Measuring Statistical Learning Across Modalities and Domains in School-Aged Children Via an Online Platform and Neuroimaging Techniques

Julie M. Schneider[1], **Anqi Hu**[1], **Jennifer Legault**[1], **Zhenghan Qi**[1]

[1] Department of Linguistics and Cognitive Science, University of Delaware

Corresponding Authors

Julie M. Schneider
juschnei@udel.edu

Zhenghan Qi
zqi@udel.edu

Citation

Schneider, J.M., Hu, A., Legault, J., Qi, Z. Measuring Statistical Learning Across Modalities and Domains in School-Aged Children Via an Online Platform and Neuroimaging Techniques. *J. Vis. Exp.* (160), e61474, doi:10.3791/61474 (2020).

Date Published

June 30, 2020

DOI

10.3791/61474

URL

jove.com/video/61474

Abstract

Statistical learning, a fundamental skill to extract regularities in the environment, is often considered a core supporting mechanism of the first language development. While many studies of statistical learning are conducted within a single domain or modality, recent evidence suggests that this skill may differ based on the context in which the stimuli are presented. In addition, few studies investigate learning as it unfolds in real-time, rather focusing on the outcome of learning. In this protocol, we describe an approach for identifying the cognitive and neural basis of statistical learning, within an individual, across domains (linguistic vs. non-linguistic) and sensory modalities (visual and auditory). The tasks are designed to cast as little cognitive demand as possible on participants, making it ideal for young school-aged children and special populations. The web-based nature of the behavioral tasks offers a unique opportunity for us to reach more representative populations nationwide, to estimate effect sizes with greater precision, and to contribute to open and reproducible research. The neural measures provided by the functional magnetic resonance imaging (fMRI) task can inform researchers about the neural mechanisms engaged during statistical learning, and how these may differ across individuals on the basis of domain or modality. Finally, both tasks allow for the measurement of real-time learning, as changes in reaction time to a target stimulus is tracked across the exposure period. The main limitation of using this protocol relates to the hour-long duration of the experiment. Children might need to complete all four statistical learning tasks in multiple sittings. Therefore, the web-based platform is designed with this limitation in mind so that tasks may be disseminated individually. This methodology will allow users to investigate how the process of statistical learning unfolds across and within domains and modalities in children from different developmental backgrounds.

Introduction

Statistical learning is an elementary skill supporting the acquisition of rule-governed combinations in language inputs[1]. Successful statistical learning ability in infants predicts later language learning success[2,3]. Variability in statistical learning skills in school-aged children has also been associated with vocabulary[4] and reading[5,6]. Difficulty in statistical learning has been proposed as one etiological mechanism underlying language impairment[7]. Despite the association between statistical learning and language outcomes in both neurotypical and atypical populations, the cognitive and the neural mechanisms underlying statistical learning remain poorly understood. In addition, previous literature has revealed that, within an individual, statistical learning ability is not uniform but independent across domains and modalities[6,8,9]. The developmental trajectory of statistical learning abilities may further vary across domains and modalities[10]. These findings emphasize the importance of assessing individual differences in statistical learning across multiple tasks throughout the course of development. However, the field first requires a more systematic investigation of the relationship between statistical learning and first language development. To address these questions, we apply innovative methods including a web-based testing platform[11] that reaches a large number of children, and laboratory-based neuroimaging techniques (functional magnetic resonance imaging, or fMRI) that examine the real-time encoding of statistical information.

Standard measures of statistical learning begin with a familiarization phase and are followed by a two-alternative forced choice (2-AFC) task[12,13]. The familiarization phase introduces a continuous stream of stimuli embedded with statistical regularities, where some stimuli are more likely to co-occur than others. The presentation of these co-occurring stimuli follows a fixed temporal order. Participants are passively exposed to the stream during the familiarization phase, followed by a 2-AFC task that tests whether the participant successfully extracted the patterns. The 2-AFC accuracy task presents two consecutive sequences: one sequence has been presented to the participant during the familiarization phase, while the other is a novel sequence, or contains part of the sequence. Above-chance accuracy on the 2-AFC would indicate successful learning at the group level. Traditional behavioral tasks assessing statistical learning generally rely upon accuracy as the outcome measure of learning. However, accuracy fails to account for the natural learning of information as it unfolds in time. A measure of real-time learning is necessary to tap into the implicit learning process of statistical learning during which children are still encoding the regularities from the inputs[14,15,16]. Various adaptations across paradigms have been developed in an effort to move away from the 2-AFC measure, towards measures of on-line learning through behavioral responses during the exposure[16]. Studies utilizing these adaptations which measure reaction time during the exposure phase found they were related to post-learning accuracy[17] with better test-retest reliability compared to that of the accuracy in adult learners[18].

Neural measures are also foundational to our understanding of how learning unfolds over time, as the implicit process by which language learning occurs likely recruits different neural resources from those used once language is learned[19]. Neural measures also provide insights into differences in cognitive specializations underlying language ability across special populations[20]. How the condition contrast is designed in an fMRI study is crucial for how we interpret patterns of neural activation during learning. One common practice is to compare brain responses during the familiarization phase between sequences containing regular patterns versus those containing the same stimuli which are ordered randomly. However, previous research implementing such a random control condition found no evidence for learning in behavior, despite neural differences between structured and random sequences. This might be due to the interference of random sequences on learning of structured sequences, as both were constructed from the same stimuli[21,22]. Other fMRI studies which utilized backward speech or earlier learning blocks as the control condition confirmed learning took place behaviorally[19,23]. However, each of these paradigms introduced its own confounding factor, such as the effect of language processing for the former case and the effect of the experimental order for the latter case. Our paradigm uses the random sequence as the control condition but mitigates their interference on participants' learning of the structured sequences. Our fMRI paradigm also implements a mixed block/event-related design, which allows for the simultaneous modeling of transient trial-related and sustained task-related BOLD signals[24]. Lastly, and more broadly, neural measures allow for the measurement of learning in populations where eliciting an explicit behavioral response may be difficult (e.g., developmental and special populations)[25].

The current protocol adopts a response time measure, in addition to traditional accuracy measures, and examines brain activation during the familiarization phase. The combination of these methods aims to provide a rich dataset for the investigation of real-time learning processes. The web-based platform offers a set of learning measures by including both response time during the exposure phase and accuracy of the 2-AFC task during the test phase. The neuroimaging protocol allows for the investigation of the underlying neural mechanisms supporting statistical learning across domains and modalities. While it is optimal to measure statistical learning within an individual using both the web-based and fMRI protocols, the tasks are designed so that they may be disseminated independently, and therefore, as two independent measures of statistical learning. The fMRI experiments included in the current protocol can help clarify how stimulus encoding, pattern extraction, and other constituent components of statistical learning are represented by particular brain regions and networks.

Protocol

All participants gave written consent to participate and study was conducted in accordance with the Institutional Review Board.

1. Overview of the statistical learning paradigm utilized in the web-based protocol

1. Include four tasks in the current paradigm: image (visual-nonlinguistic), letter (visual-linguistic), tone (auditory-nonlinguistic), and syllable (auditory-linguistic).
    1. Construct stimuli for visual tasks using 12 standalone alien cartoon images (image) and 12 letter images (letter; B, J, K, A, H, C, F, E, J, G, D, M) showing the same alien holding up 12 signs with capital letters written on them.
    2. Construct auditory stimuli using 12 English syllables (syllable; pi,pu,pa,ti,tu.ta,di,du,da,bi,bu,ba) and 12 musical tones within the same octave (tone; F,G,D,G#,C#,B,C,F#,D#,E,A,A#). The syllable stimuli can be made using an artificial speech synthesizer, and can be recorded as separate files in Praat[26,27].

