Received: 12 May 2019 | Revised: 10 December 2019 | Accepted: 12 December 2019

DOI: 10.1111/lnc3.12365

ARTICLE

WILEY

Statistical learning abilities and their relation to language

1978

Noam Siegelman

Haskins Laboratories

Correspondence
Noam Siegelman, Haskins Laboratories, 300 George street, 06511 New Haven, CT.
Email: noam.siegelman@yale.edu

Funding information
Eunice Kennedy Shriver National Institute of Child Health and Human Development, Grant/Award Number: P20HD091013

Abstract

Numerous studies on statistical learning (SL) have demonstrated humans' sensitivity to complex statistical properties in their sensory environment. These observations have had a profound impact on the study of language, highlighting statistical aspects of the linguistic input that can be learned from experience, leading to the widespread claim that SL plays a key role in language acquisition and processing. But how can this theorized link be experimentally demonstrated? One increasingly popular avenue comes from studies of individual differences, which tie individual variability in SL to variance in linguistic behavior. This review presents the theoretical advances stemming from this line of research, as well as some of the challenges it currently faces. It contends that while previous studies had an important role in establishing the existence of some coarse-grained link between SL and language, recent developments in SL research suggest that the exact nature of this relationship is more complex than originally conceived and is still far from being fully understood. I specifically discuss three outstanding challenges: (a) understanding individual differences in light of the componential nature of SL, (b) mapping the full array of SL processes given the complexity of real-world statistics, and (c) estimating the strength of current empirical evidence while taking into account both positive and null findings. Confronting these issues, I argue, is a necessary step towards a full theory of the role of SL across language.

1 | INTRODUCTION

Over the last two and a half decades, statistical learning (SL) has become a major theoretical construct in the cognitive sciences. The basic premise behind this concept is rooted in the hundred-year-old observation that humans (as well as other species) are constantly bombarded by continuous streams of sensory information (James, 1890). SL is defined as the mechanism by which cognitive systems extract regularities from such sensory stimulation, to discover its underlying structure. As such, SL is considered to play a key role in various cognitive functions, including among others segmentation of continuous input (e.g., Aslin, Saffran, & Newport, 1998), categorization (e.g., Maye, Werker, & Gerken, 2002), and prediction of upcoming signal (e.g., Dale, Duran, & Morehead, 2012). SL computations are thus considered to be involved in all domains of cognition, from low level perception of stimuli (e.g., Barakat, Seitz, & Shams, 2013) to complex cognitive functions such as music appreciation (e.g., Salimpoor, Zald, Zatorre, Dagher, & McIntosh, 2015) and face recognition (e.g., Dotsch, Hassin, & Todorov, 2017).

Yet the main motivation to study SL comes from its proposed link to one cognitive function in particular: language acquisition and use. Thus, over the past few decades, the idea that SL computations are particularly important in language has become increasingly popular. Behind this claim lies a simple argument that can be summarized in the following three steps:

Assumption I Different aspects of language can be characterized as a set of complex statistical regularities.

The view of languages as statistical creatures is not at all new. In fact, early behaviorist studies—more than seven decades before the inception of the term SL—were already concerned with whether patterns embedded in the linguistic input can give rise to high-level linguistic knowledge (see Christiansen, 2018 for a comprehensive historical review). As such, the pioneering work by Esper (1925) showed that participants can extract grammatical categories from an artificial miniature linguistic system, leading him to hypothesize that similar statistical information is also available in natural languages. Following the cognitive revolution, methodologies employing artificial miniature linguistic systems—then labeled artificial grammar learning tasks—were developed extensively by Miller (1958) and Reber (1967). The latter was particularly interested in using these paradigms to investigate "implicit learning" processes, asking how knowledge regarding complex grammatical structures can emerge without intent simply via the assimilation of the underlying structure of presented exemplars.

More recent works, within the modern framework of SL, extended this view to virtually all aspects of language. Thus, for example, works on SL suggest that writing systems can be thought of as an array of correlations between letters (e.g., Chetail, 2017) and between letters to sounds (e.g., Treiman & Kessler, 2006); phonology and phonotactics can be described as sets of co-occurrences between speech sounds (e.g., Onishi, Chambers, & Fisher, 2002), morphology can be thought of as co-occurrences between morphemes (e.g., Pacton, Fayol, & Perruchet, 2005), word knowledge as co-occurrences between words and their referents (e.g., Yu & Smith, 2007). Recent works also revisit the question regarding the statistical nature of syntax, discussing what aspects of syntax can be indeed reduced to sets of regularities spanning words and larger phrases (and see, Saffran & Wilson, 2003; Thompson & Newport, 2007 for a discussion of this more controversial issue).

Assumption II Humans can extract complex statistical regularities from the input using SL computations.

This assertion is based on the definition of SL as a theoretical construct and is supported by numerous experimental demonstrations (see below).

Then, the juxtaposition of Assumption I—languages are essentially a set of statistical regularities—and Assumption II—humans can extract such regularities via SL—leads to the widespread claim that

Conclusion SL plays an important role in language acquisition and use.

Although simple, this argument carries important theoretical implications for the study of language: it highlights aspects of language that can be learned from experience. Thus, a strong interpretation of this argument maintains that all aspects of language can be learned from experience; other, weaker versions do not preclude the possibility of some innate knowledge or constraints, but still emphasize parts of the input that are learned from exposure during development (see Lidz & Gagliardi, 2015; Saffran, 2003; Seidenberg, MacDonald, & Saffran, 2002; Yang, 2004 for discussion). In general, then, this argument goes against the Chomskyan emphasis of the domain-specific, modular, and innate nature of the language acquisition process (Chomsky, 1959) and strengthens usage-based approaches to language (e.g., Tomasello, 2003).

Looking back at the SL literature, it seems that the second part of the argument (Assumption II above)—showing that humans can indeed extract regularities via SL abilities—is well supported by a large body of empirical work. A key study in this line of research was the seminal demonstration of Saffran and her colleagues showing that 8-month-old infants can extract word-like units from continuous input based on the statistical structure of the speech stream alone (Saffran, Aslin, & Newport, 1996), and their follow-up study with adults (Saffran, Newport, & Aslin, 1996).[1] To do so, Saffran and colleagues developed a task (sometimes referred to as the "embedded pattern paradigm" or simply as "the SL task") that consists of two parts: a familiarization phase, followed by a test phase. During the familiarization phase, participants are exposed to a continuous stream of stimuli, which (unbeknownst to participants) is composed of several repeated patterns. To illustrate, the original materials comprise four tri-syllabic "words" (e.g., *tupiro*, *golabu*, *bidaku*, and *padoti*) randomized and concatenated to create a continuous speech-like stream (e.g., *bidakugolabupadotibidakutupiropadoti…*). Importantly, since there are no breaks or other prosodic cues for word boundaries, the only cue to segment the stream to its composing trisyllabic words are the co-occurrences between elements. This can be described as a set of transitional probabilities (TPs) between syllables in the stream, where the extraction of the word *tupiro* involves learning that after the syllable *tu* there is a high probability that the syllable *pi* would appear, which is then likely followed by *ro* (i.e., there is a high TP between *tu* and *pi*, and between *pi* and *ro*), but that after the syllable *ro* multiple syllables can appear, each of them with lower TPs. The second part of the SL task is a test phase, which is used to measure each participant's ability to extract the repeated patterns. In adults (and older children), the test typically consists of a series of two-alternative-forced-choice (2-AFC) trials. In each trial, participants are presented with a pattern from the familiarization phase (i.e., with high TPs between elements), and a foil, composed of elements that appeared in the familiarization phase but with lower TPs (e.g., *piropa* or *tulati*). Participants are asked to choose the pattern that they are more familiar with based on the familiarization phase and are scored according to the number of correct identifications of patterns over foils.[2]

The key to this experimental paradigm is that it can examine whether the sampled participants, on average, possess some ability to assimilate the statistical contingencies embedded in the input. Thus, if no SL computations occur, the expected mean performance in this task would be at chance level, or 50%. Saffran and colleagues showed however that participants' mean performance in this task is significantly above chance-level. This is taken as evidence for successful SL computations at the group-level, demonstrating that humans can indeed extract patterns from the input based on its statistical structure alone.

Since then, a large number of studies have replicated and extended this basic finding, demonstrating the robust and ubiquitous nature of SL. Using modified versions of the embedded pattern paradigm, studies show that SL phenomena are observed in both the auditory (e.g., Endress & Mehler, 2009), and visual (e.g., Kirkham, Slemmer, & Johnson, 2002) modality, with both verbal (e.g., Pelucchi, Hay, & Saffran, 2009) and nonverbal (e.g., Gebhart, Newport, & Aslin, 2009) material, and can be used to detect regularities comprised of adjacent or nonadjacent contingencies (e.g., Gómez, 2002; Newport & Aslin, 2004). Moreover, studies show that SL computations are affected even by subtle changes in the input's statistical structure (e.g., Bogaerts, Siegelman, & Frost, 2016), can operate without overt attention (Evans, Saffran, & Robe-Torres, 2009), and are readily available from a very young age (demonstrated already in 1–3 day newborns, Bulf, Johnson, & Valenza, 2011). Together, all of these studies point to a powerful mechanism, that may help speakers and readers in their quest to master language and use it efficiently.

