Concurrent Visual Learning of Adjacent and Nonadjacent Dependencies in Adults and Children

Lai-Sang Iao*, Jens Roeser, Lucy Justice, and Gary Jones

Department of Psychology, Nottingham Trent University, 50 Shakespeare Street, Nottingham, United Kingdom, NG1 4FQ

*Corresponding author

Address for correspondence: Department of Psychology, Nottingham Trent University, 50 Shakespeare Street, Nottingham, United Kingdom, NG1 4FQ

E-mail: lai-sang.iao@ntu.ac.uk

Phone: +44 (0)115 848 2387

Acknowledgements

This study was supported by the Quality Research funds awarded to Lai-Sang Iao from Nottingham Trent University. We thank Natalie Harrison, Adam Asmal, and Robin Clark for assisting in data collection. We also thank the participants, parents, and head teachers who made this research possible.

Author Note

Data and R scripts for the model fits of the study can be found on https://osf.io/bcyvr/ (Iao, Roeser, Justice, & Jones, 2020, December 29).

Declarations of interest: none

Concurrent Visual Learning of Adjacent and Nonadjacent Dependencies in Adults and Children

Abstract

Concurrent learning of adjacent and nonadjacent dependencies has been shown in adults only. This study extended this line of research by examining dependency-specific learning for both adjacent and nonadjacent dependencies concurrently in both adults and children. Seventy adults aged 18 to 64 (40 females, 30 males; Experiment 1) and 64 children aged 10 to 11 years (40 girls, 24 boys; Experiment 2) were tested with a new serial reaction time (SRT) task in which they were trained for 6 - 8 minutes on materials comprising equally probable adjacent and nonadjacent dependencies. They were then asked to discriminate between trained and untrained dependencies in a familiarity task. Both adults and children showed implicit concurrent learning of both adjacent and nonadjacent dependencies. The two dependency types were learnt to the same extent. However, adults showed a rapid, sustainable and dependency-specific sensitivity throughout the SRT task while children only showed a dependency-specific sensitivity to violations of expectations after exposure. When the two groups were statistically compared, only adults showed a dependency-specific learning effect after exposure. These findings are in line with the age-related improvement model of dependency learning.

Keywords: visual statistical learning; serial reaction time; nonadjacent dependency; implicit knowledge; child development

Submission date: 27/04/2020

Introduction

Humans of all ages rapidly extract statistical regularities among sensory elements across time and/or space (for recent reviews, see Conway, 2020; Saffran & Kirkham, 2018). Many of these regularities occur among temporally adjacent elements. For example, the transitional probability from one syllable to the next is high when the two syllables follow one another within a word, but low when the two syllables are spanning a word boundary. Using these transitional probabilities, adults, children and infants learn nonsense words and their boundaries in artificial languages (e.g., Aslin, Saffran, & Newport, 1998; Saffran, Aslin, & Newport, 1996; Saffran, Newport, & Aslin, 1996; Saffran, Newport, Aslin, Tunick, & Barrueco, 1997). Such statistical learning of adjacent dependencies has vastly been investigated not only across ages but also across modalities (e.g., Arciuli & Simpson, 2011; Bulf, Johnson, & Valenza, 2011; Campbell, Zimerman, Healey, Lee, & Hasher, 2012; Conway & Christiansen, 2005; Emberson, Conway, & Christiansen, 2011). However, adjacent dependencies are not the only type of dependencies that exist in natural language, music, motor and visual sequences. They coexist with nonadjacent dependencies in which two depending elements are separated by one or more intervening elements. For example, in English, there are dependencies between auxiliaries and inflectional morphemes (e.g., "*is* [kick]*ing*") and agreement between subject and verb (e.g., "the *boys* [in the red team] *are* great"). Both adjacent and nonadjacent dependencies may consist of similar elements in the same domain and co-occur in the same sequence (e.g., the root word [kick] consists of adjacent syllables with high transitional probability and is straddled by a nonadjacent dependency that consists of the auxiliary "*is*" and the inflectional morpheme "*ing*"). The two dependency types must be concurrently processed for the acquisition and development of language and various skills across domains. Therefore, it is crucial to understand how the two dependency types are concurrently processed and acquired as well as how this capacity may differ depending on age.