2. In the familiarization phase, present stimuli in a structured stream (see **Figure 1**). Feedback is not provided at any point during the familiarization or test phase.

   NOTE: Within each task, a familiarization phase is immediately followed by a test phase.

   1. For the Image (visual-nonlinguistic) task, structure 12 images into four target triplets. In the familiarization phase, repeat each of the four target triplets 24 times for a total of 96 triplets.

      NOTE: The 96 triplets are randomly concatenated into a continuous stream, with the constraint that no triplet can be immediately repeated. Images are presented one at a time in the center of the screen. Each image is presented for 800 ms with 200 ms of inter-stimulus interval. The whole familiarization phase will last for 4 min 48 s.

   2. Ensure the test phase always follows the familiarization phase and is composed of 32 two-alternative forced-choice (2AFC) questions. For each question, include 2 options: a target triplet from the familiarization phase and a triplet that was not included in the familiarization phase, referred to as a foil triplet.

      NOTE: Foil triplets are constructed so that the relative position of each image in the foil triplet is the same as the target triplet. Each target and foil triplet are presented 8 times total in a test, and each foil-target pair is repeated. The test phase consists of 32 (4 target triplets x 4 foil triplets x 2 repetitions) randomly ordered trials.

   3. For the letter (visual-linguistic) task include 12 images of capitalized letters that are organized into four target triplets (GJA, FKC, LBE, and MDH). For the test phase, create 4 foil triplets (GDE, FJH, LKA, and MBC) and pair them with the target triplets to form the 32 2AFC test trials. No letter triplet can contain any words, common acronyms, or initialisms.

   4. For the tone (auditory-nonlinguistic) task include 12 musical pure tones within the same octave (a full chromatic scale starting from middle C) and concatenate them into four target triplets (F#DE, ABC, C#A#F, and GD#G#). Unlike in the visual tasks, presentation speed is faster due to differences in auditory perceptual preference[6,28,29].

      NOTE: Each of the four target triplets is repeated 48 times for a total of 192 triplets (twice as much as the visual conditions). All triplets are concatenated into a sound stream with no triplet being repeated twice in a row. Pure tones are presented one at a time while participants view a blank screen. The duration of each tone is 460 ms with a 20 ms inter-stimulus interval. The whole stream lasts about 4 min and 36 s. As in the visual tasks, a test phase of 32 2AFC trials with pairs of target and foil triplets (F#BF, AA#G#, C#D#E, GDC) immediately follows the familiarization phase.

   5. For the syllable (auditory-linguistic) task use 12 consonant vowel (CV) syllables created and grouped into four target triplets (pa-bi-ku, go-la-tu, da-ro-pi, and ti-bu-do). The duration of each syllable and the inter-stimulus interval is the same as the tone condition. Pair four foil triplets (pa-ro-do, go-bu-ku, da-bi-tu, and ti-la-pi) with the target triplets in the test phase.

3. Randomize the order of the four statistical learning tasks across participants.

2. Participant recruitment

NOTE: While the web-based protocol and the fMRI protocol are best implemented together within a single participant, here we outline the best practices for participant recruitment for each task independently.

1. Web based participant recruitment
    1. Recruit participants age 6 years and above. Participants of any sex, race, and ethnicity may participate; however, study sample should be representative of the population.
    2. Recruit participants who are a native English speaker and have been exposed to no languages besides English before the age of 5.
    3. Ensure they report no known psychological (including ADD, depression, PTSD, and clinical anxiety) and/or neurological condition (including stroke, seizure, brain tumor, or closed head injury).
    4. Ensure participants have normal or corrected-to-normal vision (glasses or contacts are okay), normal color vision and normal hearing (no hearing aid or cochlear implant devices).
2. Task based fMRI participant recruitment
    1. Recruit participants age 6 years and above. Participants of any sex, race, and ethnicity may participate; however, study sample should be representative of the population.
    2. To be eligible, recruit participants who are native English speaker and have never been exposed to any languages besides English before the age of 5.
    3. Recruit right-handed individuals, with no known psychological (including ADD, depression, PTSD, and clinical anxiety) and neurological condition (including stroke, seizure, brain tumor, or closed head injury).
    4. Exclude participants who are pregnant, claustrophobic, taking psychoactive drugs, or have any metal in the body (including pacemakers, neural implants, metal plates or joints, shrapnel, and surgical staples).
    5. Ensure participants have normal or corrected-to-normal vision (glasses or contacts are okay), normal color vision and normal hearing (no hearing aid or cochlear implant devices).
    6. Determine the eligibility to participate in the MRI by having participants (or parents if the participant is a minor) complete an MRI Safety Screening Form.

3. Web based protocol

NOTE: The web-based statistical learning paradigm is hosted on a secure website (https://www.cogscigame.co[11]) and developed using jsPsych, a JavaScript library for creating behavioral experiments online[30] .

1. To reproduce tasks, go to DOI: 10.5281/zenodo.3820620. All scripts and materials are publicly available. Researchers can modify the scripts and run the experiments locally on any web browser as long as all the paths for the output files are set up appropriately.
2. Have participants complete a cover task where they are told to press a button when they see a certain target during all familiarization phases of each statistical learning task.
3. Target stimulus assignment for each task
    1. In the image, letter and syllable tasks, randomly choose one of the four triplets and assign the target

to the third stimulus of the triplet. In the tone task, constrain the target stimulus to only the lowest or the highest tones of the third stimulus in the triplets, and assign the target to the third stimulus of the triplet. This is done because tone stimuli are relatively harder to discriminate than other types of stimuli.

2. In the syllable and tone tasks, introduce participants to an alien and the favorite word/note in its alien language/folk music. Tell participants that they will listen to the alien's language/music and to remember to press the spacebar whenever they hear the favorite word/note" by deleting the space between / and note.

3. In the image task, tell participants to keep track of a special alien as a group of aliens line up to enter a spaceship. In the letter task, tell participants to keep track of the alien's favorite sign as the alien holds up signs for a parade. Give participant's a practice trial in both the image and letter tasks.

4. Do not provide explicit instructions about the presence of triplets.

5. Measure response time over the 24 trials in the visual tasks and over the 48 trials in the auditory tasks to assess online learning.

6. During the test phase, both a target (included in familiarization phase) and foil triplet (not included in familiarization phase) are presented to the participant. Instruct participants to then choose which one of the two is more similar to what they saw or heard in the familiarization phase. Each trial must end with a response.

4. Behavioral measures of statistical learning in the web-based protocol

1. Measure the real time learning during the familiarization phase via the linear slope of reaction time (change in reaction time throughout the familiarization phase).

2. To be considered a valid response to the target, check that the keypress must be in the time window of one stimulus before and one stimulus after the target stimulus. That is -480 ms to +960 ms relative to the onset of the target in the auditory tasks and -1000 ms to +2000 ms in the visual tasks. A keypress prior to the target is considered as anticipation and thus yields a negative reaction time.

3. To compare reaction times across conditions, transform the reaction times of each participant for each task into z scores. This normalizes the reaction times of an individual so that scores across tasks can be compared.