Importantly, however, showing that SL abilities exist does not necessarily mean that they play a role in language acquisition and processing (or in any other cognitive function). Rather, it is possible that while humans have impressive SL abilities, they do not use these mechanisms when acquiring or using language. In other words, getting back to the three-step argument above, even if indeed humans have the ability to perform SL computations (Assumption II) and even if languages are indeed rich with statistical information (Aassumption I), it still does not necessarily follow that learners and users of language employ SL computations during language acquisition and processing. Alternative views that go against the involvement of SL in language posit that SL abilities cannot scale-up to account for real-world complexity of linguistic inputs. Instead, such views assert that although SL abilities may play a limited role in some aspects of language acquisition or processing, many linguistic functions are not based on the assimilation of statistics due to input's complexity, and these are likely to be rooted in innate, domain-specific linguistic knowledge (Gervain, Nespor, Mazuka, Horie, & Mehler, 2008; Yang, 2004; see Johnson, 2012 for review).

So how can the hypothesized link between SL and language be more directly examined? One promising avenue comes from studies of individual differences. This is because the argument regarding the involvement of SL in language leads to a clear prediction at the individual level. Namely, if indeed SL abilities are linked to language abilities, then individual differences in SL abilities should be predictive of inter-individual variability in language acquisition and use. In simpler words, if indeed SL plays a key role in language, then individuals with better SL abilities should be also characterized by better linguistic abilities, whereas those who exhibit low linguistic capacities should have low SL performance.

Importantly, in order to show that individual differences in SL are related to linguistic abilities (or any other outcome), it is imperative to first show that (a) there is variation in SL abilities (i.e., that not all individuals have the same SL abilities), (b) this variance reflects a stable characteristic of the individual, and (c) it is not nested within other general abilities, such as intelligence quotient or working memory. Empirical evidence confirms these three prerequisites.

Using tasks based on the paradigm developed by Saffran and colleagues, Siegelman and Frost (2015) showed that individuals differ substantially in their ability to extract repeated patterns (triplets of syllables or sounds) from continuous sensory stream, where some participants perform at chance whereas others show near-perfect performance (see also Hunt & Aslin, 2001; Kalra, Gabrieli, & Finn, 2019; Kaufman et al., 2010 for evidence in the context of implicit learning tasks). Moreover, by retesting the same subjects using the same SL tasks, and by examining the correlations between SL performance and other cognitive measures, Siegelman and Frost showed that performance in a given SL task reflects a stable characteristic of an individual, that does not overlap with her/his intelligence quotient or memory capacities (and see Conway, Bauernschmidt, Huang, & Pisoni, 2010; Evans et al., 2009; Isbilen, McCauley, Kidd, & Christiansen, 2017; Kalra et al., 2019; Kaufman et al., 2010; Misyak & Christiansen, 2012, for corroborating evidence). These findings establish individual's SL ability as a potential unique predictor of other abilities, linguistic functions included, allowing to examine whether indeed individual SL abilities predict linguistic performance as hypothesized by theories that emphasize the role of SL in language acquisition and processing.

Individual-differences studies from recent years confirm this prediction, suggesting that individual SL performance in fact predicts variability in linguistic outcomes. In this vein, individual differences in SL performance among both children and adults were shown to correlate with abilities such as syntactic processing (Kidd, 2012; Kidd & Arciuli, 2016; Misyak, Christiansen, & Tomblin, 2010), lexical knowledge and vocabulary size (Mainela-Arnold & Evans, 2014; Shafto, Conway, Field, & Houston, 2012; Singh, Steven Reznick, & Xuehua, 2012; Spencer, Kaschak, Jones, & Lonigan, 2014), speech perception (Conway et al., 2010; Conway, Karpicke, & Pisoni, 2007; Lany, Shoaib, Thompson, & Estes, 2018), and literacy acquisition in first language (Arciuli & Simpson, 2012; Tong, Leung, & Tong, 2019; Torkildsen, Arciuli, & Wie, 2019) as well as second language (Frost, Siegelman, Narkiss, & Afek, 2013; A. Yu et al., 2019; see Arciuli, 2018 for a review). More direct evidence comes from a handful of longitudinal studies. Such studies show that infants' SL performance predicts later development of linguistic abilities (Ellis, Gonzalez, & Deák, 2014; Shafto et al., 2012) and that visual sequence learning performance at Grade 5 predicts growth in children's reading skills from the fifth to the sixth grade (van der Kleij, Groen, Segers, & Verhoeven, 2019), although clearly more longitudinal research is still needed (see Arciuli & von Koss Torkildsen, 2012 for discussion). A related line of studies examines SL abilities of clinical populations with language deficits. Here also, the prediction is that a SL deficit may explain some of the difficulties these populations have with language processing and that they would therefore be characterized by impaired SL abilities. Special focus was directed to estimating SL abilities of individuals with dyslexia (Gabay, Thiessen, & Holt, 2015; Kahta & Schiff, 2019; Sigurdardottir et al., 2017), specific language impairment (Mainela-Arnold & Evans, 2014), and agrammatism (Christiansen, Louise Kelly, Shillcock, & Greenfield, 2010), all reporting some degree of impaired SL performance among these clinical samples (see Saffran, 2018 for review). Together, then, all of these individual-differences studies point to a clear pattern, supporting the involvement of SL in language.[3]

2 | LINKING SL AND LANGUAGE: MOVING BEYOND THE PROOF OF CONCEPT STAGE

In a more critical tone, however, this line of research only constitutes a preliminary step: What is common to the vast majority of these studies is that they observed *some* correlation between some SL task and some measure of linguistic performance. As such, studies of individual differences in SL so far generally provided a proof of concept regarding the link between SL and language: showing that they have something in common. Needless to say, proof-of-concept studies have an important theoretical role: The fact that it can now be said in a relatively high degree of certainty that SL is in fact related to linguistic performance, in some form or capacity, is far from trivial. Nevertheless, I argue that the next generation of SL research is still left with the task of moving beyond the proof of concept stage, in the quest of fully unveiling the role of SL across language (and cognition in general). In the second part of this review, I outline some of the challenges that SL research currently faces, which are key to achieving a deeper theoretical insight. While discussing these challenges I also review some recent studies in the field that already go beyond the proof of concept stage. Note that although I review these open issues from the perspective of individual differences, I believe that most of them apply to studies of SL in general.

2.1 | What is the structure of SL as a theoretical construct?

Most previous studies examining the link between SL and language-treated SL as a single "black-box": Assuming that all SL computations are subserved by a single central mechanism and therefore that all types of SL computations are equally important in all aspects of linguistic performance. Historically, this domain-generality assumption stems from studies demonstrating SL computations across a wide variety of stimuli and domains (reviewed above), leading to the assertion that if SL is observed across modalities and domains, it is likely that a single mechanism is responsible to all SL phenomena (see Frost, Armstrong, Siegelman, & Christiansen, 2015 for a review). In the context of individual differences, this assumption is reflected by the fact that the selection of an SL task for a particular study is typically not a matter of deep theoretical considerations—one simply selects a SL task to be employed in the study from the large arsenal of available SL tasks (visual and auditory, involving adjacent or non-adjacent contingencies, involving learning across time or in space, etc.) and uses individuals' performance in this task as a predictor of some linguistic outcome. This begs the question: Is SL indeed a domain-general, ubiquitous ability?

A large body of empirical evidence challenges this common view. First, multiple studies provide evidence for qualitative differences in SL across sensory modalities (i.e., across SL in visual, auditory, and tactile inputs). Such studies documented different learning biases in the visual, auditory and tactile modalities (Conway & Christiansen, 2005), different effects of presentation parameters in learning visual versus auditory material (Conway & Christiansen, 2009; Emberson, Conway, & Christiansen, 2011), and different developmental trajectories for visual and auditory SL (Raviv & Arnon, 2017). In addition to these findings that support modality-specific computations in SL, other studies further point to differences in the assimilation of different types of information, even within a modality. For example, while it is well documented that humans are sensitive both to the transitional statistics and to the aggregated frequency and variance of exemplars in the input, Thiessen, Kronstein, and Hufnagle (2013) claimed that from a computational perspective, the assimilation of these two types of regularities is likely subserved by different processes. Concurring behavioral evidence also showed patterns of informational-specificity (specifically, in learning adjacent vs. non-adjacent contingencies: e.g., Newport & Aslin, 2004, and verbal vs. nonverbal material: Shufaniya & Arnon, 2018; Siegelman, Bogaerts, Elazar, Arciuli, & Frost, 2018). Lastly, joining these group-level observations are preliminary individual-differences studies revealing patterns of modality and informational-specificity in SL. As such, studies of individual-differences reported little to no correlation between auditory and visual SL performance (Misyak & Christiansen, 2012; Siegelman & Frost, 2015; but see N. Siegelman, L. Bogaerts, & A. Elazar 2018 for a case where a higher correlation is present) and between learning adjacent and non-adjacent contingencies (Romberg & Saffran, 2013; Siegelman & Frost, 2015; see also Kalra et al., 2019 for a further demonstration of low correlations between different implicit learning tasks).

