from other cognitive processes. For example, if modality differences are observed, is it due to statistical learning differences, or the role of working memory^{79–81} or attention mechanisms across modalities (e.g., object-based attention⁸²)? Alternatively, many participants impose a linguistic code on visual and auditory stimuli during statistical learning tasks⁶⁸, which makes it difficult to isolate pure auditory and visual learning. Relatedly, if participants are using explicit strategies during training and testing, these could both mask or exacerbate differences in modalities; this may particularly be the case for commonly used two-alternative forced-choice tasks^{6,72}. Finally, many publications comparing across modalities study individual differences; therefore ensuring the retest-reliability of the methods, and that they capture sufficient variance, is especially critical⁵¹.

Causality. If modality effects exist, we also need to establish what is driving them to inform any theoretical framework. Establishing causality relates closely to the issue of construct validity. While construct validity asks us to ensure we are measuring what we think we are measuring, delineating the related perceptual factors, cognitive functions and potential strategies will also inform us on how the differences emerge and their relationship with statistical learning. In this process, we should also consider the fundamental reason why statistical learning exists. In most cases, this is to influence the way we interact with the world, be that in terms of the immediate allocation of cognitive resources (e.g.,⁸³) or the long-term acquisition of speech regularities¹⁵. Better appreciating why the brain may have adapted to use statistics in a given modality could help to understand the differences or similarities in the causal relationships between statistical learning and other psychological variables.

Age effects

There are three ideas that have been advanced regarding how statistical learning abilities may interact with age. First, it has been suggested that statistical learning is age-invariant, with some studies showing that learning was unaffected by age^{18,41,84,85}. Second, like many other cognitive abilities, statistical learning might improve with age^{86–91} (however, such an effect may be due to changes in other cognitive abilities such as working memory, which are known to improve with age⁹²: see “Causality” below). Third, given the importance of statistical learning in language acquisition, it has been argued that infants and children show better statistical learning compared to adults. Infancy and early childhood are where the majority of language learning typically takes place, with some suggestion that language learning abilities decline with age^{93–95}.

Robustness. Longitudinal studies are especially important for understanding potential age effects in statistical learning; indeed, a recent longitudinal study provides support for better statistical learning in younger compared to older children between the ages of 7 and 14⁹⁶. However, there is at least some evidence supporting all three possibilities for how statistical learning and age interact (age-invariance, childhood superiority, and adult superiority), effectively weakening the robustness of all of them. Complicating matters further, there is also evidence that any age-related changes in statistical learning may be influenced by modality effects: Raviv and Arnon⁴¹ found improvements in visual statistical learning between 5- and 12-year-olds, but no changes in statistical learning in the auditory domain. However, in a similar study comparing visual statistical learning with auditory statistical learning of non-linguistic stimuli, improvements in both visual and auditory statistical learning were found between 5- and 12-year-olds⁹⁷. This suggests that any modality-specific differences may be further complicated by differences between stimuli, both of which may affect the conclusions that can be drawn from the literature regarding the developmental trajectory of statistical learning. In sum, the nature of developmental changes across age is not a robust phenomenon, further complicated by the apparent influence of modality- and stimulus-specific effects on the development of statistical learning. On the other hand, there is good evidence that humans of varying ages (infants through adults) are all capable of statistical learning, so the presence of statistical learning across ages appears to be a robust finding.

One critical methodological consideration that specifically impacts our ability to determine changes in statistical learning across development relates to the “more room to improve” effect⁹⁸. This refers to the difficulties that occur when trying to compare different age groups who have different baseline reaction times. While high variability between age groups is often controlled for by using standardized scores, in learning experiments specifically, this variability is actually a critical aspect of the learning process. Specifically, individuals who show greater learning effects typically show greater variability relative to poorer learners. Therefore, using standardized scores may obscure key learning differences, rendering comparisons between age groups uninformative. More recently, efforts have been made to overcome such challenges, for example, by controlling for average speed when using reaction time-based measures of learning⁹⁸.

Construct validity. There are issues relating to the construct validity of statistical learning that are particularly important to consider when measuring people of different ages. While statistical learning in infancy was originally measured using preferential-looking paradigms – which offer an implicit method of measuring learning when the participants are not able to provide explicit judgements—when measuring learning in adults and children, learning is typically evidenced through above chance performance in a grammaticality judgement or familiarity task. In these tasks, participants are asked to provide explicit judgements on whether a sequence fits or breaks a pattern to which they have previously been exposed. Because these tasks measure learning directly by asking participants to classify sequences (often based on ‘gut instinct’), performance may be affected by additional cognitive abilities, such as understanding the task instructions and decision-making skills, which are known to improve across development⁹⁹. As an example, “yes” biases—in this case a preference for categorizing sequences as grammatical—are commonly reported in AGL tasks with children^{100,101}, which may suggest that children have a lower threshold for classifying sequences as grammatical, that they pick up and use irrelevant features of the sequences to judge their grammaticality, or that they simply do not understand how to complete the task. In any case, it is likely that such explicit measures of statistical learning may underestimate children’s abilities, which makes comparing statistical learning across development using these tasks difficult.

More recently, alternative tasks have been developed that assess statistical learning indirectly, by measuring other variables that are facilitated by statistical learning. For example, SRT tasks^{102,103}, serial recall tasks^{104–106},

and tasks measuring neural entrainment¹⁰⁷ can assess statistical learning without requiring explicit reflection on what has been learned. Indirect measures of learning also typically provide additional benefits over direct measures in that they can measure learning over the course of the task, while learning is taking place. Therefore, indirect measures of learning may provide a useful method of comparing statistical learning abilities across age, by avoiding the requirement for more explicit processes which may not accurately represent children's performance. Additionally, these tasks provide information about the trajectory of learning across the task, which could reveal less obvious differences in learning patterns between children and adults.

Causality. As we have discussed, the developmental trajectory of statistical learning, from infancy to adulthood, is already somewhat unclear, both due to a lack of robust age effects and the issues surrounding the validity of the experimental measures used. These issues are compounded by (or possibly caused by, at least in part) challenges with understanding causal relationships between different psychological variables. Aside from statistical learning, myriad cognitive processes change and develop as we age, including attention, memory, executive functions, and language. It is very likely that these abilities contribute to performance on statistical learning tasks, and each of these, and their interactions, could impact the results of developmental studies. While establishing the causal relationships between different psychological processes is a challenge for all statistical learning research, this issue is compounded in developmental research, due to the co-development of so many other aspects of cognition. However, examining changes in other statistical learning phenomena across development using longitudinal studies may actually aid attempts to establish causality. By measuring these other related cognitive functions alongside statistical learning at various timepoints in development, we may be better placed to draw conclusions regarding which aspects of cognition – or other environmental factors such as education – influence the development of statistical learning.