4. Calculate a reaction time slope of each participant for each condition using linear regression. Input the z-normed reaction times as the dependent variable and the target trial order as the independent variable (visual: 1 to 24; auditory: 1 to 48). The slope of the linear regression line (Beta coefficient/Estimate) is the reaction time slope (RT slope).

5. Measure offline accuracy of each participant for each condition by dividing the number of correct trials from the test phase by the total number of trials (32 trials).

4. Task based fMRI protocol

1. Modifications to the statistical learning paradigm (**Figure 2**).

1. For each task, present both a structured sequence (containing statistical regularities) and a random sequence (no statistical regularities).
NOTE: Structured sequences are identical to those described for the web-based protocol (see **Figure 1**). In contrast, random sequences contain the same 12 stimuli as presented in the structured sequences but are ordered pseudo-randomly. No combinations of any three stimuli are repeated more than once.
2. Divide each sequence into six smaller blocks of equal length (24 stimuli for the visual tasks and 48 stimuli for the auditory tasks).
3. Concatenate three structured blocks, 3 random blocks, and 6 resting blocks (silence with a blank screen) in a pseudorandom order to create four runs of auditory stimuli and four runs of visual stimuli. To maximize learning of the structured sequences, ensure that the random blocks in each run contain a different domain from the structured sequence (e.g., syllable structured sequences are presented together with tone random sequences in one run, and syllable random sequences are presented together with tone structured sequences in another run).
4. Include 288 images to be presented in each run for the visual task lasting approximately 4.77 min. Include 576 sounds to be presented in the auditory task which lasts approximately 4.42 min. At the beginning of each block, present a cue about the target with a verbal and visual probe: “Now listen/look for the [TARGET]”.
5. Among the four runs of the visual task, ensure that two contain structured sequences of images and the other two contain structured sequences of letters. Among the four runs of the auditory task, ensure that two contain structured sequences of syllables and the other two contain structured sequences of tones.

2. fMRI statistical learning procedure
    1. To help make participants, especially children, comfortable in the scanner, practice the MRI scanning session first using a mock scanner[31]. A mock scanner provides a naturalistic experience similar to the actual scanning session but is typically situated in a more child-friendly environment.
    2. First introduce the child to the mock scanner, i.e., brain camera, and ensure they are comfortable before putting them in the scanner.
    3. Introduce them to their “scan-buddy” and explain that the purpose of the scan buddy is to keep them accompanied and help them if they need anything. The scan buddy will gently remind the participant to keep still if too much motion is detected by the “camera”.
    4. Once they are in the scanner, play child-friendly videos to help them acclimate to the sound and video. When they are ready, play a few pre-recorded scanner sound clips to prepare them for the noises produced by the real MRI. During this time have them practice staying still and working with the scan buddy.
    5. Introduce children to the statistical learning paradigm and have them practice outside of the scanner. This is done by having children complete a brief portion of the task on a computer, similar to the web-based protocol by performing steps 3.2.2 and 3.2.3 mentioned above.
    NOTE: The practice stimuli are the same as those utilized in the task; however, children are only exposed to the random sequence and not the structured sequences, allowing for brief habituation

to the stimuli and task demands without enabling learning of particular sequences.

6. Ensure the fMRI data collection protocol is appropriately set up on the MRI acquisition computer.
NOTE: The acquisition parameters follow the recommendations of the Adolescent Brain Cognitive Development (ABCD) Study[32] .

7. Begin the scanning session with high resolution T1-weighted scans. Acquire these using a 176-slice 3D MPRAGE (Magnetization Prepared Rapid Gradient Echo) volume scan with TR (Repetition Time) = 2500 ms, TE (Echo Time) = 2.9 ms, flip angle = 8˚, FOV (Field of View) = 25.6 cm, 256 X 256 matrix size, and 1 mm slice thickness. This acquisition will last 7.2 min.

8. To acquire functional data, use T2*-weighted echo-planar imaging with simultaneous multi-slice scans acquisition with TR= 800 ms, TE = 32 ms, flip angle = 61˚, FOV = 21 cm, and matrix = 64 x 64. In this experiment, 60 adjacent slices are acquired in an interleaved sequence with 2.5 mm slice thickness, a 21 cm FOV, and a 64 X 64 matrix, resulting in an in-plane resolution of 2.5 mm x 2.5 mm x 2.5 mm.

9. Have participants lie comfortably on the bed of the fMRI scanner with headphones that protect their ears from the scanner noise and a response pad/button box in their hand (both headphones and button box must be scanner compatible).

10. Place additional padding around their head to ensure limited head motion during data collection. Give the button response box to the participant ahead of time to record responses and counterbalance whether the left or right hand is used to press buttons across participants.

11. Give every child an option of a scan buddy. For older, neurotypical children who are comfortable without a scan buddy, give them a squeeze ball to notify the experimenter if they are distressed or need to stop. Give younger children and special populations a squeeze ball but also provide them with a scan buddy to assist them (described in 4.2.3).

12. Place the head coil over the participant's head and align the patient's position in the bed.

13. On the acquisition computer register a new participant. Enter their participant ID, date of birth, weight and height. The participant may now be inserted into the bore of the MRI.

14. Acquire T1-weighted scan while showing participants a movie.

15. Before beginning the statistical learning paradigm, give participants the instructions of each task by speaking to them through an intercom system connected to their headphones.

16. In the auditory tasks, tell participants: "Now we're going to play a button-pressing game. You will hear the aliens say words and play music. Remember to press the button in your LEFT/RIGHT hand whenever you hear the sound you are listening for. There will be 4 parts, and each part will last about 5 min."

17. In the visual tasks, tell participants: "Now you are going to see the pictures of the aliens and the letters. Whenever you see the picture you are looking for, press the button in your LEFT/RIGHT hand. You will play this 4 times in a row. It will take about 5 minutes each time."

18. Start the statistical learning paradigm on the presentation computer and acquire the task fMRI data.
19. Once the participant has completed the paradigm, stop the MRI, safely remove them from inside the scanner, and remove the head coil.
20. After data collection, transfer all MRI data from the acquisition computer to a secured server for further analyses.

3. fMRI data analyses

1. Analyze in-scanner reaction time during the fMRI task similarly to the web-based calculation of reaction time during the familiarization phase. Normalize reaction time to compare across conditions, and calculate a linear slope using the normalized reaction time for each condition of an individual.
2. When analyzing the fMRI data, first organize and convert data to Brain Imaging Data Structure[33] (BIDS) formatting using HeuDiConv[34] (https://github.com/nipy/heudiconv).
3. Preprocess these data using fMRIPrep[35,36]. This automated preprocessing pipeline combines methodology from AFNI[37], ANTs[38], Freesurfer[39], FSL[40], and Mindboggle[41] to provide scientifically rigorous and reproducible data for use in data analysis.

   NOTE: The current study implements a mixed block/event-related design. The representative results (below) treat each mini block as an event (e.g., random sequence is an event, structured sequence is an event, etc.). However, the task is also designed so that one can model each stimulus as an event.
4. Include two task regressors for each run ("image" and "letter" for the visual condition, and "syllable" and "tone" for the auditory condition) in the first-level model design. Determine task regressors by convolving a vector of event onset times with their durations with a canonical hemodynamic response function. Compute differences and means between runs within each subject for higher-level model designs. This will result in a contrast between structured and random sequences within each type of stimuli.
5. Create a group mean of activation for structured blocks compared to random blocks within each modality/domain.