Critically, all of these studies provide converging evidence that SL cannot be simply referred to as a domain-general construct. In fact, multiple recent works explicitly take this theoretical stand, and call for a refinement of SL theory, which should be thought of as a componential, rather than a unified, theoretical construct (Arciuli, 2017; Arciuli & Conway, 2018; Daltrozzo & Conway, 2014; Frost et al., 2015; Siegelman, Bogaerts, Christiansen, & Frost, 2017; Thiessen et al., 2013). Importantly, while these studies differ in their exact theoretical approach, they all share one important principle—SL involves a set of non-overlapping underlying computations. Thus, they all call for a better understanding of the nature of domain-general versus domain-specific computations in SL, accompanied by explicit mapping of its sub-components. In that sense, it seems that SL follows the footsteps of the study of other major theoretical constructs in the cognitive sciences, which started from a conceptualization as unified constructs but were later reformulated as componential abilities, such as working memory (e.g., Baddeley, 1992; Smith, Jonides, Marshuetz, & Koeppe, 1998) and attention (e.g., Knudsen, 2007).

This refined view of SL carries critical implications for studies of individual differences in SL and their relation to language. First, if SL is not a unified construct, an outstanding challenge is the mapping of the structure of SL as a multifaceted, complex, theoretical construct. To do so, the relevant sub-components of SL should be specified, and their (non)overlapping computations should be explicitly discussed (see Arciuli, 2017; N. Siegelman, L. Bogaerts, & M. H. Christiansen 2017 for detailed discussions; and see Arciuli & Conway, 2018 for a detailed discussion with a focus on clinical populations). To this aim, studies of individual differences can be particularly useful, as they can directly examine the correlations between tasks that capture different SL aspects as a means of estimating the degree of overlap between SL facets (Erickson, Kaschak, Thiessen, & Berry, 2016; Siegelman & Frost, 2015).

Second, and importantly, the multifaceted view of SL requires a modification of common methodology as well as the theoretical analyses of studies of the link between individual differences in SL and language. Concretely, if SL is a componential ability rather than a unified construct, a single task cannot cover the full scope of SL computations. To date, SL research uses multiple different tasks that are taken to similarly measure "SL abilities": from the embedded triplet paradigm modeled after Saffran and colleagues, which focuses on learning of transitional statistics, to tasks borrowed from the implicit learning literature such as the serial reaction time task or artificial grammar learning. Because the latter set of tasks also involve the detection of embedded statistical regularities, they are generally considered to fall under the umbrella term "SL" (see Christiansen, 2018; Perruchet & Pacton, 2006 for a theoretical discussion of the relations between SL and implicit learning literatures). Importantly, however, not all tasks are identical from a componential perspective of SL—as each task involves different type(s) of statistical regularities to be extracted and thus implicates different computations.

Indeed, the low magnitude of observed correlations between SL tasks suggest that they tap (at least partially) into different sub-components of SL. Therefore, from an individual-differences perspective, it is critical to understand what are the exact components tapped by each task, and to what extent different tasks do or do not overlap. Then, when predicting a given linguistic skill, each study should carefully consider the specific components of SL that are relevant to the investigated linguistic ability and come up with predictions regarding which SL task(s) should be associated with this linguistic skill based on the overlap in computations (see Misyak & Christiansen, 2012; Misyak et al., 2010; Qi, Sanchez Araujo, Georgan, Gabrieli, & Arciuli, 2019 for preliminary directions and discussions). Thus, for example, auditory sentence processing may be related more to an auditory SL task than to a visual one, while reading may rely on both visual and auditory SL (see Hung, Frost, & Pugh, 2018; Qi et al., 2019). Similarly, learning different types of statistical information is predicted to underlie specific aspects of language that are predominantly characterized by such regularities (such as distributional SL in learning phonemic categories, Maye et al., 2002; or non-adjacent SL in certain aspects of syntax, e.g., Misyak et al., 2010).

Overall, this more refined view holds the promise of moving beyond proof of concept studies—which only shows that there is *something* common between SL and language—towards a mechanistic view, specifying the exact computations that are available to learners, and the exact role each of them plays in different linguistic processes.

2.2 | SL beyond the lab: The scope of SL phenomena

As noted above, the first studies on SL focused on group-level performance, with the aim of showing that humans are capable of extracting statistical properties from the sensory input. Since the goal of these studies was mainly to demonstrate the existence of SL abilities, early SL studies were not concerned with considerations of ecological validity (that is, the extent to which the simulated statistical environment as reflected in the in-lab SL task resembles the actual challenge learners face in the real world; see Erickson & Thiessen, 2015; Frost, Armstrong, & Christiansen, 2019; Pelucchi et al., 2009 for reviews). Hence, common SL tasks involve highly simplistic statistical structures, typically including only a very low number of embedded patterns (typically 4 to 16), which are all either fully predictable (TP = 1) or of one predictability level (e.g., TP = 0.5), all of similar length (typically of 3-elements), where regularities are presented one at a time, in a unimodal manner (i.e., either visual or auditory), and with short familiarizations that span from a few minutes to about an hour. Individual-differences studies that followed borrowed their tasks from these earlier investigations of SL, simply by changing the focus from the mean performance in test, to the number of correct responses of each participant. As a result, our understanding of individual differences in SL is almost entirely limited to how individuals differ from one another in their ability to extract simplistic statistical structures and how does success under these simplistic settings relate to other cognitive functions, such as language. Is that all there is to individual-differences in SL?

My claim here is that previous studies most likely only revealed the tip of the iceberg, focusing on individual differences only in a constrained part of SL phenomena, which is captured in learning from overly simplistic input. It is very likely however that participants would display non-overlapping and qualitatively different inter-individual variance when examining SL under broader, more realistic settings. Mapping this variance is key for understanding what SL is, who is a good statistical learner, and how do individual differences determine success in real-life tasks that include a myriad of complex statistical regularities such as language acquisition and use.

To do so, the next generation of SL research in general, and individual-differences in particular, should better simulate the SL computations that learners are required to perform in order to fully learn complex structures outside the lab. For starters, the vast majority of previous SL studies assumes that learners come to the learning task tabula rasa. Real-world learning, however, constantly involves the updating of existing knowledge (see, e.g., Karuza et al., 2016; Kóbor, Horváth, Kardos, Nemeth, & Janacsek, 2019; Siegelman, Bogaerts, Elazar, et al., 2018; Siegelman, Bogaerts, Kronenfeld, & Frost, 2018; Weiss, Gerfen, & Mitchel, 2009). Thus, mapping individual differences in SL cannot focus only on how individuals differ in their ability to learn a novel set of regularities from scratch, but should also consider variance in individuals' capacity to learn novel information given existing assimilated regularities. This can be done by in-task manipulations that simulate the learning of two consecutive structures, examining how regularities acquired during exposure to the first stream impacts learning of those in consecutive input (see Starling, Reeder, & Aslin, 2018 for preliminary findings), or alternatively, by quantifying real-world statistics using corpora and assess how they impact learning of an artificial stream that requires updating of this statistical knowledge (and see Fine & Florian Jaeger, 2013 for a related work in the context of adaptation to novel syntactic structures). Relatedly, learning in the real world commonly involves multiple regularities present simultaneously (in contrast to common SL tasks in which participants are exposed to a single isolated input stream). Indeed, recent group-level studies show that learners concurrently track more than one regularity (e.g., simultaneously learning of adjacent and non-adjacent contingencies, Deocampo, King, & Conway, 2019; Romberg & Saffran, 2013; Vuong, Meyer, & Christiansen, 2016). Do all individuals display similar sensitivity to concurrent statistical information, or do learners differ in their ability to track multiple regularities? Moreover, the presence of multiple regularities also requires learners to choose where to allocate their attention given the regularities across different parts of the input. Studies suggest that on average, learners prefer to attend to stimuli that are not too predictable or too random, but with a medium level of predictability (e.g., Kidd, Piantadosi, & Aslin, 2012, 2014). Do all individuals display similar attention allocation in light of competing concurrent information, or do individuals differ in how they guide their attention to stimuli given their predictability (which may then have cascading effects on their performance in learning different inputs)? Lastly, a growing number of group-level studies of SL focus on cross-modal integration, showing that reliable information present in one modality can facilitate learning in another (e.g., Glicksohn & Cohen, 2013; Mitchel & Weiss, 2011; Thiessen, 2010). Here, too, the question is whether there are systematic individual differences in the ability to use information from one modality to learn in another and whether it predicts related linguistic outcomes (e.g., the use of facial cues to learn about speech; Mitchel, Christiansen, & Weiss, 2014, and the use of tactile gestures in word learning; Seidl, Tincoff, Baker, & Cristia, 2015).