It is widely known that adjacent dependencies in various domains can be learned and assessed since infancy and throughout the lifespan with different types of measures (e.g., Arciuli & Simpson, 2011; Bulf et al., 2011; Saffran et al., 1996; Campbell et al., 2012; Conway & Christiansen, 2005; Ferdinand & Kray, 2017; Meulemans, Van der Linden, & Perruchet, 1998). One type of these measures is termed indirect, such as the serial reaction time (SRT) task which asks participants to respond to a visual cue that changes its location and does not explicitly refer participants to the to-be-learned sequence of location change (Nissen & Bullemer, 1987). This predetermined sequence of location change is therefore unknown to participants but they respond more quickly and accurately as the sequence repeats over time. When there are trials that are inconsistent with the sequence, participants' reaction times slow down compared to trials that are consistent with the sequence, suggesting that they had learned at least some aspects of the sequence. Given that participants' response times are measured throughout the task, both learning process and outcome can be revealed. The SRT task is therefore also regarded as an online measure of dependency learning. There are also direct measures, such as a familiarity task (e.g., Conway & Christiansen, 2005; Ferdinand & Kray, 2017) which explicitly asks participants to use their knowledge obtained from a previous learning session to discriminate between trained and untrained items. These measures, that can only reveal learning outcome, are also known as offline measures. Recent studies have emphasised the importance of using different types of measures to understand how participants learn dependencies, including both adjacent and nonadjacent dependencies, during exposure and whether their knowledge can be used intentionally after exposure (e.g., Bertels, Boursain, Destrebecqz, & Gaillard, 2015; Lammertink, van Witteloostuijn, Boersma, Wijnen, & Rispens, 2019; van Witteloostuijn, Lammertink, Boersma, Wijnen, & Rispens, 2019; Vuong, Meyer, & Christiansen, 2016).

Nonadjacent dependency is different from adjacent dependency in terms of temporal and/or spatial distance between the elements that make up the dependency. This distance does not exist in adjacent dependency whereas the distance between the beginning and ending elements of a nonadjacent dependency can have one or more intervening elements (e.g., "*is* [kick]*ing*," "the *boys* [in the red team] *are* great"). Nonadjacent dependencies are thus more difficult to learn compared to adjacent dependencies. Some conditions that highlight the nonadjacent dependencies are usually required for learning them. For example, an infinite set of intervening elements that fill the distance between the same beginning and ending elements of a nonadjacent dependency (e.g., "*is* [jump]*ing*," "the *boys* [on the right side] *are* great") is one well-known condition. Under this condition, the transitional probability between the beginning element and a particular intervening element is thus very low, making the remote dependency between the nonadjacent elements more salient for learning (Gómez, 2002). Using a large set of intervening elements, several studies showed that adults, children, and infants can track and learn the dependencies between nonadjacent elements in different domains with different types of measures (Gómez, 2002; Gómez & Maye, 2005; Iao, Ng, Wong, & Lee, 2017; Lammertink et al., 2019; Misyak, Christiansen, & Tomblin, 2010). Remillard (2008, 2010) also showed that adults can implicitly learn long-distance, nonadjacent dependencies. It is thus clear that both adjacent and nonadjacent dependency learning are not restricted by age groups, measure types, and stimulus domains. However, the two types of dependency learning may be in competition given that the adjacent relationship between nonadjacent elements and intervening elements may influence the acquisition of the dependency between nonadjacent elements (e.g., the transitional probability between the beginning element of a nonadjacent dependency and intervening elements is not low enough for the dependency between the nonadjacent elements to be salient for learning; Creel et al., 2004; Gebhart, Newport, & Aslin, 2009; Gómez, 2002; Gómez & Maye, 2005; Grama, Kerkhoff, & Wijnen, 2016; Newport & Aslin, 2004; Onnis, Monaghan, Richmond, & Chater, 2005). This could be due to an adjacency bias that learners tend to focus on adjacent relationships unless these adjacent relationships are too weak (Gómez, 2002). In this sense, concurrent learning of both dependency types could be difficult.