Representative Results

Web-based Behavioral Results

Given the current protocol is designed for easy dissemination with developmental populations, we have included preliminary web-based results based on data from 22 developing school-aged children (Mean (M) age = 9.3 years, Standard Deviation (SD) age = 2.04 years, range = 6.2-12.6 years, 13 girls). In the web-based statistical learning task, children performed significantly better than 0.5 chance-level on all conditions, indicating successful statistical learning at the group level (see **Table 1** for statistics; **Figure 3**). Mean reaction time slope was negative and significantly below 0 in the syllable condition ($M = -0.01$, $SD = 0.02$, $t(14) = -2.36$, one-tailed $p = .02$) and marginally significant in the letter condition ($M = -0.02$, $SD = 0.06$, $t(15) = -1.52$, one-tailed $p = .07$, **Figure 4**), suggesting a faster acceleration of target detection during the familiarization phase in the linguistic tasks. Mean reaction time slope was not significantly different from zero in the image condition ($M = 0.02$, SD =

0.04, $t(17) = 1.54$, one-tailed $p > .1$) or the tone condition ($M = 0.005$, $SD = 0.02$, $t(15) = -5.7 \times 10^{-17}$, one-tailed $p > .1$), despite evidence of learning in the offline measures of accuracy. Cronbach's alpha was 0.75 for the Letter task, 0.09 for the Syllable task, 0.67 for the Tone task, and 0.86 for the Image task. Correlations between implicit measures (RT slope) and explicit measures (accuracy) of statistical learning identify a significant relationship for the Image task ($R = -.48$, $p = 0.04$) and Letter task ($R = -.54$, $p = 0.03$). Inter-task correlations further suggest that the four tasks may have a modest degree of overlapping learning mechanism (**Figure 5**). While accuracy on both visual tasks was highly correlated ($R = .60$, $p = 0.02$), they were also positively associated with accuracy on the Syllable task (Image $R = .66$, $p = 0.01$; Letter $R = .85$, $p < 0.001$).

fMRI Results

Preliminary fMRI results were based on data from nine developing school-aged children. These nine children were a subset of the 22 children included in the web-based behavioral results, as not all children came to the lab to complete the fMRI portion of the study. All nine completed the auditory statistical learning tasks (*M* age = 10.77 years, *SD* = 1.96 years, range = 7.7-13.8 years, 4 girls) and seven completed the visual statistical learning tasks (*M* age = 11.41 years, *SD* = 2.37 years, range = 7.7-13.8 years, 4 girls). When comparing structured blocks to random blocks, significant clusters were observed in all four conditions (**Figure 6**). In the syllable condition, greater activation was found at the left superior temporal gyrus, right insula/frontal operculum, and anterior cingulate gyrus. In the tone condition, greater activation was found at left middle temporal gyri, bilateral angular gyri, left frontal pole, right lateral occipital cortex, right insula, and right frontal operculum. In the letter condition, greater activation was found at the left planum temporal. In the image condition, greater activation was found at the right lateral occipital cortex. These preliminary findings suggest that children's neural activation patterns differ across learning of statistical regularities depending on the modality and domain of the presented stimuli. The current task design is sensitive to these differences and can identify task-specific regions of activation similar to past studies[20, 25].

fMRI Behavioral Results

To demonstrate learning in the fMRI portion of this study, we have included in-scanner behavioral results from 28 adults (M age = 20.8, SD = 3.53, 20 females), as the data from 9 children was not enough to compute reliable statistics. Our findings in adults indicate that learning successfully occurred in all tasks for the structured sequence, supported by significantly quicker response time in the structured as compared to the random condition, except in the case of the tone task (see **Table 2** for statistics).

Taken together, our web-based measures of accuracy, and increased activation for structured versus random sequences in the scanner, indicate this protocol may be implemented with developmental populations to gauge statistical learning across domains and modalities within an individual. Our behavioral MRI results in an adult population further emphasis the utility of this protocol in measuring learning of structured sequences as it unfolds in real-time, as well as the ability to implement the web-based and fMRI protocols independently.

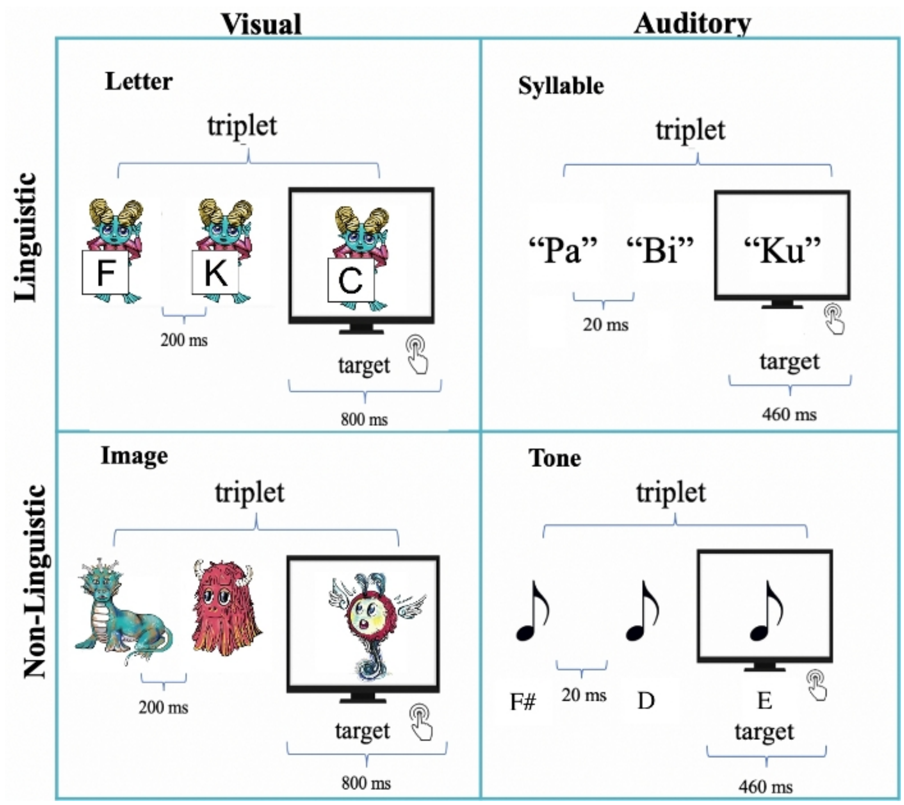

Figure 1: Familiarization phase of all four statistical learning tasks. Example triplets across each task are depicted in this figure. Each visual stimulus appeared for 800 ms with a 200 ms ISI, and each auditory stimulus was heard for 460 ms with a 20 ms ISI. Please click here to view a larger version of this figure.