I wish to emphasize that taking into account the complexity of real-world learning tasks does not mean that researchers should draw an infinite number of distinctions between different types of SL computations. In fact, recent work by N. Siegelman, L. Bogaerts, A. Elazar, (2018) demonstrate the opposite—how considering real-world constraints may lead to the unveiling of similarities across seemingly non-overlapping SL processes. Concretely, this study found that when the involvement of prior knowledge is equated across modalities (i.e., when both modalities use nonverbal stimuli that do not implicate prior knowledge), visual and auditory SL performance are highly correlated. This finding suggests that previous reports of modality-specificity effects in SL may have been in part due to different involvement of prior expectations in visual and auditory SL paradigms (since the former typically uses abstract visual shapes, that do not implicate prior expectations, in contrast to the latter which typically uses spoken syllables). To emphasize, this example shows that taking into consideration constraints that affect SL computations in the wild may result in a better understanding of not only the dissimilarities between SL components, but also of commonalities between SL processes that were previously unobserved due to intervening factors.

2.3 | The strength of current evidence: A cause for concern?

The number of studies showing a correlation between SL abilities and linguistic outcomes is large (and is apparently constantly growing), presenting a seemingly strong body of evidence regarding the link between the two. Recently, however, some studies have raised concerns regarding the replicability of these correlations, as well as regarding their magnitude. As such, two recent studies failed to find any correlations between implicit and statistical learning measures and different outcomes of language and literacy in children and adults (Schmalz, Moll, Mulatti, & Schulte-Körne, 2019; West, Vadillo, Shanks, & Hulme, 2018). Other studies did find some correlations, but these were either of a small magnitude (Spencer et al., 2014) or limited to only specific tasks (Qi et al., 2019). Joining these studies are recent meta-analyses suggesting that the evidence in support of a SL deficit among populations with language impairments is currently inconclusive. Thus, while an earlier meta-analysis provided evidence for a deficit of individuals with specific language impairment in the serial reaction time task (Lum, Ullman, & Conti-Ramsden, 2013), more recent meta-analyses on SL in dyslexia, which examined also other tasks (artificial grammar learning and embedded pattern paradigm), suggest that the observed deficits may be limited to only some tasks and/or affected by publication bias (Schmalz, Altoè, & Mulatti, 2017; van Witteloostuijn, Boersma, Wijnen, & Rispens, 2017). Do these studies undermine the validity of earlier studies relating SL abilities to language?

Here, I argue that to fully understand the implications of these null (or partial) findings, we first need to answer a basic question: What determines the magnitude of a correlation between a given SL task and a given linguistic outcome? One critical factor is the task's reliability: Whether a task produces a measure that consistently ranks individuals (e.g., Nunnally & Bernstein, 1994), or, in simpler words—whether high performing individuals in some task are consistently classified as high performers, whereas low performers are consistently classified as such. In the context of individual differences, reliability is critical as it limits the predictive validity of the task—since, as noted already by Spearman (1904), reliability presents an upper bound to the correlation between a task and an outcome measure. A recent work showed that tasks developed for group-level investigations are often characterized by compromised reliability (Hedge, Powell, & Sumner, 2018). This is because tasks developed for group-level investigations strive to minimize between-subject variability (considered in group-level analysis as error variance), while individual-differences studies require measures that are sensitive enough to detect such variation. In the context of individual differences in SL, the question therefore is whether current SL tasks—generally originating from earlier group-level investigations in the field—produce reliable signal at the individual-level. Without reliable measurement of SL abilities, one cannot accurately assess the strength of the evidence in favor (or against) the association between SL and any other outcome.[4]

Indeed, the issue of the reliability of SL tasks has received increasingly more attention over the past few years. It started from studies showing how SL tasks developed for group-level investigations are generally characterized by low to mediocre reliability both in adults (Bogaerts, Siegelman, Ben-Porat, & Frost, 2017; Siegelman & Frost, 2015), and perhaps even more so in children (Arnon, 2019; West et al., 2018). As a result, the development of more reliable SL tasks is now recognized as a key methodological challenge, with recent works already presenting some promising advances. First, research shows that simply altering easy-to-control task parameters—such as the number of trials in the test, their task demands, and their difficulty level, can substantially improve SL task reliability in adults (Siegelman, Bogaerts, & Frost, 2017). In parallel, other studies aim to further improve SL measurement by adding a continuous (or online) assessment of SL, focusing on the familiarization phase as subjects actually extract the regularities from the input stream. Behavioral studies using online measures typically use a task examining RTs to the stimuli during the exposure phase, and assess the magnitude of speed up to predictable compared to unpredictable items (Batterink, 2017; Kuppuraj, Duta, Thompson, & Bishop, 2018; Siegelman, Bogaerts, & Frost, 2017). Such online measures present an alternative to the reliance on the commonly used offline measures of SL (such as 2AFC tests), which only tap into performance after learning has occurred, and thus may interfere with the representations actually learned during familiarization (Siegelman, Bogaerts, Christiansen, et al., 2017). Lastly, studies have also developed measures that are based on changes in related cognitive functions as a result of SL, examining the impact of statistical contingencies on short-term memory (Isbilen et al., 2017), or on attention as reflected in neural entrainment (Batterink & Paller, 2017). This line of studies not only provides potentially improved methods for assessing SL, but from a theoretical standpoint can also shed light on the cognitive processes impacted by and related to SL.

Together, these recent efforts suggest that SL measurement can indeed be vastly improved and calls for the widespread use of such improved measures in future research, particularly in studies of individual-differences. In parallel, however, still more methodological work is needed to fully estimate—and improve—the reliability of the wide arsenal of available SL tasks. This seems to be particularly true in studies with children, where recent reliability estimates present mixed findings. Thus, some studies report reasonable reliability estimates within a session (i.e., split-half reliability or internal consistency, Tong et al., 2019; Torkildsen et al., 2019; van der Kleij et al., 2019), but others reveal low test–retest reliability (Arnon, 2019; West et al., 2018), suggesting that current SL tasks still do not tap into a stable characteristic of a child (see also Conway, Arciuli, Lum, & Ullman, 2019; Krishnan & Watkins, 2019; West, Vadillo, Shanks, & Hulme, 2019 for commentaries discussing the impact of task reliability on outcomes of individual- and group-level studies). In general, more research is needed to fully understand how both task-related factors (e.g., number of trials, input modality, type of embedded regularities, and measurement domain) and sample-related properties (in particular—participants' age) influence task reliability, and how to maximize the reliability of a given task for a given population. To this aim, future studies can combine data from more than one form of measurement, for example by using composite offline–online behavioral metrics (Siegelman, Bogaerts, Kronenfeld, et al., 2018), or by combining behavioral measures with neural indices from electroencephalography recordings (e.g., Batterink & Paller, 2017; Vasuki, Sharma, Ibrahim, & Arciuli, 2017). In addition, note that when estimating the strength of the association between a SL task and a linguistic outcome, one should also consider the reliability of the linguistic task, as it similarly constrains the magnitude of the observed correlation between the two.

Importantly, however, the association between a given SL task and a given linguistic outcome is also determined by theoretical reasons that have to do with the joint computations that are involved in the predictor and the outcome. Discussions of the link between a SL task and a linguistic outcome should therefore not only consider the reliability of the predictor(s) and the outcome, but also the validity of the investigated SL component given the statistical computations that underlie the outcome at hand. Throughout this paper, I have reviewed evidence that SL is not a single, unified construct. Rather, SL is a multifaceted construct, shaped and constrained by the complexity of real-world statistics, and SL research has only recently begun to map all processes that are available to learners as they extract regularities from the input. Similarly, linguistic outcomes are not a unified construct, and the statistical computations that are involved in each sub-component of language vary, given the relevant linguistic information to be extracted (i.e., what are the relevant contingencies—whether they are adjacent or non-adjacent, in what modality are they presented—print, speech, or both, what prior knowledge they implicate, and how do they relate to concurrent and sequential regularities both within and across modalities). Moreover, these computations are shaped not only by the linguistic task, but also by the idiosyncrasies of the statistical structure of a given language (i.e., the relevant statistics to be learned in English are very different than that of Chinese; see Frost, 2012; Hung et al., 2018 for a discussion in the context of cross-linguistic differences in reading). A full understanding of individual differences in SL and their role in language requires then an explicit theory regarding the exact computations that are tapped by each SL task, and the available statistical information that needs to be extracted from the sensory input in order to achieve successful attainment of different linguistic skills in different languages. This perspective is compatible with individual-differences studies already showing that some linguistic outcomes are only predicted by some SL sub-component(s), but not others (Misyak & Christiansen, 2012; Qi et al., 2019; van der Kleij et al., 2019). Future studies should further refine the predictions regarding when and why one should expect positive relations between a SL measure and a linguistic outcome. Importantly, both positive and null results are equally informative in testing these predictions, demonstrating when a SL component and a linguistic outcome overlap, but also when they do not.