Species differences

Over the last two decades, much comparative research has investigated the extent to which statistical learning abilities might be shared by species other than humans. This research has used a wide range of methods (see Construct Validity, below), but here we will primarily focus on those that assess learning under incidental conditions, particularly in nonhuman primates, as opposed to explicit, operant training (which is common in birds and rodents). These approaches have been used to demonstrate statistical learning abilities in a wide range of primates (tamarins^{108–110}, marmosets^{21,111}, squirrel monkeys¹¹², rhesus macaques^{68,111,113,114}, baboons^{115–117}, and chimpanzees^{118,119}). Beyond the general existence of statistical learning in nonhuman animals, considerable amounts of comparative research have focused on the specific types of statistical regularities that animals are able to learn, with some studies suggesting there may be differences across species.

Robustness. The general phenomenon that statistical learning is conserved in other primates (and likely more broadly in the animal kingdom), appears to be highly robust^{42,120}. However, while all animals tested learn relatively simple 'adjacent dependencies' (where stimuli co-occur immediately in space or time), 'nonadjacent dependencies' (in which intervening elements separate these stimuli), may prove more difficult to learn (for a review, see¹²¹). Moreover, more complex classes of stimuli (e.g., governed by supra-regular grammars, e.g.,¹⁰⁸) have typically proved too complex for nonhuman primates to learn, at least under incidental conditions^{108,122} (although for a successful demonstration using operant approaches, see¹²³). Therefore, the central phenomenon, that core statistical learning abilities exist across the primate taxa (and likely more broadly in the animal kingdom), appears to be highly robust. However, as the complexity of these statistical relationships increases, species differences become pronounced. Here, it is important to note that these results may not represent differences in statistical learning per se, but may

instead relate to other species differences (e.g., perception, attention, motivation, etc.) as will be discussed in 'Causality', below.

Construct validity. A central challenge of studying animal cognition is the requirement that tasks are both feasible in nonhuman animals and also accurately measure the constructs of interest, in this case, statistical learning. Early research in nonhuman primates made use of paradigms originally developed to test preverbal infants, such as preferential looking or habituation/dishabituation paradigms measuring head turn responses^{21,108–110,112}, or equivalent responses using eye-tracking^{68,111,114}. While these approaches allow reasonable comparisons of monkey and infant data, most adult human statistical learning research uses alternative methods (e.g., grammaticality judgement tasks), which are not feasible in nonhuman animals. A second popular method for testing nonhuman primate statistical learning is the use of SRT tasks, which have been combined with a range of different artificial grammars (e.g.,^{113,115,116}), in which animals make a series of rapid responses with a joystick or on a touchscreen. These tasks can allow direct comparisons with humans (as in ref. 116). However, SRT tasks may rely on somewhat different, more procedural learning processes compared to preferential looking paradigms, raising further challenges in assessing construct validity and particularly to drawing comparisons across studies and experimental designs. Additionally, adaptations to these designs allow the investigation of how primates use statistical regularities to encode sequences of stimuli into increasingly large 'chunks', over the course of training^{124,125}. Given this range of approaches, many cross-species comparisons (either within a single comparative study (e.g.,^{111,114}), or those aggregating across multiple studies) are based on disparate methods, which may, at least to some extent, measure different constructs. In cases where similar patterns of learning are observed, despite such methodological differences, it may be reasonable to conclude that not only do those similarities in statistical learning exist, but that they are also robust even to these methodological differences. However, when differences are observed, this leads to difficulties in concluding whether it is animals' statistical learning abilities that differ, or whether different methods and approaches used across species may actually measure somewhat different psychological constructs.

Causality. The clearest overarching conclusion to be taken from the comparative literature on statistical learning is that many species are capable of learning at least relatively simple statistical relationships. However, when statistical regularities become more complex, species differences appear to emerge. This raises the challenging question of determining whether these differences result from more limited statistical learning abilities in nonhuman animals (as has often been concluded) or whether other differences in perception, motivation, attention, or other psychological variables, may contribute to these effects. One interesting demonstration of this may be the recent move to testing nonhuman primates using more operant training approaches. Using traditional, incidental learning approaches monkeys have not been shown to learn complex, supra-regular grammars. However, using an explicit, operant training approach in which rhesus macaques were rewarded for correct responses over many thousands of trials, monkeys were capable of learning these types of grammars^{123,126}. It is highly likely that these tasks recruit different learning systems, but these task differences also introduce many other differences, including perceptual differences as well as different levels of motivation and attention, due to the promise of reward for correctly completed trials. This highlights the interplay between different psychological variables and the challenges of isolating statistical learning from other factors. While this specific example applies to learning in nonhuman primates, the same issues of perception, motivation, attention, etc., are likely relevant to human studies of statistical learning, and represent one of the major challenges to theory building in statistical learning.

Atypical statistical learning in developmental dyslexia

Finally, there has been much interest in examining whether statistical learning is atypical in different developmental disorders related to language and communication. Because statistical learning appears to be crucial for learning spoken and written language in typical development¹¹, it is possible that certain language and communication disorders might arise from atypical statistical learning¹²⁷. Developmental dyslexia is one such disorder that has been examined through the lens of statistical learning (e.g.,^{128–131}). Note that while we focus here on developmental dyslexia, some of the same questions about robustness, validity, and causality are likely relevant for many other developmental disorders.