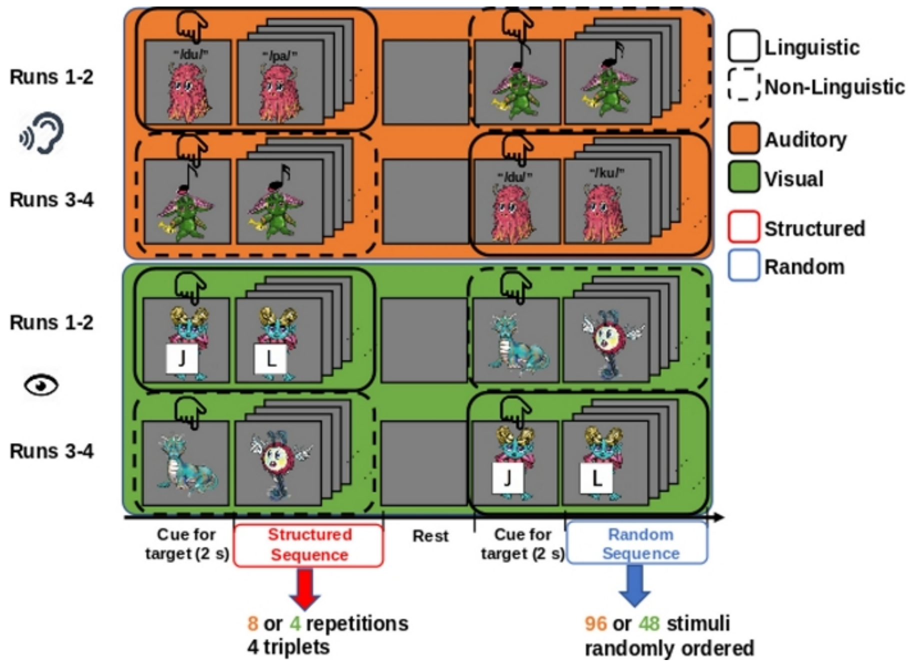

Figure 2: Familiarization modification for fMRI statistical learning tasks. The fMRI task was similar to the web-based familiarization phase but introduced a random sequence that was counterbalanced across domains. Please click here to view a larger version of this figure.

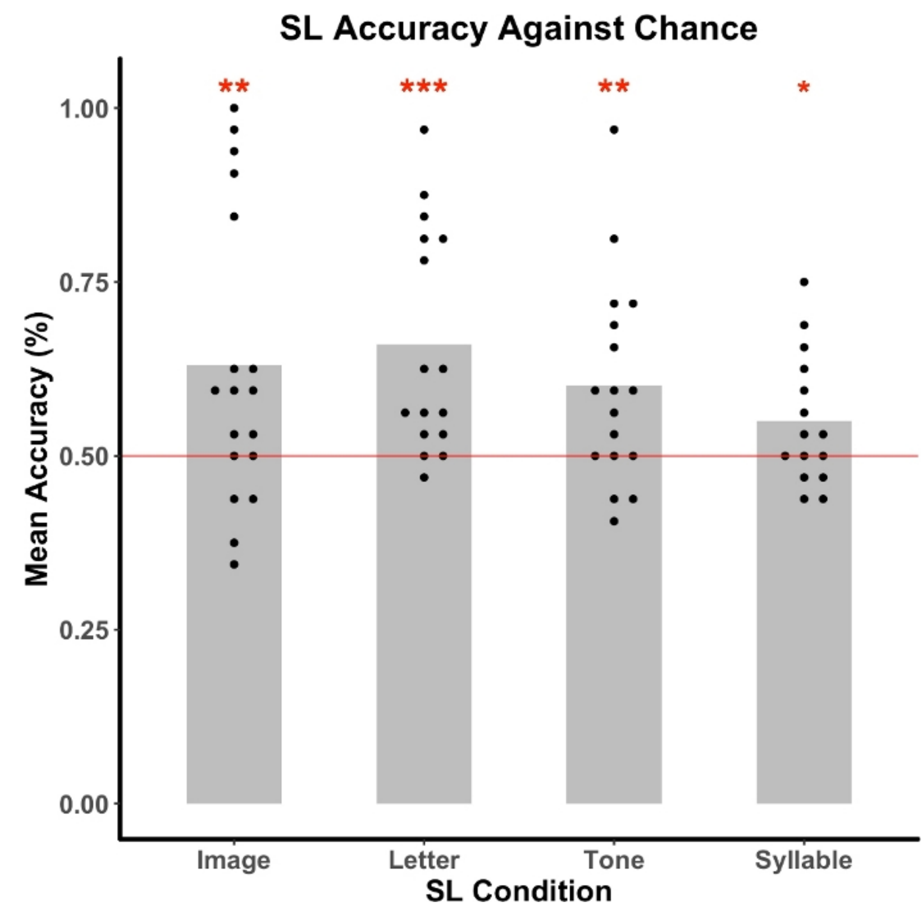

Figure 3: Average statistical learning (SL) accuracy in the web-based task compared against chance-level. Results indicate individuals performed significantly above chance on all four tasks, ***one-tailed $p < .001$, ** < 0.01, * < 0.05. Please click here to view a larger version of this figure.