3 | CONCLUSIONS

For more than two decades, researchers have aimed to understand the role of SL across cognition in general and language in particular. The prism of individual differences has opened a novel way for investigating this issue, and indeed a growing body of evidence documents a portion of shared variance between individual variability in SL and different aspects of linguistic behavior. In parallel, however, recent advances have revealed the complexity of SL as a theoretical construct, highlighting various sub-components and processes that fall under the umbrella of SL computations. This calls for further refinements of the methodologies used to capture variability in SL abilities as well as the theoretical discussions regarding the exact computations that are used by learners as they assimilate different types of linguistic information, towards a better understanding of the role of SL in language.

ACKNOWLEDGEMENT

The author received funding from the Rothschild Yad-Hanadiv foundation and the National Institute of Child Health & Human Development (NICHD Grant: P20HD091013) as a postdoctoral fellow. I wish to thank Louisa Bogaerts and Alexis K. Black for helpful comments.

ORCID

Noam Siegelman https://orcid.org/0000-0002-5005-4838

ENDNOTES

[1] As noted above, early demonstrations of related abilities exist in the "implicit learning" literature, such as the early work of Esper (1925) and the seminal artificial grammar learning studies by Reber from the 1960s (e.g., Reber, 1967). Here, I focus on the more recent works in the field, typically labeled as "statistical learning" studies. I return to the issue of the relations between SL and implicit learning tasks when discussing the componential nature of SL, below.

[2] Note that the first study by Saffran and colleagues was conducted with infants, and hence, no explicit decisions were of course involved in the test: rather, it was based on a comparison of looking times at targets and foils (Saffran et al., 1996).

[3] The current focus on individual differences does not mean to undermine the importance of other lines of research in establishing the link between SL and linguistic abilities. Notably, a large body of evidence points to shared neuro-circuitry between SL and language, including neuro-imaging studies (e.g., Christiansen, Conway, & Onnis, 2012; Karuza et al., 2013) as well as neuropsychological studies with clinical populations (e.g., Christiansen et al., 2010). Another intriguing (albeit smaller) line of studies suggest that a SL training may produce improvements in linguistic skills (e.g., Hoen et al., 2003). Together with studies of individual differences which are the focus here, these works provide converging evidence for a link between SL and language. Nonetheless, I believe that the challenges described below pertain also to most of these studies.

[4] To be clear, this issue is not at all unique to SL research. Researchers have dealt with challenges related to task (un)reliability from the inception of experimental psychology (see e.g., classic work on the reliability of change/difference scores, e.g., Cronbach & Furby, 1970; Lord, 1958). These issues have recently regained the attention of the research community and discussed in detail with particular focus on how task reliability impacts the replicability of individual- and group-level study outcomes (e.g., Hedge et al., 2018; Lebel & Paunonen, 2011; Loken & Gelman, 2017).

REFERENCES

Arciuli, J. (2017). The multi-component nature of statistical learning. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *372*(1711), 20160058. https://doi.org/10.1098/rstb.2016.0058

Arciuli, J. (2018). Reading as statistical learning. *Language, Speech, and Hearing Services in Schools*, *49*(3), 634–643. https://doi.org/10.1044/2018_lshss-stlt1-17-0135

Arciuli, J., & Conway, C. M. (2018). The promise—and challenge—of statistical learning for elucidating atypical language development. *Current Directions in Psychological Science*, *27*(6), 492–500. https://doi.org/10.1177/0963721418779977

Arciuli, J., & Simpson, I. C. (2012). Statistical learning is related to reading ability in children and adults. *Cognitive Science*, *36*(2), 286–304. https://doi.org/10.1111/j.1551-6709.2011.01200.x

Arciuli, J., & von Koss Torkildsen, J. (2012). Advancing our understanding of the link between statistical learning and language acquisition: The need for longitudinal data. *Frontiers in Psychology*, *3*(324). https://doi.org/10.3389/fpsyg.2012.00324

Arnon, I. (2019). Do current statistical learning tasks capture stable individual differences in children? An investigation of task reliability across modality. *Behavior Research Methods.* https://doi.org/10.3758/s13428-019-01205-5

Aslin, R. N., Saffran, J. R., & Newport, E. L. (1998). Computation of conditional probability statistics by 8-month-old infants. *Psychological Science*, *9*(4), 321–324. https://doi.org/10.1111/1467-9280.00063

Baddeley, A. (1992). Working memory. *Science*, *255*(5044), 556–559. https://doi.org/10.1126/science.1736359

Barakat, B. K., Seitz, A. R., & Shams, L. (2013). The effect of statistical learning on internal stimulus representations: Predictable items are enhanced even when not predicted. *Cognition*, *129*(2), 205–211. https://doi.org/10.1016/j.cognition.2013.07.003

Batterink, L. J. (2017). Rapid statistical learning supporting word extraction from continuous speech. *Psychological Science*, *28*(7), 921–928. https://doi.org/10.3837/tiis.0000.00.000

Batterink, L. J., & Paller, K. A. (2017). Online neural monitoring of statistical learning. *Cortex*, *90*, 31–45. https://doi.org/10.1016/j.cortex.2017.02.004

Bogaerts, L., Siegelman, N., Ben-Porat, T., & Frost, R. (2017). Is the Hebb repetition task a reliable measure of individual differences in sequence learning? *The Quarterly Journal of Experimental Psychology*, *71*(4), 892–905. https://doi.org/10.1080/17470218.2017.1307432

Bogaerts, L., Siegelman, N., & Frost, R. (2016). Splitting the variance of statistical learning performance: A parametric investigation of exposure duration and transitional probabilities. *Psychonomic Bulletin & Review*, *23*(4), 1250–1256. https://doi.org/10.3758/s13423-015-0996-z

Bulf, H., Johnson, S. P., & Valenza, E. (2011). Visual statistical learning in the newborn infant. *Cognition*, *121*, 127–132. https://doi.org/10.1016/j.cognition.2011.06.010

Chetail, F. (2017). What do we do with what we learn? Statistical learning of orthographic regularities impacts written word processing. *Cognition*, *163*, 103–120. https://doi.org/10.1016/j.cognition.2017.02.015

Chomsky, N. (1959). A review of B.F. Skinner's verbal behavior. *Language*, *35*, 26–58. https://doi.org/citeulike-article-id:263746

Christiansen, M. H. (2018). Implicit statistical learning: A tale of two literatures. *Topics in Cognitive Science*, *11*(3), 468–481. https://doi.org/10.1111/tops.12332

Christiansen, M. H., Conway, C. M., & Onnis, L. (2012). Similar neural correlates for language and sequential learning: Evidence from event-related brain potentials. *Language and Cognitive Processes*, *27*(2), 231–256. https://doi.org/10.1080/01690965.2011.606666

Christiansen, M. H., Louise Kelly, M., Shillcock, R. C., & Greenfield, K. (2010). Impaired artificial grammar learning in agrammatism. *Cognition*, *116*, 382–393. https://doi.org/10.1016/j.cognition.2010.05.015

Conway, C. M., Arciuli, J., Lum, J. A. G., & Ullman, M. T. (2019). Seeing problems that may not exist: A reply to West et al.'s (2018) questioning of the procedural deficit hypothesis. *Developmental Science*, *22*(4), e12814. https://doi.org/10.1111/desc.12814

Conway, C. M., Bauernschmidt, A., Huang, S. S., & Pisoni, D. B. (2010). Implicit statistical learning in language processing: Word predictability is the key. *Cognition*, *114*(3), 356–371. https://doi.org/10.1016/j.cognition.2009.10.009

Conway, C. M., & Christiansen, M. H. (2005). Modality-constrained statistical learning of tactile, visual, and auditory sequences. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *31*(1), 24–39. https://doi.org/10.1037/0278-7393.31.1.24

Conway, C. M., & Christiansen, M. H. (2009). Seeing and hearing in space and time: Effects of modality and presentation rate on implicit statistical learning. *European Journal of Cognitive Psychology*, *21*(4), 561–580. https://doi.org/10.1080/09541440802097951