Robustness. Based on several meta-analytic reviews, which can be useful for systematically assessing robustness across studies (e.g.,³⁸), the extent of a statistical learning impairment in dyslexia remains unclear. On the one hand, Lum et al.'s¹³² systematic review showed evidence of a statistical learning impairment in those with dyslexia, based on studies using SRT tasks in adults and children. Likewise, van Witteloostuijn et al.'s¹³³ meta-analysis also found evidence of impaired learning in dyslexia, based on studies using AGL paradigms. However, their analysis suggested the presence of publication bias, which makes it harder to gauge the robustness of the relationship between statistical learning and dyslexia. Similarly, Schmalz et al.'s¹³⁴ systematic review of statistical learning in dyslexia using both SRT and AGL tasks also noted the presence of publication bias and concluded that there is “insufficient high-quality data to draw conclusions.” Finally, Singh and Conway¹³⁵ reviewed recent studies that were not included in the three prior meta-analytic and systematic reviews and concluded that there was some evidence for a statistical learning impairment in adults with dyslexia (with 11 of 19 studies showing impairment), but that the evidence for an impairment in children with dyslexia was weaker (with only 6 of 14 studies showing impairment). Singh and Conway¹³⁵ proposed that some of the inconsistent findings could be due to heterogeneity in the tasks and methods used to assess statistical learning, the heterogeneity of dyslexia itself, and the role of publication bias. Thus, the finding of impaired statistical learning in developmental dyslexia does not appear to be a robust phenomenon.

Construct validity. As mentioned earlier, statistical learning research uses a variety of task methodologies, and it is currently unclear to what extent they measure the same construct. This in turn can lead to discrepancies across studies intending to measure statistical learning in dyslexia. For instance, if two studies examine statistical learning in dyslexia using different tasks, and the findings are inconsistent with each other, it is unclear whether such differences are due to differences in the use of the tasks or some other factor^{11,12,135}. The solution, of course, is not to focus on only one statistical learning task, but instead to carry out systematic investigations of dyslexia using a variety of tasks, chosen based on theoretical reasons³³, and used across multiple studies. And as is true for the other phenomena reviewed above, construct validity requires increasing our understanding of how different tasks engage different brain systems and therefore different underlying cognitive processes (e.g., the SRT task is heavily dependent on the basal ganglia^{136,137}, while the triplet segmentation task instead appears to rely on a combination of neocortical sensory networks and the hippocampus²⁷).

Causality. One issue for research investigating statistical learning in dyslexia is that the findings are essentially correlational in nature: observing a statistical learning deficit in individuals with dyslexia compared to a control group is consistent with the idea that the learning deficit caused reading problems, but it is not the only possible conclusion. Some other cognitive function (i.e., a “third variable” such as attention or phonological processing) may impact both statistical learning performance and reading ability. Moreover, statistical learning impairments may not cause developmental dyslexia

but rather result from it. To establish the causal nature of a statistical learning-dyslexia link would require experimentally influencing statistical learning ability—through the use of neurostimulation (e.g., TMS, tDCS^{138,139} or possibly dual-task interference tasks^{140,141})—and then observing how such manipulations impact reading ability (i.e., inducing a “virtual lesion”¹⁴²). However, such attempts run into Eronen and Bringmann's³⁰ aforementioned “fat-handed” problem: where direct attempts to experimentally manipulate statistical learning ability may instead affect related cognitive processes such as attention or working memory. Further complicating attempts at causality is the heterogeneity within dyslexia in terms of variability in reading¹⁴³ and spelling ability profiles^{30,144}. There are also often co-occurring conditions with developmental dyslexia (e.g., attention, executive function, and motor skills^{143,145–147}). The heterogeneity and possible comorbidity of developmental dyslexia presents a further challenge to establishing causality, because it is possible that even if statistical learning has a causal influence on reading ability and disability, it may only be the case for certain individuals.

Discussion and recommendations

In sum, the statistical learning phenomena reviewed here vary in their robustness and have similar though also unique issues related to construct validity and causality. Some issues are generalizable to much of statistical learning research (e.g., separating statistical learning abilities from other psychological processes such as memory, attention and motivation), whereas others are more specific to each area (e.g., unavoidable methodological differences when testing different ages or species). As discussed above, the general phenomenon of statistical learning—incidentally acquiring information from environmental stimuli—is highly robust. Based on the present evidence, modality effects, specifically the auditory-serial and visual-spatial advantages for learning, also appears to be a relatively robust finding. Likewise, the presence of statistical learning across age and across nonhuman species also appear to be robust, though it is likely that not all species learn all types of patterns equally well. The extent to which statistical learning changes with age, as well as to what extent statistical learning is impaired in developmental dyslexia, do not seem to be robust phenomena. For all four areas of research, crucial questions about construct validity and causality must be addressed to ensure an accurate understanding of the phenomena in question. While we focused on these four phenomena as ones that have received substantial coverage in the research literature, we acknowledge that many other relevant statistical learning phenomena also exist, and may face similar challenges related to robustness, validity, and causality.

While each statistical learning research phenomenon varies in the extent and nature of these challenges, and thus may require unique solutions, we believe the following recommendations are general enough to broadly deal with many of the most important and common issues facing statistical learning research. We also note that some of these recommendations may be useful for other areas of psychological research that face similar challenges.

There may be value in taking a “bottom-up” approach to theory building

Eronen and Bringmann³⁰ (see also^{29,148–150}) suggested that psychology needs more “phenomenon-driven” research, with the focus on identifying what are the robust phenomena, so that we know what findings theories must encapsulate. Discovering new (and robust) phenomena will help us constrain the possible theory space so that we avoid attempting to develop theoretical frameworks to explain findings that turn out to not be robust. Furthermore, as our understanding of robust phenomena increases, this allows one to see abstract patterns across them that would not be otherwise possible without such knowledge. Note, a bottom-up approach to theory-building does not preclude a top-down approach to experimental design or hypothesis-driven research; in fact, these are complementary, not antagonistic, approaches.

Relevant concepts need to be clearly and transparently defined

As mentioned earlier, statistical learning is sometimes defined differently across studies and assessed in different ways^{11,53}, potentially causing serious issues whereby a single term might be used to refer to several quite different cognitive or neural processes¹⁴. Until we have a clear and accurate understanding of relevant constructs and how best to measure them, it will be impossible to know what the robust phenomena are and how constructs causally relate to one another. Thus, while much recent focus has been on the reliability of statistical learning measures^{50,51}, and rightly so, the issue of construct validity is also critical to consider. Relatedly, it is important to point out that using different measures for statistical learning is not inherently counterproductive (see next recommendation); however, we should be aware that using different methodological approaches for studying statistical learning may mean we are not measuring the same processes.

The field will benefit from a diversity of statistical learning measures

While it is important to be mindful of methodological differences across statistical learning tasks, a wide range of methods are needed to fully understand the relevant phenomena. As Eronen and Bringmann³⁰ pointed out, robustness is improved if a given phenomenon is observed across multiple methods and measures. Thus, we suggest that future research investigating statistical learning phenomena—even within individual studies—should incorporate multiple measures of statistical learning, to determine the robustness and generalizability of any such findings, rather than focusing on a single task⁹. A more complete understanding of statistical learning requires a deep exploration of how differences in the type of task, nature of the input, individual differences, and other factors affect the learning process and the resulting knowledge that is acquired. Even so, a diversity of measures is only valuable if we know what they are measuring (i.e., again, good construct validity is necessary).