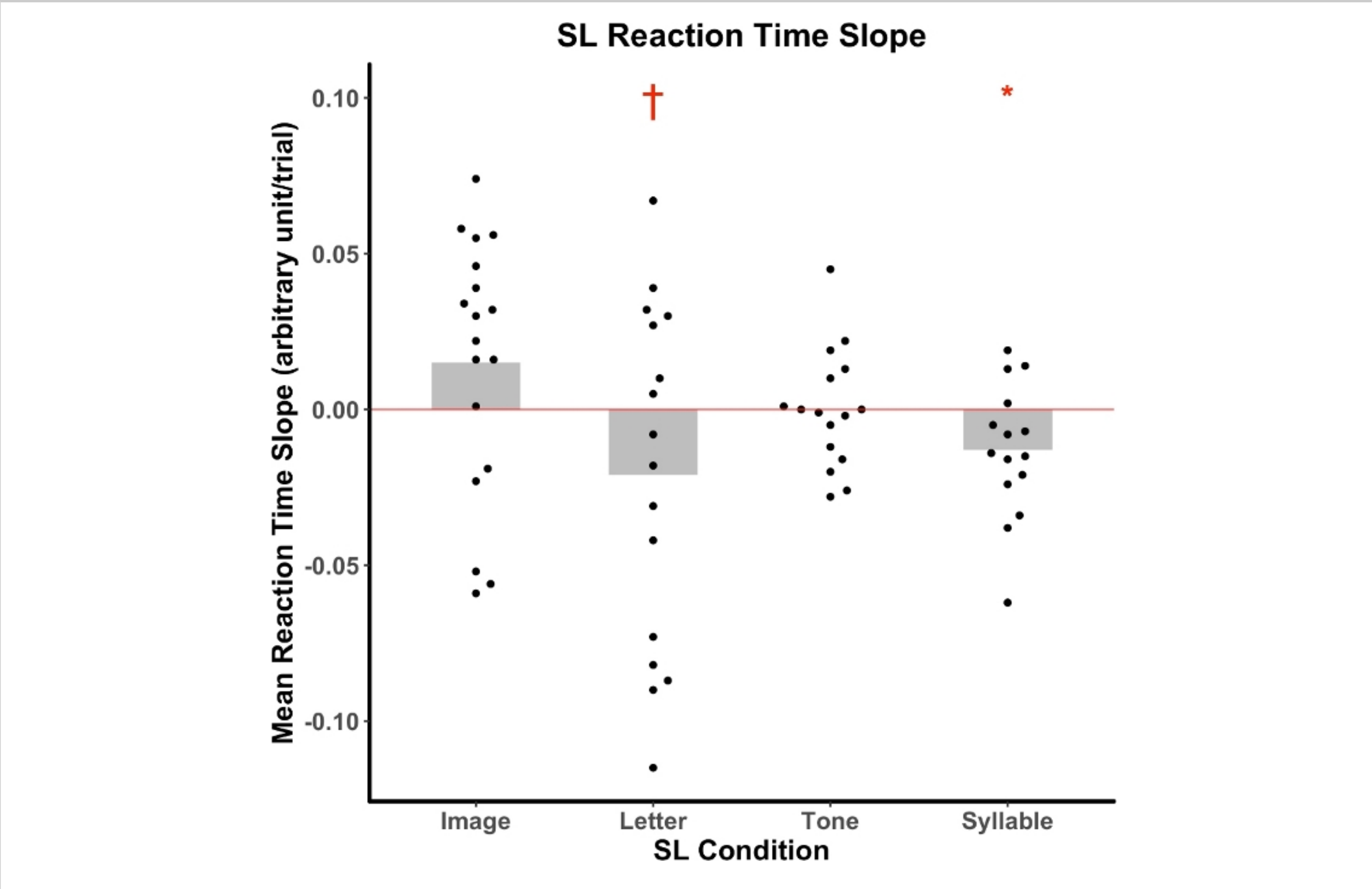

Figure 4: Mean reaction time slope in the web-based task against zero. A more negative slope indicates faster acceleration in the target detection during familiarization. Target detection significantly improved over the course of exposure during the syllable task. †one-tailed p = .07, * < .05. Please click here to view a larger version of this figure.

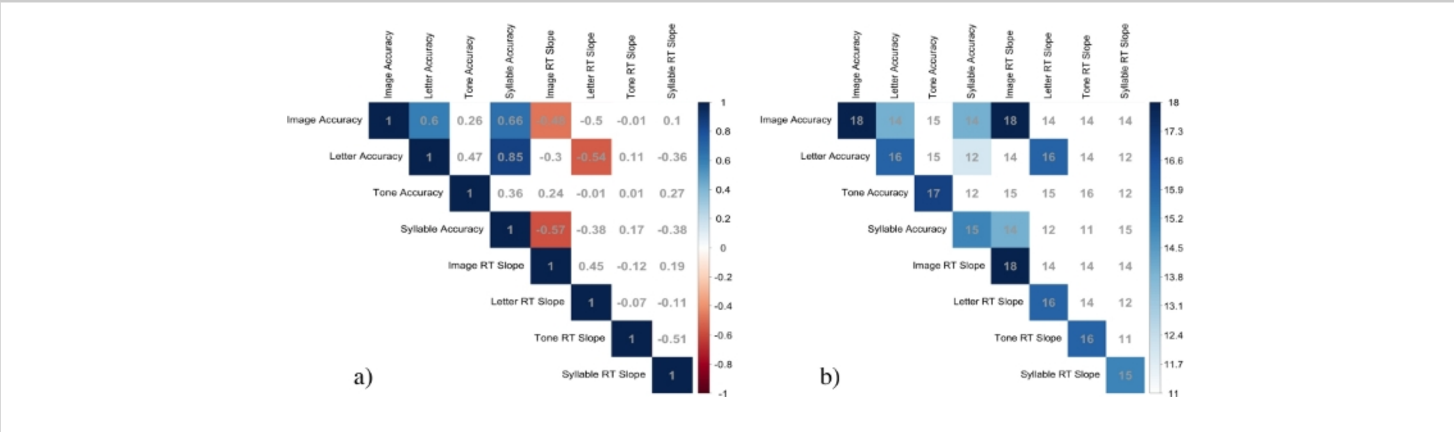

Figure 5: Web-based between-task correlations across all four statistical learning tasks. (**a**) Non-significant values at an alpha of .05 are shown with a white background. All comparisons with a colored background denote significant effects. (**b**) Sample size for each pairwise comparison. Please click here to view a larger version of this figure.

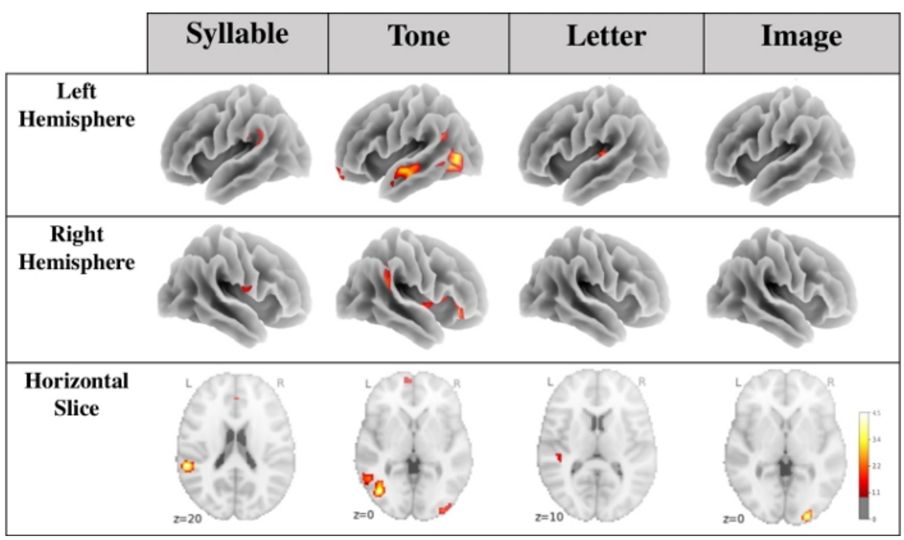

Figure 6: Neural activation at the group-level for structured blocks compared to random blocks within each modality and domain. Significant clusters were thresholded at voxel-level p < 0.001 and cluster-level p < 0.05 for each task. Horizontal slices were selected to depict the cluster with the maximum z-value. The color bar in the bottom, right corner reflects the same scale for all plots. Please click here to view a larger version of this figure.

Condition	Mean	Standard Deviation	One-tailed T-test
Image	0.63	0.21	$t(17) = 2.64$, $p = .009$
Letter	0.66	0.16	$t(15) = 3.98$, $p < .001$
Tone	0.60	0.15	$t(16) = 2.83$, $p = .006$
Syllable	0.55	0.1	$t(14) = 2.06$, $p = .03$

Table 1: Web-based accuracy by condition. One-sample t-tests represent group differences compared to 0.5 chance-level.

	Structured		Random		
Condition	**Mean**	**Standard Deviation**	**Mean**	**Standard Deviation**	**Paired Samples T-test**
Image	468.1	76.04	493.4	60.33	$t(27) = -2.01$, $p = .05$
Letter	374.72	143.59	502.1	68.75	$t(27) = -4.97$, $p <.001$
Tone	426.37	169.10	407.68	162.63	$t(26) = 0.67$, $p = .51$*
Syllable	589.3	180.95	679.9	55.99	$t(26) = -2.51$, $p = .02$*
*One subject had too few button presses to compute a value for the tone or syllable task.					

Table 2: MRI behavioral performance differences on random versus structured sequences across all four tasks in adults. Paired-samples t-tests represent group differences in learning of structured versus random sequences.

Discussion

The methods presented in the current protocol provide a multimodal paradigm for understanding the behavioral and neural indices of statistical learning across the course of development. The current design allows for the identification of individual differences in statistical learning ability across modalities and domains, which can be used for future investigation of the relationship between statistical learning and language development. Since an individuals' statistical learning ability is found to vary across domains and modalities[6,8,9], it is optimal if participants complete all four tasks. Findings from typically developing children and adults indicate that an individuals' performance across statistical learning domains/modalities can differentially relate to vocabulary[4] and reading[5,6] outcomes. Therefore, we recommend additional measures of cognitive and language abilities be taken to relate to the measures of statistical learning taken in the current protocol.