Conway, C. M., Karpicke, J., & Pisoni, D. B. (2007). Contribution of implicit sequence learning to spoken language processing: Some preliminary findings with hearing adults. *Journal of Deaf Studies and Deaf Education*, *12*(3), 317–334. https://doi.org/10.1093/deafed/enm019

Cronbach, L. J., & Furby, L. (1970). How we should measure "change": Or should we? *Psychological Bulletin*, *74*(1), 68–80. https://doi.org/10.1037/h0029382

Dale, R., Duran, N. D., & Morehead, J. R. (2012). Prediction during statistical learning, and implications for the implicit/explicit divide. *Advances in Cognitive Psychology*, *8*(2), 196–209. https://doi.org/10.2478/v10053-008-0115-z

Daltrozzo, J., & Conway, C. M. (2014). Neurocognitive mechanisms of statistical-sequential learning: What do event-related potentials tell us? *Frontiers in Human Neuroscience*, *8*(437). https://doi.org/10.3389/fnhum.2014.00437

Deocampo, J. A., King, T. Z., & Conway, C. M. (2019). Concurrent learning of adjacent and nonadjacent dependencies in visuo-spatial and visuo-verbal sequences. *Frontiers in Psychology*, *10*, 1107. https://doi.org/10.3389/fpsyg.2019.01107

Dotsch, R., Hassin, R. R., & Todorov, A. (2017). Statistical learning shapes face evaluation. *Nature Human Behaviour*, *1*(1), 1–6. https://doi.org/10.1038/s41562-016-0001

Ellis, E. M., Gonzalez, M. R., & Deák, G. O. (2014). Visual prediction in infancy: What is the association with later vocabulary? *Language Learning and Development*, *10*(1), 36–50. https://doi.org/10.1080/15475441.2013.799988

Emberson, L. L., Conway, C. M., & Christiansen, M. H. (2011). Timing is everything: Changes in presentation rate have opposite effects on auditory and visual implicit statistical learning. *Quarterly Journal of Experimental Psychology*, *64*, 1021–1040. https://doi.org/10.1080/17470218.2010.538972

Endress, A. D., & Mehler, J. (2009). The surprising power of statistical learning: When fragment knowledge leads to false memories of unheard words. *Journal of Memory and Language*, *60*, 351–367. https://doi.org/10.1016/j.jml.2008.10.003

Erickson, L. C., Kaschak, M. P., Thiessen, E. D., & Berry, C. A. S. (2016). Individual differences in statistical learning: Conceptual and measurement issues. *Collabra*, *2*(1), 14. https://doi.org/10.1525/collabra.41

Erickson, L. C., & Thiessen, E. D. (2015). Statistical learning of language: Theory, validity, and predictions of a statistical learning account of language acquisition. *Developmental Review*, *37*, 66–108.

Esper, E. A. (1925). A technique for the experiment investigation of associative interference in artificial linguistic material. *Language Monographs*, *1*, 1–47.

Evans, J., Saffran, J., & Robe-Torres, K. (2009). Statistical learning in children with specific language impairment. *Journal of of Speech, Language, and Hearing Research*, *52*, 321–335.

Fine, A. B., & Florian Jaeger, T. (2013). Evidence for implicit learning in syntactic comprehension. *Cognitive Science*, *37*(3), 578–591. https://doi.org/10.1111/cogs.12022

Frost, R. (2012). Towards a universal model of reading. *Behavioral and Brain Sciences*, *35*(5), 263–279. https://doi.org/10.1017/S0140525X11001841

Frost, R., Armstrong, B. C., & Christiansen, M. H. (2019). Statistical learning research: A critical review and possible new directions. *Psychological Bulletin*, *145*(12), 1128–1153. https://doi.org/10.1037/bul0000210

Frost, R., Armstrong, B. C., Siegelman, N., & Christiansen, M. H. (2015). Domain generality versus modality specificity: The paradox of statistical learning. *Trends in Cognitive Sciences*, *19*(3), 117–125. https://doi.org/10.1016/j.tics.2014.12.010

Frost, R., Siegelman, N., Narkiss, A., & Afek, L. (2013). What predicts successful literacy acquisition in a second language? *Psychological Science*, *24*(7), 1243–1252. https://doi.org/10.1177/0956797612472207

Gabay, Y., Thiessen, E. D., & Holt, L. L. (2015). Impaired statistical learning in developmental dyslexia. *Journal of Speech, Language, and Hearing Research*, *58*, 934–945. https://doi.org/10.1044/2015

Gebhart, A. L., Newport, E. L., & Aslin, R. N. (2009). Statistical learning of adjacent and nonadjacent dependencies among nonlinguistic sounds. *Psychonomic Bulletin & Review*, *16*, 486–490. https://doi.org/10.3758/PBR.16.3.486

Gervain, J., Nespor, M., Mazuka, R., Horie, R., & Mehler, J. (2008). Bootstrapping word order in prelexical infants: A Japanese-Italian cross-linguistic study. *Cognitive Psychology*, *57*(1), 56–74. https://doi.org/10.1016/j.cogpsych.2007.12.001

Glicksohn, A., & Cohen, A. (2013). The role of cross-modal associations in statistical learning. *Psychonomic Bulletin & Review*, *20*, 1161–1169. https://doi.org/10.3758/s13423-013-0458-4

Gómez, R. L. (2002). Variability and detection of invariant structure. *Psychological Science*, *13*, 431–436. https://doi.org/10.1111/1467-9280.00476

Hedge, C., Powell, G., & Sumner, P. (2018). The reliability paradox: Why robust cognitive tasks do not produce reliable individual differences. *Behavior Research Methods*, *50*(3), 1166–1186. https://doi.org/10.3758/s13428-017-0935-1

Hoen, M., Golembiowski, M., Guyot, E., Deprez, V., Caplan, D., & Dominey, P. F. (2003). Training with cognitive sequences improves syntactic comprehension in agrammatic aphasics. *Neuroreport*, *14*(3), 495–499. https://doi.org/10.1097/00001756-200303030-00040

Hung, Y.-H., Frost, S. J., & Pugh, K. R. (2018). Domain generality and specificity of statistical learning and its relation with reading ability. In T. Lachmann & T. Weis (Eds.), *Reading and dyslexia: From basic functions to higher order cognition* (pp. 33–55). Cham, Switzerland: Springer. https://doi.org/10.1007/978-3-319-90805-2_2

Hunt, R. H., & Aslin, R. N. (2001). Statistical learning in a serial reaction time task: Access to separable statistical cues by individual learners. *Journal of Experimental Psychology. General*, *130*, 658–680. https://doi.org/10.1037/0096-3445.130.4.658

Isbilen, E. S., McCauley, S. M., Kidd, E., & Christiansen, M. H. (2017). Testing statistical learning implicitly: A novel chunk-based measure of statistical learning. In G. Gunzelmann, A. Howes, T. Tenbrink, & E. Davelaar (Eds.), *Proceedings of the 39th annual conference of the cognitive science society* (pp. 564–569). Austin, TX: Cognitive Science Society.

James, W. (1890). *The principles of psychology* (Vol. 1). New York: Holt. https://doi.org/10.1037/10538-000

Johnson, E. K. (2012). Bootstrapping language: Are infant statisticians up to the job? In P. Rebuschat & J. Williams (Eds.), *Statistical learning and language acquisition* (pp. 55–89). Berlin, Germany: Mouton de Gruyter. https://doi.org/10.1515/9781934078242.55

Kahta, S., & Schiff, R. (2019). Deficits in statistical leaning of auditory sequences among adults with dyslexia. *Dyslexia*, *25*(2), 142–157. https://doi.org/10.1002/dys.1618

Kalra, P. B., Gabrieli, J. D. E., & Finn, A. S. (2019). Evidence of stable individual differences in implicit learning. *Cognition*, *190*, 199–211. https://doi.org/10.1016/j.cognition.2019.05.007

Karuza, E. A., Li, P., Weiss, D. J., Bulgarelli, F., Zinszer, B. D., & Aslin, R. N. (2016). Sampling over nonuniform distributions: A neural efficiency account of the primacy effect in statistical learning. *Journal of Cognitive Neuroscience*, *28*(10), 1484–1500. https://doi.org/10.1162/jocn_a_00990

Karuza, E. A., Newport, E. L., Aslin, R. N., Starling, S. J., Tivarus, M. E., & Bavelier, D. (2013). The neural correlates of statistical learning in a word segmentation task: An fMRI study. *Brain and Language*, *127*, 46–54. https://doi.org/10.1016/j.bandl.2012.11.007

Kaufman, S. B., DeYoung, C. G., Gray, J. R., Jiménez, L., Brown, J., & Mackintosh, N. (2010). Implicit learning as an ability. *Cognition*, *116*, 321–340. https://doi.org/10.1016/j.cognition.2010.05.011