There is value in methods that can be commonly used across different modalities, ages, species, and populations

While we recommend incorporating a diversity of (valid) statistical learning measures, some areas of research will benefit from the use of common methods. For instance, using equivalent methods and measures may be important when attempting to examine potential commonalities and/or differences across modalities, age, species, or populations. As pointed out above, this may be especially true when interpreting observed learning differences (e.g., across age, or species, etc.). If a difference across age or species is observed when using different methods or tasks, then it is not clear if the differences are due to differences in methods or actual differences due to age or species. For instance, some methods such as the SRT paradigm have been used successfully across species and across much of the human lifespan^{94,116}; similarly, neural measures can also be adapted across different ages, species, in addition to acquiring behavioral data.

Statistical learning should not be studied in a vacuum

To better understand statistical learning phenomena, it is also important to measure as many other relevant processes and variables as possible to see what other factors may be associated with statistical learning⁹. For instance, several lines of research suggest that there may be a competitive relationship between executive functions and statistical learning^{71,141}. Consequently, if a difference in statistical learning is uncovered between species, or between ages, it is important to determine if such differences are due to statistical learning per se or to differences in executive function, or some other psychological variable that may also relate to statistical learning, such as motivation or attention. To solve this issue requires both incorporating well-validated measures not just of statistical learning but of other relevant psychological constructs and considering the interactions between them.

Interactions among phenomena should be explored

For convenience, the four sample phenomena evaluated above were presented largely independent of one another; however, in reality, the various

phenomena interact in potentially complex ways. For instance, as noted above, there is some suggestion that the developmental trajectory of statistical learning is affected by modality, with improvements in visual but not auditory statistical learning being found across age. Likewise, modality effects observed in humans may or may not generalize to nonhumans⁶⁸. There is also evidence that statistical learning impairments in developmental dyslexia may depend at least somewhat on age and possibly sensory modality as well (see¹³⁵). Thus, to fully evaluate and understand each phenomenon may require understanding their interactions with one another.

Conclusion

Despite the surge of interest in statistical learning, there are a number of critical challenges facing this area of research^{9,11–13}. We have argued that these challenges are similar to those facing the psychological sciences more generally. Relying on recent conceptual advances in contending with the theory crisis³⁰, we have evaluated a sample of statistical learning research areas in terms of the robustness of each phenomenon, issues with construct validity, and difficulties with establishing causality. We have also offered recommendations for how to deal with these issues in a way that we believe will lead to promising areas for theory-building, specifically by focusing on ways to better establish robustness, validity, and causality in statistical learning research.

We hope this analysis and the recommendations contained within will improve future attempts at constructing theories to explain various statistical learning phenomena. Such theories would need to be pitched at both a higher level—i.e., what neural mechanisms and cognitive processes underlie statistical learning itself³² and how they relate to other psychological and cognitive constructs—but also within the embedded sub-phenomena reviewed here—i.e., theories to explain how and why differences exist (or not) across modality, age, species, and developmental disorders. The development of such theories will be crucially dependent upon the continued identification and exploration of the robust phenomena, the refinement of a diverse set of valid measures, and an understanding of the causal relationships relevant to statistical learning.

Data availability

No datasets were generated or analysed during the current study.

Received: 12 February 2025; Accepted: 25 August 2025;

Published online: 29 September 2025

References

- Aslin, R. N. Statistical learning: a powerful mechanism that operates by mere exposure. *WIREs Cogn. Sci.* **8**, e1373 (2017).
- Frost, R., Armstrong, B. C., Siegelman, N. & Christiansen, M. H. Domain generality versus modality specificity: the paradox of statistical learning. *Trends Cogn. Sci.* **19**, 117–125 (2015).
- Saffran, J. R. & Kirkham, N. Z. Infant statistical learning. *Annu. Rev. Psychol.* **69**, 181–203 (2018).
- Saffran, J. R., Aslin, R. N. & Newport, E. L. Statistical learning by 8-month-old infants. *Science* **274**, 1926–1928 (1996).
- Reber, A. S. Implicit learning of artificial grammars. *J. Verbal Learn. Verbal Behav.* **6**, 855–863 (1967).
- Christiansen, M. H. Implicit statistical learning: a tale of two literatures. *Top. Cogn. Sci.* **11**, 468–481 (2019).
- 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).
- Sherman, B. E., Graves, K. N. & Turk-Browne, N. B. The prevalence and importance of statistical learning in human cognition and behavior. *Curr. Opin. Behav. Sci.* **32**, 15–20 (2020).
- Frost, R., Armstrong, B. C. & Christiansen, M. H. Statistical learning research: a critical review and possible new directions. *Psychol. Bull.* **145**, 1128–1153 (2019).