So far, to our knowledge, there have only been five studies examining concurrent learning of both adjacent and nonadjacent dependencies, all testing adults only. Two of the five studies used a direct offline measure (i.e., the 2-alternative forced-choice test) to assess participants' learning outcome by explicitly asking them to use their knowledge obtained from a previous learning session to identify trained items in sets of two choices (Creel et al., 2004; Romberg & Saffran, 2013). Both studies reported concurrent learning of both adjacent and nonadjacent dependencies in both audio-musical and audio-verbal domains when adjacent relationships were weakened to an extent that the adjacency bias could be reduced but not overridden completely. In addition to direct offline measures, Vuong et al. (2016) also used an indirect online measure (i.e., a modified SRT task) to reveal both learning process and outcome by requiring participants to make responses on a matrix of written pseudowords according to a sequence of spoken pseudowords as quickly and accurately as possible rather than directly assessing their knowledge. The modified SRT task was thus similar to the aforementioned SRT task that learning was inferred if participants responded more quickly to the ending elements of the dependencies as the training trials proceeded and slowed down for the violation trials in which the ending elements no longer followed the beginning elements of the dependencies. Their results demonstrated concurrent adjacent and nonadjacent dependency learning in the verbal domain in both types of measures despite no weakening of the dependencies between adjacent elements. However, a reaction time change from training to violation trials may reflect familiarity versus unfamiliarity of the trials rather than the dependencies per se that were embedded in the trials. Moreover, the study involved extended and multiple learning sessions.

The last two of the five studies used an indirect offline measure (i.e., a memory-based reproduction task) to assess participants' learning outcome by requiring them to recall each sequence without directly testing their knowledge of the dependencies embedded in the sequences (Conway et al., 2020; Deocampo, King, & Conway (2019). Both studies showed concurrent adjacent and nonadjacent dependency learning in visuo-verbal domain with a single 20-minute session. Deocampo et al. (2019) also showed concurrent adjacent and nonadjacent dependency learning in visuo-spatial domain using the same paradigm. The fact that the five studies used different measures and tested different domains suggests that concurrent learning of both adjacent and nonadjacent dependencies may not be restricted by measure types and stimulus domains. What is still unknown, compared to the previous literature on dependency learning, is whether concurrent learning of both adjacent and nonadjacent dependencies is restricted to adults only. The extent to which children may concurrently learn both adjacent and nonadjacent dependencies like adults is essential for understanding the acquisition and development of language and various skills across domains.

There have been three developmental models relating to adjacent and nonadjacent dependency learning (for a review, see Zwart, Vissers, Kessels, & Maes, 2019). One is the age invariance model, suggesting that children and adults do not differ. However, it is based on findings from studies that examined adjacent dependency learning only with both indirect and direct measures (e.g., Ferdinand & Kray, 2017; Meulemans et al., 1998). Previous findings that learning of adjacent or nonadjacent dependencies via indirect measures may not be demonstrated in direct measures were also age invariant (Lammertink et al., 2019;

Meulemans et al., 1998; Thomas & Nelson, 2001; van Witteloostuijn et al., 2019). The second model suggests that adults or older individuals learn better than younger individuals. It is based on findings from studies that tested either adjacent or nonadjacent dependency learning using either indirect or direct measures (e.g., Arciuli & Simpson, 2011; Gómez & Maye, 2005; Thomas et al., 2004). The last model suggests a gradual decline in learning nonadjacent dependencies across the lifespan using an indirect measure (Janacsek, Fiser, & Nemeth, 2012; Juhasz, Nemeth, & Janacsek, 2019; Nemeth, Janacsek, & Fiser, 2013). The difference between the models could be due to the complexity of the sequences tested in the studies according to Thomas et al. (2004), Janacsek et al. (2012), and Ferdinand and Kray (2017). For example, Thomas et al.'s to-be-learned sequence included two ambiguous items that were followed equally often by all other items whereas Meulemans et al.'s (1998) included only one. Thus, Thomas et al. suggested that their to-be-learned sequence was more complex than Meulemans et al.'s and that this increase in sequence complexity elicited the developmental effect found in their study but not Meulemans et al.'s. So far, there has not yet been a study examining how concurrent adjacent and nonadjacent dependency learning may or may not differ depending on age despite its importance in continuous lifelong learning and developing proficiency in natural languages and various skills across domains. The current study therefore sought to be the first to fill this gap by investigating concurrent learning of both adjacent and nonadjacent dependencies in both adults and children.