Research has reported reasonable internal consistency and test-retest reliability of these statistical learning tasks for adults[8,42]. However, concerns about task reliability for children[42] and a recent discussion on general measurement issues[9] indicates an urgent need to develop measures of statistical learning, that take into account children's developmental characteristics. While our previous research, as well as the preliminary data from the current protocol, indicates high internal consistency for the non-linguistic statistical learning tasks in school-aged children between 8 and 16 years old[6], our research also confirmed a less satisfying task reliability, particularly in auditory linguistic statistical learning which has been reported before[42]. The differences in internal consistency between tasks are particularly intriguing in light of recent findings on the impact of a learner's prior linguistic experiences on statistical learning outcomes[18,43,44]. Language and reading development change rapidly during the school years. The learnability of each auditory linguistic triplet might differ substantially within each child, depending on their developmental stage and current language abilities. Combining our protocol with other individual difference measures will offer an exciting opportunity to study the cascading effect between existing skills and subsequent learning underlying the heterogeneity of statistical learning performance across the course of development.

An important benefit of the current design is in its' utility for measuring statistical learning via an online web-platform. Researchers should be aware of the following when considering the accuracy of reaction time measurements via a web browser. de Leeuw and Motz (2016)[45] found the response times measured via a web browser were approximately 25 ms longer than those measured via other standard data presentation software. Importantly, this delay was found to be constant across trials. Because our measure of real-time learning in the web-based tasks is the slope of change in reaction time, the effects of the delay in reaction time has been minimized using within-subject comparisons. de Leeuw (2015)[30] has also acknowledged that reaction time measured via jsPsych may be affected by factors such as the processing speed of the computer or the number of tasks loaded in the background. To minimize these effects, we recommend normalizing response time within each individual participant before computing the response time slope[30].

The current protocol, providing robust methods to demonstrate large variability in learning behavior across domains and modalities, is designed to investigate individual differences of statistical learning. However, this protocol is not suitable for investigating questions such as whether visual statistical learning is inherently easier than auditory statistical learning. The interpretation of group-level performance differences between tasks is difficult due to all the confounding factors that we are not able to control, such as stimuli familiarity[14,43,46,47], sensory salience, and processing speed[28]. Related to stimuli familiarity, it is well established that an individual's prior experiences with the stimuli may influence their statistical learning performance. Additionally, the visual and auditory tasks are difficult to directly compare due to differences in the salience of the stimuli and presentation rate across these modalities. Therefore, our methods are designed with the aim of investigating individual differences in statistical learning. However, with advanced fMRI analysis approaches, our protocol is suitable for studying theoretical questions about the nature of statistical learning, for example we can ask which brain networks are sensitive to regularities in each domain and how the patterns of neural engagement differ/ overlap.

The current protocol was developed to be child-friendly and easily accessible to maximize research in neurotypical and atypical populations. During the implementation of this protocol with young children or those with developmental disorders, a critical step is to give breaks between each SL task to avoid fatigue. Each condition of the web-based tasks can be disseminated individually to ease cognitive demands. Prior to scanning, the mock scanner can be used to reduce child anxiety and head motion in preparation for the real fMRI task. An additional issue researcher should be aware of relates to a general concern when conducting any neuroimaging study: motion. A rotational head movement of just 0.3 mm can cause artifacts to manifest. In an effort to minimize the likelihood of motion artifacts, the current protocol has limited each run to last less than 5 minutes[48]. Participants should be encouraged to stay still during each 5-minute run but allowed to move or stretch between runs in order to reduce motion during actual scanning. We also recommend rigorous data analysis techniques to correct motion-related artifacts on the fMRI data[49].

Given the critical contribution of statistical learning ability on later language acquisition, it is necessary to develop more comprehensive and reliable measures that assess both real time and offline learning of statistical regularities. The current proposal is a first step towards delineating how individual differences in statistical learning ability based on domain/modality may account for variations in later language outcomes.

The current protocol, providing robust methods to demonstrate large variability in learning behavior across domains and modalities, is designed to investigate individual differences of statistical learning. However, this protocol is not suitable for investigating questions such as whether visual statistical learning is inherently easier than auditory statistical learning. The interpretation of group-level performance differences between tasks is difficult due to all the confounding factors that we are not able to control

Disclosures

The authors have nothing to disclose.

Acknowledgments

We thank Yoel Sanchez Araujo and Wendy Georgan for their contribution in the initial design of the web-based platform. We thank An Nguyen and Violet Kozloff for their work on improving the web-based statistical learning tasks, implementing the fMRI tasks, and piloting the tasks in adult participants. We thank Violet Kozloff and Parker Robbins for their contribution in assisting data collection in children. We thank Ibrahim Malik, John Christopher, Trevor Wigal, and Keith Schneider at the Center for Biological and Brain Imaging at the University of Delaware for their assistance in neuroimaging data collection. This work is funded in part by the National Institute on Deafness and other Communication Disorders (PI: Qi; NIH 1R21DC017576) and the National Science Foundation Directorate for Social, Behavioral & Economic Sciences (PI: Schneider, Co-PI: Qi & Golinkoff; NSF 1911462).

References

1. Saffran, J. R., Newport, L. L., Aslin, R. N. Word Segmentation: The Role of Distributional Cues. *Journal of Memory and Language.* **35**, 606-621 (1996).
2. Newman, R., Ratner, N. B., Jusczyk, A. M., Jusczyk, P. W., Dow, K. A. Infants' early ability to segment the conversational speech signal predicts later language development: A retrospective analysis. *Developmental Psychology.* **42** (4), 643-655 (2006).
3. Graf Estes, K., Gluck, S. C. W., Grimm, K. J. Finding patterns and learning words: Infant phonotactic knowledge is associated with vocabulary size. *Journal of Experimental Child Psychology.* **146**, 34-49 (2016).
4. Evans, J. L., Saffran, J. R., Robe-Torres, K. Statistical Learning in Children With Specific Language Impairment. *Journal of Speech, Language, and Hearing Research.* **52** (2), 321-335 (2009).
5. Arciuli, J., Simpson, I. C. Statistical Learning Is Related to Reading Ability in Children and Adults. *Cognitive Science.* **36** (2), 286-304 (2012).
6. Qi, Z., Sanchez Araujo, Y., Georgan, W. C., Gabrieli, J. D. E., Arciuli, J. Hearing Matters More Than Seeing: A Cross-Modality Study of Statistical Learning and Reading Ability. *Scientific Studies of Reading.* **23** (1), 101-115 (2019).
7. Walenski, M., Tager-Flusberg, H. B., Ullman, M. T. Language in autism. *Understanding autism: From Basic Neuroscience to Treatment.* eds Moldin S. O., Rubenstein J. L. R., Boca Raton: Taylor and Francis Books, 175-203 (2006).
8. Siegelman, N., Frost, R. Statistical learning as an individual ability: Theoretical perspectives and empirical evidence. *Journal of Memory and Language.* **81**, 105-120 (2015).
9. Erickson, L. C., Kaschak, M. P., Thiessen, E. D., Berry, C. A. S. Individual Differences in Statistical Learning: Conceptual and Measurement Issues. *Collabra.* **2** (1), 14 (2016).
10. Shufaniya, A., Arnon, I. Statistical Learning Is Not Age-Invariant During Childhood: Performance Improves With Age Across Modality. *Cognitive Science.* **42** (8), 3100-3115 (2018).
11. Qi, Z. et al. An online platform for visual and auditory statistical learning for school-aged children (Version 1.0.0). *Zenodo.* (2020).
12. Conway, C. M., Christiansen, M. H. Statistical learning within and between modalities: Pitting abstract against stimulus-specific representations. *Psychological Science.* **17** (10), 905-912 (2006).
13. Saffran, J. R., Johnson, E. K., Aslin, R. N., Newport, E. L. Statistical learning of tone sequences by human infants and adults. *Cognition.* **70** (1), 27-52 (1999).
14. Siegelman, N., Bogaerts, L., Elazar, A., Arciuli, J., Frost, R. Linguistic entrenchment: Prior knowledge impacts statistical learning performance. *Cognition.* **177**, 198-213 (2018).
15. Armstrong, B. C., Frost, R., Christiansen, M. H. The long road of statistical learning research: past, present and future. *Philosophical Transactions of the Royal Society B: Biological Sciences.* **372** (1711), 20160047 (2017).
16. Siegelman, N., Bogaerts, L., Christiansen, M. H., Frost, R. Towards a theory of individual differences in statistical learning. *Transactions of the Royal Society B.* **372** (1711), 20160059 (2017).