10. 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).
11. Arciuli, J. & Conway, C. M. The promise—and challenge—of statistical learning for elucidating atypical language development. *Curr. Dir. Psychol. Sci.* **27**, 492–500 (2018).
12. Bogaerts, L., Frost, R. & Christiansen, M. H. Integrating statistical learning into cognitive science. *J. Mem. Lang.* **115**, 104167 (2020).
13. Siegelman, N. Statistical learning abilities and their relation to language. *Lang. Linguist. Compass* **14**, e12365 (2020).
14. Willingham, D. B. & Preuss, L. The death of implicit memory. *Psyche* **2**, 1–10 (1995).
15. Erickson, L. C. & Thiessen, E. D. Statistical learning of language: theory, validity, and predictions of a statistical learning account of language acquisition. *Dev. Rev.* **37**, 66–108 (2015).
16. Thiessen, E. D., Girard, S. & Erickson, L. C. Statistical learning and the critical period: how a continuous learning mechanism can give rise to discontinuous learning. *WIREs Cogn. Sci.* **7**, 276–288 (2016).
17. Thiessen, E. D. & Erickson, L. C. Beyond word segmentation: a two-process account of statistical learning. *Curr. Dir. Psychol. Sci.* **22**, 239–243 (2013).
18. Thiessen, E. D., Kronstein, A. T. & Hufnagle, D. G. The extraction and integration framework: a two-process account of statistical learning. *Psychol. Bull.* **139**, 792–814 (2013).
19. Perruchet, P. What mechanisms underlie implicit statistical learning? Transitional probabilities versus chunks in language learning. *Top. Cogn. Sci.* **11**, 520–535 (2019).
20. Perruchet, P. & Pacton, S. Implicit learning and statistical learning: one phenomenon, two approaches. *Trends Cogn. Sci.* **10**, 233–238 (2006).
21. Reber, S. A. et al. Common marmosets are sensitive to simple dependencies at variable distances in an artificial grammar. *Evol. Hum. Behav.* **40**, 214–221 (2019).
22. Arciuli, J. The multi-component nature of statistical learning. *Philos. Trans. R. Soc. B Biol. Sci.* **372**, 20160058 (2017).
23. Daltrozzo, J. & Conway, C. M. Neurocognitive mechanisms of statistical-sequential learning: what do event-related potentials tell us? *Front. Hum. Neurosci.* **8**, 437 (2014).
24. Dehaene, S., Meyniel, F., Wacongne, C., Wang, L. & Pallier, C. The neural representation of sequences: from transition probabilities to algebraic patterns and linguistic trees. *Neuron* **88**, 2–19 (2015).
25. Savalia, T., Shukla, A. & Bapi, R. S. A unified theoretical framework for cognitive sequencing. *Front. Psychol.* **7**, 1821 (2016).
26. Gigerenzer, G. Personal reflections on theory and psychology. *Theory Psychol.* **20**, 733–743 (2010).
27. Muthukrishna, M. & Henrich, J. A problem in theory. *Nat. Hum. Behav.* **3**, 221–229 (2019).
28. Oberauer, K. & Lewandowsky, S. Addressing the theory crisis in psychology. *Psychon. Bull. Rev.* **26**, 1596–1618 (2019).
29. Borsboom, D., van der Maas, H. L. J., Dalege, J., Kievit, R. A. & Haig, B. D. Theory construction methodology: a practical framework for building theories in psychology. *Perspect. Psychol. Sci.* **16**, 756–766 (2021).
30. Eronen, M. I. & Bringmann, L. F. The theory crisis in psychology: how to move forward. *Perspect. Psychol. Sci.* **16**, 779–788 (2021).
31. Marr, D. Vision: A computational investigation into the human representation and processing of visual information. <https://doi.org/10.7551/mitpress/9780262514620.001.0001> (The MIT Press, 2010).
32. Conway, C. M., Janacek, K., Buffington, J. & Ullman, M. T. The what, how, and where of statistical learning. *Nature Reviews Neuroscience*. (under review).
33. MacLeod, C. M. Half a century of research on the Stroop effect: an integrative review. *Psychol. Bull.* **109**, 163–203 (1991).
34. Stroop, J. R. Studies of interference in serial verbal reactions. *J. Exp. Psychol.* **18**, 643–662 (1935).
35. Greenwald, A. G. & Banaji, M. R. Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychol. Rev.* **102**, 4–27 (1995).
36. Eronen, M. I. Robustness and reality. *Synthese* **192**, 3961–3977 (2015).
37. Eronen, M. I. Robust realism for the life sciences. *Synthese* **196**, 2341–2354 (2019).
38. Isbilen, E. S. & Christiansen, M. H. Statistical learning of language: a meta-analysis into 25 years of research. *Cogn. Sci.* **46**, e13198 (2022).
39. Conway, C. M. & Christiansen, M. H. Modality-constrained statistical learning of tactile, visual, and auditory sequences. *J. Exp. Psychol. Learn. Mem. Cogn.* **31**, 24–39 (2005).
40. Pothos, E. An entropy model for artificial grammar learning. *Front. Psychol.* **1**, 16 (2010).
41. 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**, e12593 (2018).
42. Santolin, C. & Saffran, J. R. Constraints on statistical learning across species. *Trends Cogn. Sci.* **22**, 52–63 (2018).
43. Turk-Browne, N. B., Jungé, J. A. & Scholl, B. J. The automaticity of visual statistical learning. *J. Exp. Psychol. Gen.* **134**, 552–564 (2005).
44. Pedraza, F. et al. Evidence for a competitive relationship between executive functions and statistical learning. *npj Sci. Learn.* **9**, 30 (2024).
45. Saffran, J. R. Statistical learning as a window into developmental disabilities. *J. Neurodev. Disord.* **10**, 35 (2018).
46. Borsboom, D., Mellenbergh, G. J. & van Heerden, J. The concept of validity. *Psychol. Rev.* **111**, 1061–1071 (2004).
47. Flake, J. K., Pek, J. & Hehman, E. Construct validation in social and personality research: current practice and recommendations. *Soc. Psychol. Personal. Sci.* **8**, 370–378 (2017).
48. Zumbo, B. D. & Chan, E. K. H. Setting the stage for validity and validation in social, behavioral, and health sciences: trends in validation practices. In *Validity and Validation in Social, Behavioral, and Health Sciences* (eds Zumbo, B. D. & Chan, E. K. H.) 3–8 https://doi.org/10.1007/978-3-319-07794-9_1 (Springer International Publishing, Cham, 2014).
49. Arnon, I. Do current statistical learning tasks capture stable individual differences in children? An investigation of task reliability across modality. *Behav. Res.* **52**, 68–81 (2020).
50. Siegelman, N., Bogaerts, L., Christiansen, M. H. & Frost, R. Towards a theory of individual differences in statistical learning. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **372**, 20160059 (2017).
51. Siegelman, N., Bogaerts, L. & Frost, R. Measuring individual differences in statistical learning: Current pitfalls and possible solutions. *Behav. Res.* **49**, 418–432 (2017).
52. Siegelman, N. & Frost, R. Statistical learning as an individual ability: theoretical perspectives and empirical evidence. *J. Mem. Lang.* **81**, 105–120 (2015).
53. Bogaerts, L., Siegelman, N. & Frost, R. Statistical learning and language impairments: toward more precise theoretical accounts. *Perspect. Psychol. Sci.* **16**, 319–337 (2021).
54. Nissen, M. J. & Bullemer, P. Attentional requirements of learning: Evidence from performance measures. *Cogn. Psychol.* **19**, 1–32 (1987).
55. Hebb, D. Distinctive features of learning in the higher animal. in *Brain Mechanisms of Learning* 37–46 (Oxford, UK: Blackwell, 1961).
56. Yu, C. & Smith, L. B. Rapid word learning under uncertainty via cross-situational statistics. *Psychol. Sci.* **18**, 414–420 (2007).
57. Chun, M. M. & Jiang, Y. Contextual cueing: implicit learning and memory of visual context guides spatial attention. *Cogn. Psychol.* **36**, 28–71 (1998).