To understand how adults and children concurrently learn both adjacent and nonadjacent dependencies during exposure and whether their knowledge can be used intentionally after exposure, the current study developed a new paradigm that involved an indirect online measure and a direct offline measure. This paradigm is very brief and can be easily adapted into a child-friendly version. It was first used in Experiment 1 to examine whether dependency-specific learning occurred for both types of dependencies concurrently in adults and whether their knowledge could then be used intentionally after a very brief exposure. As such, Experiment 1 extended previous studies in two aspects. One was a full-range online measure of dependency-specific learning for both types of dependencies throughout the indirect online measure. This measure was crucial when dependencies were embedded in sequences/trials in an indirect online measure that relied on a reaction time change from training to violation sequences/trials to infer learning such as the SRT task. It is because a reaction time change from training to violation sequences/trials in this case may reflect a familiarity effect of the sequences/trials or a learning effect of the dependencies embedded in the sequences/trials (i.e., dependency-specific learning) or both. To avoid this uncertainty, which occurred in Vuong et al. (2016), a full-range online measure of dependency-specific learning was achieved by including a baseline of reaction times that did not concern the dependencies and comparing it against the reaction times for the dependencies throughout the indirect online measure. If participants showed a familiarity effect of the training sequences rather than dependency-specific learning, the reaction times for the dependencies should get faster to the same extent as the baseline as exposure proceeded and then slower to the same extent as the baseline when the dependencies were disrupted. If participants showed dependency-specific learning for both types of dependencies concurrently, the reaction times for both dependency types should be faster than the baseline during exposure and then slower specifically for the dependencies but not the baseline when the dependencies were disrupted. This dependency-specific learning would thus be independent of sequence learning as well as confounding effects due to fatigue and task requirement of mapping stimuli to response keys given that these effects would be consistent across stimuli that were part of the dependences and stimuli that were not part of the dependences (i.e., the baseline). As such, it could be considered as ‘pure statistical learning’ (Kóbor et al., 2018; Nemeth et al., 2013; Simor et al., 2019; Toth-Faber, Janacsek, Szollosi, Keri, & Nemeth, 2020). The other extension from previous studies was an examination of concurrent adjacent and nonadjacent dependency learning with a much briefer exposure than those in Vuong et al.'s, Deocampo et al.'s (2019) and Conway et al.'s (2020) studies.

The same paradigm as in Experiment 1 was then adapted into a child-friendly version in Experiment 2 to investigate the same questions in children. The two experiments together provided insights on whether concurrent learning of both adjacent and nonadjacent dependencies was age invariant. If it was age invariant, children should show a similar pattern of findings across both indirect online and direct offline measures as the adults did in Experiment 1. If it developed and changed from childhood to adulthood, children should show a different pattern of findings, whether in the indirect online measure or the direct offline measure or both, compared to that shown in the adults from Experiment 1. Further analysis was conducted to compare the findings across adults and children. Ethical approval for this study "Concurrent Visual Learning of Adjacent and Nonadjacent Dependencies in Adults and Children" was obtained from the Business, Law, and Social Sciences College Research Ethics Committee of Nottingham Trent University (No. 2016/24).

Experiment 1

Experiment 1 first examined whether dependency-specific learning occurred for both adjacent and nonadjacent dependencies concurrently in adults using a SRT task. It also tested whether adults' knowledge of the dependencies obtained in the SRT task could then be used intentionally. This was assessed by a familiarity task as the direct offline measure.

Method

Participants. Seventy adults (40 females, 30 males) aged 18 to 64 years with a mean age of 29.84 years (*SD* = 12.57) were recruited within Nottinghamshire, United Kingdom and volunteered to participate in the experiment after providing informed consents. All participants spoke English as their native language.

Materials and Design. Nine black and white visual shapes, adapted from Fiser and Aslin (2001; Figure 1), were used. Four of the nine shapes were randomly assigned to form the dependencies (e.g., Sign A, B, C and D). The remainders (e.g., Sign *e*, *f*, *g*, *h* and *i*) were used as intervening elements. Two sets of dependencies were constructed. Set 1 consisted of AB as the adjacent dependency and C-D as the nonadjacent dependency whereas Set 2 consisted of AD as the adjacent dependency and C-B as the nonadjacent dependency. Each participant was randomly assigned to one of the two sets. The ending elements of the dependencies (i.e., B and D) always appeared in the 4th and 8th positions of a 9-element sequence (e.g., *ef*AB*g*C*h*D*i* or *i*C*h*D*gf*AB*e*; see example sequences in Figure 1) in the SRT task. Each element was presented one by one and stayed on screen until a key response was recorded.