Kidd, C., Piantadosi, S. T., & Aslin, R. N. (2012). The goldilocks effect: Human infants allocate attention to visual sequences that are neither too simple nor too complex. *PLoS ONE*, *7*(5), e36399. https://doi.org/10.1371/journal.pone.0036399

Kidd, C., Piantadosi, S. T., & Aslin, R. N. (2014). The goldilocks effect in infant auditory attention. *Child Development*, *85*, 1795–1804. https://doi.org/10.1111/cdev.12263

Kidd, E. (2012). Implicit statistical learning is directly associated with the acquisition of syntax. *Developmental Psychology*, *48*(1), 171–184. https://doi.org/10.1037/a0025405

Kidd, E., & Arciuli, J. (2016). Individual differences in statistical learning predict children's comprehension of syntax. *Child Development*, *87*(1), 184–193. https://doi.org/10.1111/cdev.12461

Kirkham, N. Z., Slemmer, J. A., & Johnson, S. P. (2002). Visual statistical learning in infancy: Evidence for a domain general learning mechanism. *Cognition*, *83*(2), B35–B42. https://doi.org/10.1016/S0010-0277(02)00004-5

van der Kleij, S. W., Groen, M. A., Segers, E., & Verhoeven, L. (2019). Sequential implicit learning ability predicts growth in reading skills in typical readers and children with dyslexia. *Scientific Studies of Reading*, *23*(1), 77–88. https://doi.org/10.1080/10888438.2018.1491582

Knudsen, E. I. (2007). Fundamental components of attention. *Annual Review of Neuroscience*, *30*, 57–78. https://doi.org/10.1146/annurev.neuro.30.051606.094256

Kóbor, A., Horváth, K., Kardos, Z., Nemeth, D., & Janacsek, K. (2019). Perceiving structure in unstructured stimuli: Implicitly acquired prior knowledge impacts the processing of unpredictable transitional probabilities. *BioRxiv*, 738419. https://doi.org/10.1101/738419

Krishnan, S., & Watkins, K. E. (2019). A challenge for the procedural deficit hypothesis: How should we measure sequential learning in childhood? *Developmental Science*, *22*(4), e12815. https://doi.org/10.1111/desc.12815

Kuppuraj, S., Duta, M., Thompson, P., & Bishop, D. (2018). Online incidental statistical learning of audiovisual word sequences in adults: A registered report. *Royal Society Open Science*, *5*(2), 171678. https://doi.org/10.1098/rsos.171678

Lany, J., Shoaib, A., Thompson, A., & Estes, K. G. (2018). Infant statistical-learning ability is related to real-time language processing. *Journal of Child Language*, *45*(2), 368–391. https://doi.org/10.1017/S0305000917000253

Lebel, E. P., & Paunonen, S. V. (2011). Sexy but often unreliable: The impact of unreliability on the replicability of experimental findings with implicit measures. *Personality and Social Psychology Bulletin*, *37*(4), 570–583. https://doi.org/10.1177/0146167211400619

Lidz, J., & Gagliardi, A. (2015). How nature meets nurture: Universal grammar and statistical learning. *Annual Review of Linguistics*, *1*(1), 333–353. https://doi.org/10.1146/annurev-linguist-030514-125236

Loken, E., & Gelman, A. (2017). Measurement error and the replication crisis. *Science*, *355*(6325), 584–585. https://doi.org/10.1126/science.aal3618

Lord, F. M. (1958). Further problems in the measurement of growth. *Educational and Psychological Measurement*, *18*(3), 437–451. https://doi.org/10.1177/001316445801800301

Lum, J. A. G., Ullman, M. T., & Conti-Ramsden, G. (2013). Procedural learning is impaired in dyslexia: Evidence from a meta-analysis of serial reaction time studies. *Research in Developmental Disabilities*, *34*(10), 3460–3476. https://doi.org/10.1016/j.ridd.2013.07.017

Mainela-Arnold, E., & Evans, J. L. (2014). Do statistical segmentation abilities predict lexical-phonological and lexical-semantic abilities in children with and without SLI? *Journal of Child Language*, *41*(2), 327–351. https://doi.org/10.1017/S0305000912000736

Maye, J., Werker, J. F., & Gerken, L. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition*, *82*(3), 101–111. https://doi.org/10.1016/S0010-0277(01)00157-3

Miller, G. A. (1958). Free recall of redundant strings of letters. *Journal of Experimental Psychology*, *56*, 485–491. https://doi.org/10.1037/h0044933

Misyak, J. B., & Christiansen, M. H. (2012). Statistical learning and language: An individual differences study. *Language Learning*, *62*, 302–331. https://doi.org/10.1111/j.1467-9922.2010.00626.x

Misyak, J. B., Christiansen, M. H., & Tomblin, J. B. (2010). On-line individual differences in statistical learning predict language processing. *Frontiers in Psychology*, *1*, 31. https://doi.org/10.3389/fpsyg.2010.00031

Mitchel, A. D., Christiansen, M. H., & Weiss, D. J. (2014). Multimodal integration in statistical learning: Evidence from the McGurk illusion. *Frontiers in Psychology*, *5*, 407. https://doi.org/10.3389/fpsyg.2014.00407

Mitchel, A. D., & Weiss, D. J. (2011). Learning across senses: Cross-modal effects in multisensory statistical learning. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *37*, 1081–1091. https://doi.org/10.1037/a0023700

Newport, E. L., & Aslin, R. N. (2004). Learning at a distance I. statistical learning of non-adjacent dependencies. *Cognitive Psychology*, *48*, 127–162. https://doi.org/10.1016/S0010-0285(03)00128-2

Nunnally, J., & Bernstein, I. (1994). *Psychometric theory*, 3rd edn. McGraw-Hill, New York (3). Retrieved from http://scholar.google.com.vn/scholar?q=Psychometric+Theory+3rd&btnG=&hl=vi&as_sdt=0,5#0

Onishi, K. H., Chambers, K. E., & Fisher, C. (2002). Learning phonotactic constraints from brief auditory experience. *Cognition*, *83*(1), B13–B23. https://doi.org/10.1016/S0010-0277(01)00165-2

Pacton, S., Fayol, M., & Perruchet, P. (2005). Children's implicit learning of graphotactic and morphological regularities. *Child Development*, *76*(2), 324–339. https://doi.org/10.1111/j.1467-8624.2005.00848_a.x

Pelucchi, B., Hay, J. F., & Saffran, J. R. (2009). Statistical learning in a natural language by 8-month-old infants. *Child Development*, *80*(3), 674–685. https://doi.org/10.1111/j.1467-8624.2009.01290.x

Perruchet, P., & Pacton, S. (2006). Implicit learning and statistical learning: One phenomenon, two approaches. *Trends in Cognitive Sciences*, *10*, 233–238. https://doi.org/10.1016/j.tics.2006.03.006

Qi, Z., Sanchez Araujo, Y., Georgan, W. C., Gabrieli, J. D. E., & Arciuli, J. (2019). Hearing matters more than seeing: A cross-modality study of statistical learning and reading ability. *Scientific Studies of Reading*, *23*(1), 101–115. https://doi.org/10.1080/10888438.2018.1485680

Raviv, L., & Arnon, I. (2017). The developmental trajectory of children's auditory and visual statistical learning abilities: Modality-based differences in the effect of age. *Developmental Science*, *21*(4), e12593. https://doi.org/10.1111/desc.12593

Reber, A. S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior*, *6*(6), 855–863. https://doi.org/10.1016/S0022-5371(67)80149-X

Romberg, A. R., & Saffran, J. R. (2013). All together now: Concurrent learning of multiple structures in an artificial language. *Cognitive Science*, *37*, 1290–1320. https://doi.org/10.1111/cogs.12050

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*(5294), 1926–1928. https://doi.org/10.1126/science.274.5294.1926

Saffran, J. R. (2003). Statistical language learning: Mechanisms and constraints. *Current Directions in Psychological Science*, *12*(4), 110–114. https://doi.org/10.1111/1467-8721.01243

Saffran, J. R. (2018). Statistical learning as a window into developmental disabilities. *Journal of Neurodevelopmental Disorders*, *10*, 35. https://doi.org/10.1186/s11689-018-9252-y

Saffran, J. R., Newport, E. L., & Aslin, R. N. (1996). Word segmentation: The role of distributional cues. *Journal of Memory and Language*, *35*(4), 606–621. https://doi.org/10.1006/jmla.1996.0032

Saffran, J. R., & Wilson, D. P. (2003). From syllables to syntax: Multilevel statistical learning by 12-month-old infants. *Infancy*, *4*(2), 273–284. https://doi.org/10.1207/S15327078IN0402_07

Salimpoor, V. N., Zald, D. H., Zatorre, R. J., Dagher, A., & McIntosh, A. R. (2015). Predictions and the brain: How musical sounds become rewarding. *Trends in Cognitive Sciences*, *19*(2), 86–91. https://doi.org/10.1016/j.tics.2014.12.001