17. Krogh, L., Vlach, H. A., Johnson, S. P. Statistical Learning Across Development: Flexible Yet Constrained. *Frontiers in Psychology.* **3** (JAN), 598 (2013).

18. Siegelman, N., Bogaerts, L., Elazar, A., Arciuli, J., Frost, R. Linguistic entrenchment: Prior knowledge impacts statistical learning performance. *Cognition.* **177** (August 2017), 198-213 (2018).

19. Plante, E., Patterson, D., Dailey, N. S., Kyle, R. A., Fridriksson, J. Dynamic changes in network activations characterize early learning of a natural language. *Neuropsychologia.* **62** (1), 77-86 (2014).

20. Milne, A. E., Wilson, B., Christiansen, M. H., Petkov, C., Marslen-Wilson, W. Structured sequence learning across sensory modalities in humans and nonhuman primates This review comes from a themed issue on The evolution of language. *Current Opinion in Behavioral Sciences.* **21**, 39-48 (2018).

21. Mcnealy, K., Mazziotta, J., Dapretto, M. Cracking the Language Code: Neural Mechanisms Underlying Speech Parsing. *Journal of Neuroscience.* **26** (29), 7629-7639 (2006).

22. McNealy, K., Mazziotta, J. C., Dapretto, M. Age and experience shape developmental changes in the neural basis of language-related learning. *Developmental Science.* **14** (6), 1261-1282 (2011).

23. Karuza, E. A., Emberson, L. L., Aslin, R. N. Combining fMRI and behavioral measures to examine the process of human learning. *Neurobiology of Learning and Memory.* **109**, 193-206 (2014).

24. Petersen, S. E., Dubis, J. W. The mixed block/event-related design. *NeuroImage.* **62** (2), 1177-1184 (2012).

25. Batterink, L. J., Paller, K. A., Reber, P. J. Understanding the Neural Bases of Implicit and Statistical Learning. *Topics in Cognitive Science.* **11**, 482-503 (2019).

26. Boersma, P., Weenink, D. Boersma, P., Weenink, D. Praat: doing phonetics by computer [Computer program]. *Version.* 6.1.14 (2020).

27. Boersma, P. Praat, a system for doing phonetics by computer. *Glot International.* **5** (9), 341-345 (2001).

28. Conway, C. M., Christiansen, M. H. 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 (2009).

29. 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. *Quarterly Journal of Experimental Psychology.* **64** (5), 1021-1040 (2011).

30. de Leeuw, J. R. jsPsych: A JavaScript library for creating behavioral experiments in a Web browser. *Behavior Research Methods.* **47** (1), 1-12 (2015).

31. Kana, R. K. et al. Probing the brain in autism using fMRI and diffusion tensor imaging. *Journal of Visualized Experiments.* (55), e3178 (2011).

32. Casey, B. J. et al. The Adolescent Brain Cognitive Development (ABCD) study: Imaging acquisition across 21 sites. *Developmental Cognitive Neuroscience.* **32**, 43-54 (2018).

33. Gorgolewski, K. J. et al. The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments. *Scientific Data.* **3** (1), 1-9 (2016).

34. Halchenko, Y. *nipy/heudiconv v0.6.0 (Version v0.6.0).* Zenodo. (2019).

35. Esteban, O. et al. FMRIPrep: a robust preprocessing pipeline for functional MRI. *Nature Methods.* **16**, 111-116 (2019).

36. Esteban, O. et al. Analysis of task-based functional MRI data preprocessed with fMRIPrep. *bioRxiv.* 694364 (2020).

37. Cox, R. W. AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomedical Research.* **29** (3), 162-173 (1996).

38. Avants, B. B. et al. A reproducible evaluation of ANTs similarity metric performance in brain image registration. *NeuroImage.* **54** (3), 2033-2044 (2011).

39. Dale, A. M., Fischl, B., Sereno, M. I. Cortical surface-based analysis: I. Segmentation and surface reconstruction. *NeuroImage.* **9** (2), 179-194 (1999).

40. Woolrich, M. W. et al. Bayesian analysis of neuroimaging data in FSL. *NeuroImage.* **45** (1 Suppl), S173-86 (2009).

41. Klein, A. et al. Mindboggling morphometry of human brains. *PLoS Computational Biology.* **13** (2), e1005350 (2017).

42. Arnon, I. Do current statistical learning tasks capture stable individual differences in children? An investigation of task reliability across modality. *Behavior Research Methods.* **52** (1), 68-81 (2020).

43. Finn, A. S., Hudson Kam, C. L. The curse of knowledge: First language knowledge impairs adult learners' use of novel statistics for word segmentation. *Cognition.* **108** (2), 477-499 (2008).

44. Krogh, L., Vlach, H. A., Johnson, S. P. Statistical learning across development: Flexible yet constrained. *Frontiers in Psychology.* **3** (JAN), 598 (2013).

45. De Leeuw, J. R., Motz, B. A. Psychophysics in a Web browser? Comparing response times collected with JavaScript and Psychophysics Toolbox in a visual search task. *Behavior Research Methods.* **48**, 1-12 (2016).

46. Lew-Williams, C., Saffran, J. R. All words are not created equal: Expectations about word length guide infant statistical learning. *Cognition.* **122** (2), 241-246 (2012).

47. Poulin-Charronnat, B., Perruchet, P., Tillmann, B., Peereman, R. Familiar units prevail over statistical cues in word segmentation. *Psychological Research.* **81** (5), 990-1003 (2017).

48. Leonard, J., Flournoy, J., Lewis-de los Angeles, C. P., Whitaker, K. How much motion is too much motion? Determining motion thresholds by sample size for reproducibility in developmental resting-state MRI. *Research Ideas and Outcomes.* **3**, e12569 (2017).

49. Power, J. D. et al. Ridding fMRI data of motion-related influences: Removal of signals with distinct spatial and physical bases in multiecho data. *Proceedings of the National Academy of Sciences of the United States of America.* **115** (9), e2105-e2114 (2018).