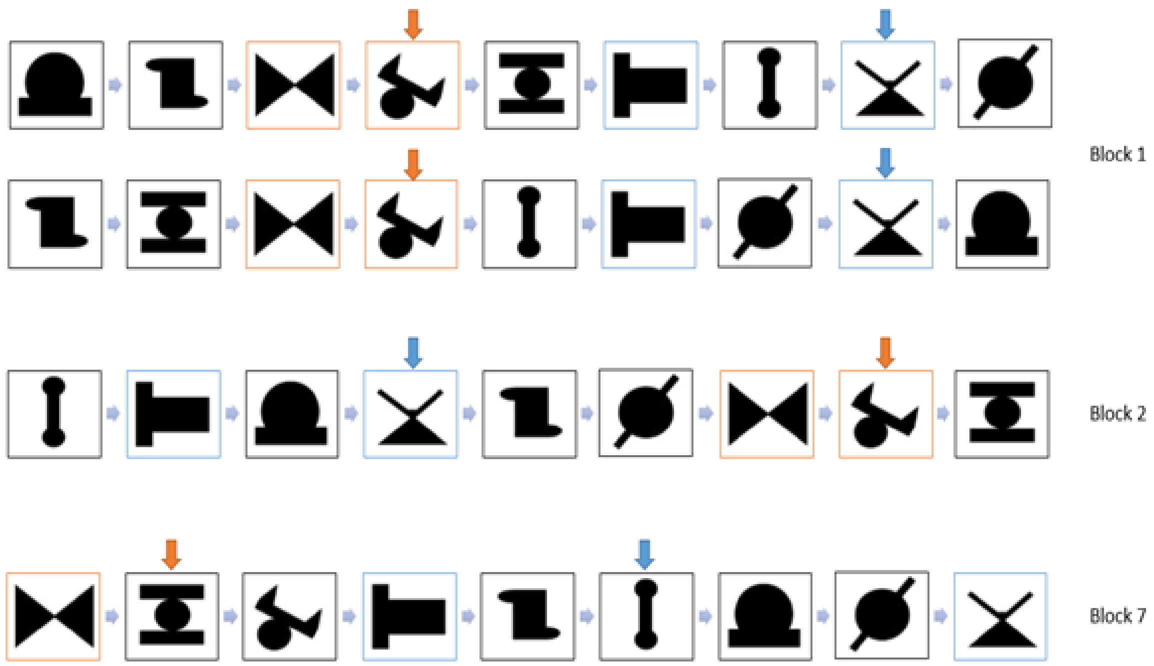

Figure 1

Stimuli, each approximately 3 x 3 cm, are presented one by one on screen in 9-element sequences (e.g., efABgChDi). For demonstration, AB as adjacent dependency is marked by orange (light grey) boxes, C-D as nonadjacent dependency is marked by blue (dark grey) boxes, and ef--g-h-i as intervening elements are marked by black boxes. Arrows between stimuli in a sequence represent stimuli's presentation order. The first two sequences are example sequences in Block 1. The third sequence is an example sequence in Block 2, in which the presentation order of the dependency types is counterbalanced. The last sequence is an example sequence in Block 7, in which the dependencies are disrupted. Disruption of adjacent dependency is highlighted by orange (light grey) arrows whereas disruption of nonadjacent dependency is highlighted by blue (dark grey) arrows. Note that neither boxes nor arrows are presented in the actual task. See the colour version of this figure online. Stimuli were adapted with permission from "Unsupervised statistical learning of higher-order spatial structures from visual scenes," by J. Fiser, R. N. Aslin, 2001, Psychological Science, 12, p. 500. Copyright 2001 by Sage.