Schmalz, X., Altoè, G., & Mulatti, C. (2017). Statistical learning and dyslexia: A systematic review. *Annals of Dyslexia*, *67*(2), 147–162. https://doi.org/10.1007/s11881-016-0136-0

Schmalz, X., Moll, K., Mulatti, C., & Schulte-Körne, G. (2019). Is statistical learning ability related to reading ability, and if so, why? *Scientific Studies of Reading*, *23*(1), 64–76. https://doi.org/10.1080/10888438.2018.1482304

Seidenberg, M. S., MacDonald, M. C., & Saffran, J. R. (2002). Neuroscience: Does grammar start where statistics stop? *Science*, *298*(5593), 553–554. https://doi.org/10.1126/science.1078094

Seidl, A., Tincoff, R., Baker, C., & Cristia, A. (2015). Why the body comes first: Effects of experimenter touch on infants' word finding. *Developmental Science*, *18*(1), 155–164. https://doi.org/10.1111/desc.12182

Shafto, C. L., Conway, C. M., Field, S. L., & Houston, D. M. (2012). Visual sequence learning in infancy: Domain-general and domain-specific associations with language. *Infancy*, *17*, 247–271. https://doi.org/10.1111/j.1532-7078.2011.00085.x

Shufaniya, A., & Arnon, I. (2018). Statistical learning is not age-invariant during childhood: Performance improves with age across modality. *Cognitive Science*, *42*(8), 3100–3115. https://doi.org/10.1111/cogs.12692

Siegelman, N., Bogaerts, L., Christiansen, M. H., & Frost, R. (2017). Towards a theory of individual differences in statistical learning. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *372*(1711), 20160059. https://doi.org/10.1098/rstb.2016.0059

Siegelman, N., Bogaerts, L., Elazar, A., Arciuli, J., & Frost, R. (2018). Linguistic entrenchment: Prior knowledge impacts statistical learning performance. *Cognition*, *177*, 198–213. https://doi.org/10.1016/j.cognition.2018.04.011

Siegelman, N., Bogaerts, L., & Frost, R. (2017). Measuring individual differences in statistical learning: Current pitfalls and possible solutions. *Behavior Research Methods*, *49*(2), 418–432. https://doi.org/10.3758/s13428-016-0719-z

Siegelman, N., Bogaerts, L., Kronenfeld, O., & Frost, R. (2018). Redefining "learning" in statistical learning: What does an online measure reveal about the assimilation of visual regularities? *Cognitive Science*, *42*(53), 692–727. https://doi.org/10.1111/cogs.12556

Siegelman, N., & Frost, R. (2015). Statistical learning as an individual ability: Theoretical perspectives and empirical evidence. *Journal of Memory and Language*, *81*, 105–120.

Sigurdardottir, H. M., Danielsdottir, H. B., Gudmundsdottir, M., Hjartarson, K. H., Thorarinsdottir, E. A., & Kristjánsson, Á. (2017). Problems with visual statistical learning in developmental dyslexia. *Scientific Reports*, *7*(1), 606. https://doi.org/10.1038/s41598-017-00554-5

Singh, L., Steven Reznick, J., & Xuehua, L. (2012). Infant word segmentation and childhood vocabulary development: A longitudinal analysis. *Developmental Science*, *15*, 482–495. https://doi.org/10.1111/j.1467-7687.2012.01141.x

Smith, E. E., Jonides, J., Marshuetz, C., & Koeppe, R. A. (1998). Components of verbal working memory: Evidence from neuroimaging. *Proceedings of the National Academy of Sciences of the United States of America*, *95*(3), 876–882. https://doi.org/10.1073/pnas.95.3.876

Spearman, C. (1904). The proof and measurement of association between two things. *The American Journal of Psychology*, *15*(1), 72–101. https://doi.org/10.2307/1422689

Spencer, M., Kaschak, M. P., Jones, J. L., & Lonigan, C. J. (2014). Statistical learning is related to early literacy-related skills. *Reading and Writing*, *28*(4), 467–490. https://doi.org/10.1007/s11145-014-9533-0

Starling, S. J., Reeder, P. A., & Aslin, R. N. (2018). Probability learning in an uncertain world: How children adjust to changing contingencies. *Cognitive Development*, *48*, 105–116. https://doi.org/10.1016/j.cogdev.2018.06.008

Thiessen, E. D. (2010). Effects of visual information on adults' and infants' auditory statistical learning. *Cognitive Science*, *34*(6), 1093–1106. https://doi.org/10.1111/j.1551-6709.2010.01118.x

Thiessen, E. D., Kronstein, A. T., & Hufnagle, D. G. (2013). The extraction and integration framework: A two-process account of statistical learning. *Psychological Bulletin*, *139*, 792–814. https://doi.org/10.1037/a0030801

Thompson, S. P., & Newport, E. L. (2007). Statistical learning of syntax: The role of transitional probability. *Language Learning and Development*, *3*(1), 1–42. https://doi.org/10.1207/s15473341lld0301_1

Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.

Tong, X., Leung, W. W. S., & Tong, X. (2019). Visual statistical learning and orthographic awareness in Chinese children with and without developmental dyslexia. *Research in Developmental Disabilities*, *92*, 103443. https://doi.org/10.1016/j.ridd.2019.103443

Torkildsen, J. v. K., Arciuli, J., & Wie, O. B. (2019). Individual differences in statistical learning predict children's reading ability in a semi-transparent orthography. *Learning and Individual Differences*, *69*, 60–68. https://doi.org/10.1016/J.LINDIF.2018.11.003

Treiman, R., & Kessler, B. (2006). Spelling as statistical learning: Using consonantal context to spell vowels. *Journal of Educational Psychology*, *98*(3), 642–652. https://doi.org/10.1037/0022-0663.98.3.642

Vasuki, P. R. M., Sharma, M., Ibrahim, R. K., & Arciuli, J. (2017). Musicians' online performance during auditory and visual statistical learning tasks. *Frontiers in Human Neuroscience*, *11*, 114. https://doi.org/10.3389/fnhum.2017.00114

Vuong, L. C., Meyer, A. S., & Christiansen, M. H. (2016). Concurrent statistical learning of adjacent and non-adjacent dependencies. *Language Learning*, *66*(1), 8–30. https://doi.org/10.1111/lang.12137

Weiss, D. J., Gerfen, C., & Mitchel, A. D. (2009). Speech segmentation in a simulated bilingual environment: A challenge for statistical learning? *Language Learning and Development: The Official Journal of the Society for Language Development*, *5*(1), 30–49. https://doi.org/10.1080/15475440802340101

West, G., Vadillo, M. A., Shanks, D. R., & Hulme, C. (2018). The procedural learning deficit hypothesis of language learning disorders: We see some problems. *Developmental Science*, *21*(2), e12552. https://doi.org/10.1111/desc.12552

West, G., Vadillo, M. A., Shanks, D. R., & Hulme, C. (2019). The procedural deficit hypothesis of language learning disorders: We still see some serious problems. *Developmental Science*, *22*(4), e12813. https://doi.org/10.1111/desc.12813

van Witteloostuijn, M., Boersma, P., Wijnen, F., & Rispens, J. (2017). Visual artificial grammar learning in dyslexia: A meta-analysis. *Research in Developmental Disabilities*, *70*, 126–137. https://doi.org/10.1016/j.ridd.2017.09.006

Yang, C. D. (2004). Universal grammar, statistics or both? *Trends in Cognitive Sciences*, *8*, 451–456. https://doi.org/10.1016/j.tics.2004.08.006

Yu, A., Chen, M. S. Y., Cherodath, S., Hung, D. L., Tzeng, O. J. L., & Wu, D. H. (2019). Neuroimaging evidence for sensitivity to orthography-to-phonology conversion in native readers and foreign learners of Chinese. *Journal of Neurolinguistics*, *50*, 53–70. https://doi.org/10.1016/j.jneuroling.2018.07.002

Yu, C., & Smith, L. B. (2007). Rapid word learning under uncertainty via cross-situational statistics: Research article. *Psychological Science*, *18*(5), 414–420. https://doi.org/10.1111/j.1467-9280.2007.01915.x

AUTHOR BIOGRAPHY

Noam Siegelman's research is concerned with statistical learning, reading, and their intersection. His work focuses mainly on individual- and cross-linguistic differences, examining how individuals differ in their learning abilities, how writing systems vary in their statistical properties, and how does the interaction between individuals' learning capacities and the structure of their writing system co-determine literacy skills of a given individual in a given language. Noam holds a PhD in Cognitive Sciences from the Hebrew University of Jerusalem and is currently a post-doctoral fellow at Haskins laboratories.

How to cite this article: Siegelman N. Statistical learning abilities and their relation to language. *Lang Linguist Compass*. 2019;e12365. https://doi.org/10.1111/lnc3.12365