58. Eronen, M. I. Causal discovery and the problem of psychological interventions. *New Ideas Psychol.* **59**, 100785 (2020).
59. Woodward, J. Methodology, ontology, and interventionism. *Synthese* **192**, 3577–3599 (2015).
60. Center, E. G., Federmeier, K. D. & Beck, D. M. The brain's sensitivity to real-world statistical regularity does not require full attention. *J. Cogn. Neurosci.* **36**, 1715–1740 (2024).
61. Sohoglu, E. & Chait, M. Detecting and representing predictable structure during auditory scene analysis. *eLife* **5**, e19113 (2016).
62. Romberg, A. R. & Saffran, J. R. Statistical learning and language acquisition. *WIREs Cogn. Sci.* **1**, 906–914 (2010).
63. Nemeth, D., Hallgató, E., Janacsek, K., Sándor, T. & Londe, Z. Perceptual and motor factors of implicit skill learning. *NeuroReport* **20**, 1654–1658 (2009).
64. Hallgató, E., Györi-Dani, D., Pekár, J., Janacsek, K. & Nemeth, D. The differential consolidation of perceptual and motor learning in skill acquisition. *Cortex* **49**, 1073–1081 (2013).
65. Conway, C. M. & Christiansen, M. H. Seeing and hearing in space and time: effects of modality and presentation rate on implicit statistical learning. *Eur. J. Cogn. Psychol.* **21**, 561–580 (2009).
66. Emberson, L. L., Conway, C. M. & Christiansen, M. H. Timing is everything: changes in presentation rate have opposite effects on auditory and visual implicit statistical learning. *Q. J. Exp. Psychol.* **64**, 1021–1040 (2011).
67. Lukics, K. S. & Lukács, Á. Modality, presentation, domain and training effects in statistical learning. *Sci. Rep.* **12**, 20878 (2022).
68. Milne, A. E., Petkov, C. I. & Wilson, B. Auditory and visual sequence learning in humans and monkeys using an artificial grammar learning paradigm. *Neuroscience* **389**, 104–117 (2018).
69. Walk, A. M. & Conway, C. M. Cross-domain statistical–sequential dependencies are difficult to learn. *Front. Psychol.* **7**, 250 (2016).
70. Milne, A. E., Wilson, B. & Christiansen, M. Structured sequence learning across sensory modalities in humans and nonhuman primates. *Curr. Opin. Behav. Sci.* **21**, 39–48 (2018).
71. Vékony, T. et al. Modality-specific and modality-independent neural representations work in concert in predictive processes during sequence learning. *Cereb. Cortex* **33**, 7783–7796 (2023).
72. Bogaerts, L., Siegelman, N., Christiansen, M. H. & Frost, R. Is there such a thing as a 'good statistical learner'? *Trends Cogn. Sci.* **26**, 25–37 (2022).
73. Polyanskaya, L. et al. Intermodality differences in statistical learning: phylogenetic and ontogenetic influences. *Ann. N. Y. Acad. Sci.* **1511**, 191–209 (2022).
74. Conway, C. M. & Christiansen, M. H. Statistical learning within and between modalities: pitting abstract against stimulus-specific representations. *Psychol. Sci.* **17**, 905–912 (2006).
75. Durrant, S. J., Cairney, S. A. & Lewis, P. A. Cross-modal transfer of statistical information benefits from sleep. *Cortex* **78**, 85–99 (2016).
76. Ordin, M., Polyanskaya, L. & Samuel, A. G. An evolutionary account of intermodality differences in statistical learning. *Ann. N. Y. Acad. Sci.* **1486**, 76–89 (2021).
77. Zhou, H., van der Ham, S., de Boer, B., Bogaerts, L. & Raviv, L. Modality and stimulus effects on distributional statistical learning: sound vs. sight, time vs. space. *J. Mem. Lang.* **138**, 104531 (2024).
78. Ren, J. & Wang, M. Development of statistical learning ability across modalities, domains, and languages. *J. Exp. Child Psychol.* **226**, 105570 (2023).
79. Baddeley, A. D. & Hitch, G. Working memory. in *Psychology of Learning and Motivation* (ed Bower, G. H.) **8**, 47–89 (Academic Press, 1974).
80. Lehnert, G. & Zimmer, H. D. Modality and domain specific components in auditory and visual working memory tasks. *Cogn. Process* **9**, 53–61 (2008).
81. Zimmer, H. D. Visual and spatial working memory: From boxes to networks. *Neurosci. Biobehav. Rev.* **32**, 1373–1395 (2008).
82. Shinn-Cunningham, B. G. Object-based auditory and visual attention. *Trends Cogn. Sci.* **12**, 182–186 (2008).
83. Milne, A. E., Zhao, S., Tampakaki, C., Bury, G. & Chait, M. Sustained pupil responses are modulated by predictability of auditory sequences. *J. Neurosci.* **41**, 6116–6127 (2021).
84. Jost, E., Conway, C. M., Purdy, J. D., Walk, A. M. & Hendricks, M. A. Exploring the neurodevelopment of visual statistical learning using event-related brain potentials. *Brain Res.* **1597**, 95–107 (2015).
85. 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).
86. Arciuli, J. & Simpson, I. C. Statistical learning in typically developing children: The role of age and speed of stimulus presentation. *Dev. Sci.* **14**, 464–473 (2011).
87. Kirkham, N. Z., Slemmer, J. A. & Johnson, S. P. Visual statistical learning in infancy: evidence for a domain general learning mechanism. *Cognition* **83**, B35–B42 (2002).
88. Kirkham, N. Z., Slemmer, J. A., Richardson, D. C. & Johnson, S. P. Location, location, location: development of spatiotemporal sequence learning in infancy. *Child Dev.* **78**, 1559–1571 (2007).
89. Lukács, Á. & Kemény, F. Development of different forms of skill learning throughout the lifespan. *Cogn. Sci.* **39**, 383–404 (2015).
90. 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).
91. Vaidya, C. J., Huger, M., Howard, D. V. & Howard, J. H. Jr. Developmental differences in implicit learning of spatial context. *Neuropsychology* **21**, 497 (2007).
92. Gathercole, S. E., Pickering, S. J., Ambridge, B. & Wearing, H. The structure of working memory from 4 to 15 years of age. *Dev. Psychol.* **40**, 177 (2004).
93. Birdsong, D. *Second Language Acquisition and the Critical Period Hypothesis* (Routledge, 1999).
94. 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).
95. Newport, E. L. Maturational constraints on language learning. *Cogn. Sci.* **14**, 11–28 (1990).
96. Tóth-Fáber, E., Farkas, B. C., Harmath-Tánczos, T., Nemeth, D. & Janacsek, K. Longitudinal evidence for decreasing statistical learning abilities across childhood. <https://doi.org/10.31234/osf.io/gj3hq> (2024).
97. Shufaniya, A. & Arnon, I. Statistical learning is not age-invariant during childhood: performance improves with age across modality. *Cogn. Sci.* **42**, 3100–3115 (2018).
98. Juhasz, D., Nemeth, D. & Janacsek, K. Is there more room to improve? The lifespan trajectory of procedural learning and its relationship to the between- and within-group differences in average response times. *PLoS ONE* **14**, e0215116 (2019).
99. Lammertink, I., Van Witteloostuijn, M., Boersma, P., Wijnen, F. & Rispens, J. Auditory statistical learning in children: novel insights from an online measure. *Appl. Psycholinguist.* **40**, 279–302 (2019).
100. Lammertink, I., Boersma, P., Wijnen, F. & Rispens, J. Children with developmental language disorder have an auditory verbal statistical learning deficit: evidence from an online measure. *Lang. Learn.* **70**, 137–178 (2020).
101. van der Lely, H. K., Jones, M. & Marshall, C. R. Who did Buzz see someone? Grammaticality judgement of wh-questions in typically developing children and children with Grammatical-SLI. *Lingua* **121**, 408–422 (2011).
102. Jenkins, H. E., Leung, P., Smith, F., Riches, N. & Wilson, B. Assessing processing-based measures of implicit statistical