The SRT task consisted of seven blocks in total. Each block was composed of five 9-element sequences. Blocks 1 to 6 were learning blocks in which the same adjacent and nonadjacent dependencies were repeated in each sequence so that the transitional probability between the beginning element (e.g., A) and the ending element (e.g., B) of a dependency was always 1.0. The presentation order of the dependency types in a sequence was counterbalanced across blocks. For example, if the adjacent dependency appeared in the 3rd and 4th positions and the nonadjacent dependency appeared in the 6th and 8th positions of the sequences in Block 1, the adjacent dependency would appear in the 7th and 8th positions and the nonadjacent dependency would appear in the 2nd and 4th positions of the sequences in Block 2. The five intervening elements randomly appeared in the rest of the positions of a sequence without re-appearing in the same position within the same block so that the transitional probability between the intervening elements (e.g., *f* or *g*) and a dependency (e.g., AB) was always .20. Block 7 was a transfer block in which the nine elements appeared in random positions for each of the five sequences with a restriction that the dependencies must be disrupted. Nine keys in the middle of a laptop keyboard (i.e., R, T, Y, F, G, H, V, B, N) were used as response keys to the nine elements in random assignment with a restriction that R and N were for the beginning elements of the dependencies versus Y and V for the ending elements of the dependencies for all participants. Each key was indicated by its corresponding element rather than its original letter. A cardboard cut-out was also used to cover the rest of the keyboard, exposing only the nine response keys.

The familiarity task consisted of 20 trials of 3-element sequences (e.g., *f*AB and C*h*B): 5 involved the adjacent dependency and 5 involved the nonadjacent dependency from each of the two sets. The order of the trials was randomized for each participant. Participants were asked whether they had encountered the sequence previously. Two keys on the keyboard (i.e., K and L) were used as 'Yes' and 'No' response keys to each of the trials. Another cardboard cut-out was used to cover the rest of the keyboard, exposing the 'Yes' and 'No' response keys only.

The two tasks were presented on a computer using OpenSesame (version 3.0.5). In the SRT task, participants had to press the key associated with its shape as quickly and accurately as possible. An example file of the task (child-friendly version used in Experiment 2) can be found on https://osf.io/bcyvr/ (Iao, Roeser, Justice, & Jones, 2020, December 29). In the familiarity task, participants were asked to identify the patterns in which the shapes appeared.

Procedure. Participants were tested individually in a laboratory by an experimenter. They first completed the SRT task in which they were trained on materials comprising both adjacent and nonadjacent dependencies. Depending on reaction time, this task lasted approximately 5 minutes. Before starting the task, the principles of the task were explained verbally and visually. Participants were told to press a corresponding key as fast and accurately as possible using their preferred hand when a shape appeared on screen. Therefore, the shapes were presented one at a time and each stayed on screen until a key response was recorded from any of the corresponding keys. A brief practice phase, comprising 3 examples that appeared in random order for 3 times, was performed to ensure understanding. The task then started, running from Block 1 to 7. No feedback was given. An encouraging slide saying "Good job! Keep going!" was presented after each block which allowed a break when necessary. Note that participants were not told about the regularities presented in the sequences throughout the task.

The familiarity task asked participants to identify the patterns in which the shapes appeared. Each shape was shown for 1250 milliseconds in groups of three and participants were required to determine whether they had seen the shape sequence before by pressing a 'Yes' or 'No' key. Feedback (i.e., "Good job! Keep going!") was given after each response, no matter if the response was correct or not. At the end of the task, participants were thanked and were given the opportunity to ask questions.

Results

Participant's median response accuracy was 99% (IQR = 1.70, range = 91%-100%) in the SRT task, suggesting that responses to each element were made with high accuracy. For the analysis of the data obtained from the SRT task, we removed incorrect responses (1.46%). Based on previous SRT studies (e.g., Thomas & Nelson, 2001), responses that were faster than 100 msecs and slower than 10,000 msecs (0.05%) were considered to be outliers and were also removed. Reaction times for the beginning elements of the dependencies were not included in the analysis because they were neither considered as baseline nor index of learning (as in Vuong et al., 2016; 24%). Finally, we excluded G-key responses where participants tended to rest their hands (11%). This key was not involved in any of the dependencies.

Data were analysed in Bayesian linear mixed effects models (Gelman et al., 2014; Kruschke, 2014; McElreath, 2016). The R package brms (Bürkner, 2017, 2018) was used to model the data using the probabilistic programming language Stan (Carpenter et al., 2016; Hoffman & Gelman, 2014; Stan Development Team, 2015).

Our statistical inference was based on the modelled parameter distributions expressed as the most probable posterior parameter values μ as well as their 95% Highest Posterior Density Interval (henceforth, HPDI) – the shortest interval containing 95% of the posterior probability mass (Kruschke, Aguinis, & Joo, 2012; Sorensen, Hohenstein, & Vasishth, 2016). These modelled data allow for direct statistical inference of patterns in the data and differences between conditions after accounting for sample specific variance.[1]

Data and R scripts for the model fits of the study can be found on https://osf.io/bcyvr/ (Iao et al., 2020, December 29).

Serial Reaction Time. The reaction times for the responses to the ending elements of the dependencies were extracted for Block 1 to 7 and modelled in Bayesian linear mixed effects models. The reaction times for elements that were not part of an adjacent or nonadjacent dependency were included in the model and treated as baseline. Model predictors were Dependency (levels: dependency, baseline), Adjacency (levels: adjacent, nonadjacent), Block (levels: 1-7) and all by-Block 2-way interactions of Dependency and Adjacency.

Reaction time data were fitted with shifted-lognormal distributions as frequently used in modelling literature (Heathcote, Brown, & Cousineau, 2004; Heathcote, Popiel, & Mewhort,

[1] Models were fitted with weakly informative priors (see McElreath, 2016) and run with 8,000 iterations on 3 chains with a warm-up of 4,000 iterations and no thinning. Model convergence was confirmed by the Rubin-Gelman statistic (Gelman & Rubin, 1992) and inspection of the Markov chain Monte Carlo chains.

1991; Wagenmakers & Brown, 2007). All models were fitted with random intercepts for participants and random intercepts for each stimulus image (Barr, Levy, Scheepers, & Tily, 2013; Bates, Kliegl, Vasishth, & Baayen, 2015) to account for random variance associated with individual differences between participants and image-specific variation. Partially pooled by-participant slope adjustments were included for all model predictors and trial order within blocks to account for the variance associated with repetitions within block.[2] This type of analysis differs from previous SRT studies which used aggregated response times (i.e., median reaction times; e.g., Meulemans et al., 1998; Thomas & Nelson, 2001). The inference based on aggregated data reported in these studies is problematic. This is because aggregating data artificially removes the within-block variance, reduces the natural skew of reaction time values, and therefore biases the inferential outcome. Instead we specified our model such that it accounted for both random error variance and the natural skew found in response time data.

The Bayesian model was used to infer the parameter distributions for each Dependency and by Block after accounting variance attributed to participants, stimulus image and repetitions. These statistically derived parameter values can be found in Figure 2.

[2] In brms syntax: RT ~ Block * Dependency * Adjacency + repetition + (Block * Dependency * Adjacency + repetition | participant) + (1 | stimulus).