- learning: three serial reaction time experiments do not reveal artificial grammar learning. *PLoS ONE* **19**, e0308653 (2024).
103. Misyak, J. B., Christiansen, M. H. & Tomblin, J. B. On-line individual differences in statistical learning predict language processing. *Front. Psychol.* **1**, 31 (2010).
 104. Conway, C. M., Bauernschmidt, A., Huang, S. S. & Pisoni, D. B. Implicit statistical learning in language processing: word predictability is the key. *Cognition* **114**, 356–371 (2010).
 105. Isbilen, E. S., McCauley, S. M., Kidd, E. & Christiansen, M. H. Statistically induced chunking recall: a memory-based approach to statistical learning. *Cogn. Sci.* **44**, e12848 (2020).
 106. Kidd, E. et al. Measuring children’s auditory statistical learning via serial recall. *J. Exp. Child Psychol.* **200**, 104964 (2020).
 107. Moreau, C. N., Joanisse, M. F., Mulgrew, J. & Batterink, L. J. No statistical learning advantage in children over adults: evidence from behaviour and neural entrainment. *Dev. Cogn. Neurosci.* **57**, 101154 (2022).
 108. Fitch, W. T. & Hauser, M. D. Computational constraints on syntactic processing in a nonhuman primate. *Science* **303**, 377–380 (2004).
 109. Newport, E. L., Hauser, M. D., Spaepen, G. & Aslin, R. N. Learning at a distance II. Statistical learning of non-adjacent dependencies in a non-human primate. *Cogn. Psychol.* **49**, 85–117 (2004).
 110. Saffran, J. et al. Grammatical pattern learning by human infants and cotton-top tamarin monkeys. *Cognition* **107**, 479–500 (2008).
 111. Wilson, B. et al. Auditory artificial grammar learning in macaque and marmoset monkeys. *J. Neurosci.* **33**, 18825–18835 (2013).
 112. Ravignani, A., Sonnweber, R.-S., Stobbe, N. & Fitch, W. T. Action at a distance: dependency sensitivity in a New World primate. *Biol. Lett.* **9**, 20130852 (2013).
 113. Heimbauer, L. A., Conway, C. M., Christiansen, M. H., Beran, M. J. & Owren, M. J. Visual artificial grammar learning by rhesus macaques (*Macaca mulatta*): exploring the role of grammar complexity and sequence length. *Anim. Cogn.* **21**, 267–284 (2018).
 114. Wilson, B., Smith, K. & Petkov, C. I. Mixed-complexity artificial grammar learning in humans and macaque monkeys: evaluating learning strategies. *Eur. J. Neurosci.* **41**, 568–578 (2015).
 115. Malassis, R., Rey, A. & Fagot, J. Non-adjacent dependencies processing in human and non-human primates. *Cogn. Sci.* **42**, 1677–1699 (2018).
 116. Rey, A., Minier, L., Malassis, R., Bogaerts, L. & Fagot, J. Regularity extraction across species: associative learning mechanisms shared by human and non-human primates. *Top. Cogn. Sci.* **11**, 573–586 (2019).
 117. Yeaton, J., Tosatto, L., Fagot, J., Grainger, J. & Rey, A. Simple questions on simple associations: regularity extraction in non-human primates. *Learn Behav.* **51**, 392–401 (2023).
 118. Endress, A. D., Cahill, D., Block, S., Watumull, J. & Hauser, M. D. Evidence of an evolutionary precursor to human language affixation in a non-human primate. *Biol. Lett.* **5**, 749–751 (2009).
 119. Watson, S. K. et al. Nonadjacent dependency processing in monkeys, apes, and humans. *Sci. Adv.* **6**, eabb0725 (2020).
 120. Wilson, B., Marslen-Wilson, W. D. & Petkov, C. I. Conserved sequence processing in primate frontal cortex. *Trends Neurosci.* **40**, 72–82 (2017).
 121. Wilson, B. et al. Non-adjacent dependency learning in humans and other animals. *Top. Cogn. Sci.* **12**, 843–858 (2020).
 122. van Heijningen, C. A. A., de Visser, J., Zuidema, W. & ten Cate, C. Simple rules can explain discrimination of putative recursive syntactic structures by a songbird species. *Proc. Natl Acad. Sci.* **106**, 20538–20543 (2009).
 123. Jiang, X. et al. Production of supra-regular spatial sequences by Macaque monkeys. *Curr. Biol.* **28**, 1851–1859.e4 (2018).
 124. Tosatto, L., Fagot, J., Nemeth, D. & Rey, A. Chunking as a function of sequence length. *Anim. Cogn.* **28**, 2 (2024).
 125. Tosatto, L., Fagot, J., Nemeth, D. & Rey, A. The evolution of chunks in sequence learning. *Cogn. Sci.* **46**, e13124 (2022).
 126. Fitch, W. T. Bio-linguistics: monkeys break through the syntax barrier. *Curr. Biol.* **28**, R695–R697 (2018).
 127. Plante, E. & Gómez, R. L. Learning without trying: the clinical relevance of statistical learning. *Lang. Speech Hear. Serv. Sch.* **49**, 710–722 (2018).
 128. Ballan, R., Durrant, S. J., Manoach, D. S. & Gabay, Y. Failure to consolidate statistical learning in developmental dyslexia. *Psychon. Bull. Rev.* **30**, 160–173 (2023).
 129. Daikoku, T. et al. Neural correlates of statistical learning in developmental dyslexia: an electroencephalography study. *Biol. Psychol.* **181**, 108592 (2023).
 130. Gabay, Y., Thiessen, E. D. & Holt, L. L. Impaired statistical learning in developmental dyslexia. *J. Speech Lang. Hear. Res.* **58**, 934–945 (2015).
 131. Ozernov-Palchik, O., Qi, Z., Beach, S. D. & Gabrieli, J. D. E. Intact procedural memory and impaired auditory statistical learning in adults with dyslexia. *Neuropsychologia* **188**, 108638 (2023).
 132. Lum, J. A. G., Ullman, M. T. & Conti-Ramsden, G. Procedural learning is impaired in dyslexia: evidence from a meta-analysis of serial reaction time studies. *Res. Dev. Disabil.* **34**, 3460–3476 (2013).
 133. van Witteloostuijn, M., Boersma, P., Wijnen, F. & Rispens, J. Visual artificial grammar learning in dyslexia: a meta-analysis. *Res. Dev. Disabil.* **70**, 126–137 (2017).
 134. Schmalz, X., Altoè, G. & Mulatti, C. Statistical learning and dyslexia: a systematic review. *Ann. Dyslexia* **67**, 147–162 (2017).
 135. Singh, S. & Conway, C. M. Unraveling the interconnections between statistical learning and dyslexia: a review of recent empirical studies. *Front. Hum. Neurosci.* **15**, 734179 (2021).
 136. Doyon, J. et al. Contributions of the basal ganglia and functionally related brain structures to motor learning. *Behav. Brain Res.* **199**, 61–75 (2009).
 137. Ullman, M. T., Earle, F. S., Walenski, M. & Janacsek, K. The neurocognition of developmental disorders of language. *Annu. Rev. Psychol.* **71**, 389–417 (2020).
 138. Ambrus, G. G. et al. When less is more: Enhanced statistical learning of non-adjacent dependencies after disruption of bilateral DLPFC. *J. Mem. Lang.* **114**, 104144 (2020).
 139. Smalle, E. H. M., Panouilleres, M., Szmalec, A. & Möttönen, R. Language learning in the adult brain: disrupting the dorsolateral prefrontal cortex facilitates word-form learning. *Sci. Rep.* **7**, 13966 (2017).
 140. Hendricks, M. A., Conway, C. M. & Kellogg, R. T. Using dual-task methodology to dissociate automatic from nonautomatic processes involved in artificial grammar learning. *J. Exp. Psychol. Learn. Mem. Cogn.* **39**, 1491–1500 (2013).
 141. Smalle, E. H. M., Daikoku, T., Szmalec, A., Duyck, W. & Möttönen, R. Unlocking adults’ implicit statistical learning by cognitive depletion. *Proc. Natl Acad. Sci.* **119**, e2026011119 (2022).
 142. Woollams, A. M., Madrid, G. & Lambon Ralph, M. A. Using neurostimulation to understand the impact of pre-morbid individual differences on post-lesion outcomes. *Proc. Natl Acad. Sci.* **114**, 12279–12284 (2017).
 143. Snowling, M. J., Hayiou-Thomas, M. E., Nash, H. M. & Hulme, C. Dyslexia and developmental language disorder: comorbid disorders with distinct effects on reading comprehension. *J. Child Psychol. Psychiatry* **61**, 672–680 (2020).
 144. Georgiou, G. K., Martinez, D., Vieira, A. P. A. & Guo, K. Is orthographic knowledge a strength or a weakness in individuals with dyslexia? Evidence from a meta-analysis. *Ann. Dyslexia* **71**, 5–27 (2021).

145. Catts, H. W., Adlof, S. M., Hogan, T. P. & Weismer, S. E. Are specific language impairment and dyslexia distinct disorders? *J. Speech Lang. Hear. Res.* **48**, 1378–1396 (2005).
146. Pennington, B. F. et al. Individual prediction of dyslexia by single versus multiple deficit models. *J. Abnorm. Psychol.* **121**, 212–224 (2012).
147. Gooch, D., Hulme, C., Nash, H. M. & Snowling, M. J. Comorbidities in preschool children at family risk of dyslexia. *J. Child Psychol. Psychiatry* **55**, 237–246 (2014).
148. Haig, B. D. Detecting psychological phenomena: taking bottom-up research seriously. *Am. J. Psychol.* **126**, 135–153 (2013).
149. Houwer, J. D. Why the cognitive approach in psychology would profit from a functional approach and vice versa. *Perspect. Psychol. Sci.* **6**, 202–209 (2011).
150. Hughes, S., De Houwer, J. & Perugini, M. The functional-cognitive framework for psychological research: controversies and resolutions. *Int. J. Psychol.* **51**, 4–14 (2016).
151. Darwin, C. *On the Origin of Species by Means of Natural Selection* (John Murray, London, 1859).

Acknowledgements

Preparation of this manuscript was supported by a Grinnell College Faculty Scholarship Competitive Grant (CMC), a Wellcome Trust Grant (213686/Z/18/Z) (AEM), and by the National Institutes of Health's Office of the Director, Office of Research Infrastructure Programs (P51OD011132) (BW).

Author contributions

All authors (C.M.C., H.E.J., A.E.M., S.S., and B.W.) contributed to the writing of the manuscript and read and approved the final version.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to Christopher M. Conway.

Reprints and permissions information is available at <http://www.nature.com/reprints>

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

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025