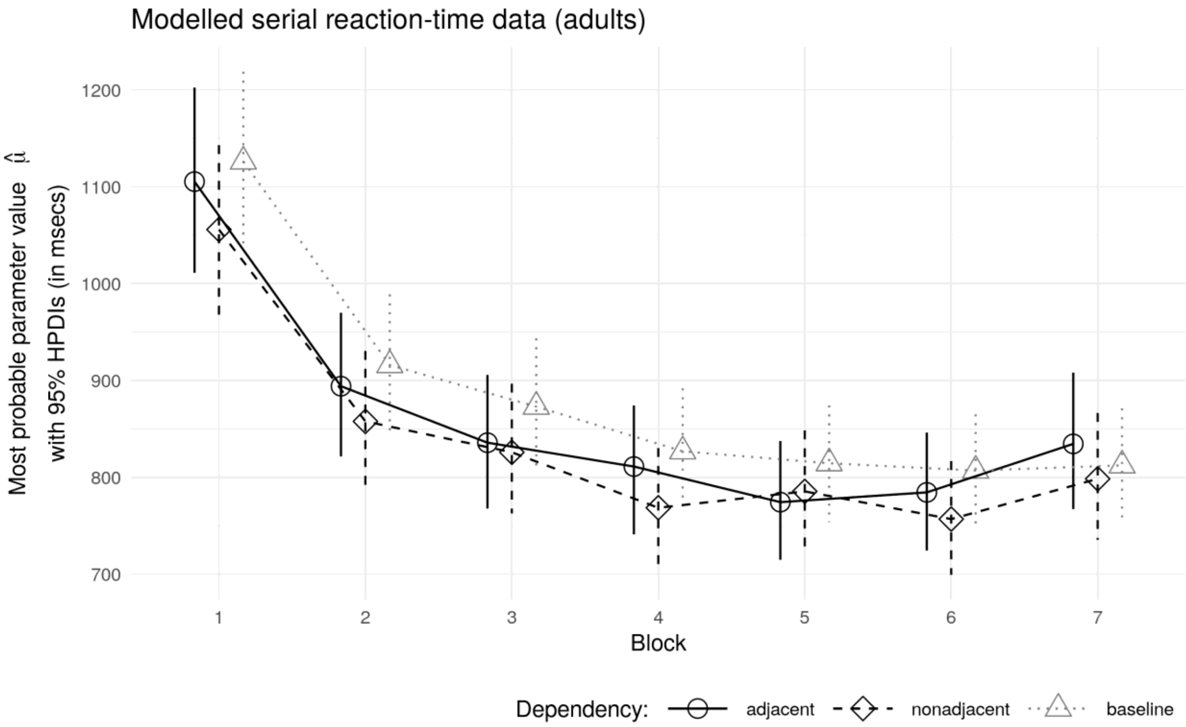

Figure 2

Summary of posterior reaction time data inferred from the Bayesian mixed effects model. Dots indicate the most probable a posteriori parameter $\hat{\mu}$, and error bars show 95% HPDIs.

Figure 2 shows that reaction times speeded up from Block 1 to 4, remained relatively stable until they slowed down from Block 6 to 7. From Block 1 to 6, reaction times were consistently faster for nonadjacent dependencies ($\hat{\mu}$ = -48 msecs, 95% HPDI[-104, 0]) but not for adjacent dependencies ($\hat{\mu}$ = -28 msecs, 95% HPDI[-77, 22]) compared to the baseline. This is evident as the HPDI for nonadjacent dependencies compared to the baseline does not include a difference of 0 msecs (thus, the absence of a difference) as a possible parameter value, while the HPDI for adjacent dependencies compared to the baseline does indeed include 0 msecs.

Only in Block 5, reaction times were found to be faster for adjacent dependencies compared to the baseline ($\hat{\mu}$ = -39 msecs, 95% HPDI[-78, 0]). This early adaptation to nonadjacent dependencies and late adaptation to adjacent dependencies was supported by an interaction of Adjacency and Block 4-5 ($\hat{\mu}$ = -15 msecs, 95% HPDI[-28, -2]).

Importantly we found evidence for an interaction of Dependency and Block 6-7 ($\hat{\mu}$ = 22 msecs, 95% HPDI[-47, 4]) with a posterior probability of P($\hat{\mu}$ < 0) = .95. The latter statement indicates the probability that the true parameter value of $\hat{\mu}$ is smaller than 0. In other words, there was a 95% probability that there was indeed a negative interaction between adjacent and nonadjacent dependencies versus baseline and the transition from Block 6 to 7. In the transfer Block 7 we found a slow-down of 50 msecs (95% HPDI[-7, 101]) for adjacent dependencies and 43 msecs (95% HPDI[-6, 92]) for nonadjacent dependencies, but for the baseline the slow-down was only 5 msecs (95% HPDI[-17, 29]) and thus virtually absent. From the posterior estimates we calculated the probability to observe longer values (i.e., a positive difference) in Block 7 compared to Block 6. The slow-down in Block 7 was found with a probability of P($\Delta\hat{\mu}$ > 0) = .97 for adjacent dependencies, a probability of P($\Delta\hat{\mu}$ > 0) = .96 for nonadjacent dependencies, and a probability of P($\Delta\hat{\mu}$ > 0) = .68 for baseline.

These results suggest that although nonadjacent dependencies were learnt faster than adjacent dependencies, the disruption of the dependencies in the transfer block affected responses to both types of the dependencies to the same extent while having almost no effect to the baseline stimuli. A summary of the model's fixed effects can be found in Appendix A1.

Explicit judgement. Participants' explicit judgements on trained and untrained trials in the familiarity task, based on accuracy, revealed chance-level responses for both adjacent (*M* = .51, *SD* = 0.5) and nonadjacent dependencies (*M* = .53, *SD* = 0.5). We evaluated chance-level performance in the context of equal variance Gaussian signal detection theory (EVSDT). Bayesian generalised linear mixed effects models (probit link function) provide a statistical tool to infer the EVSDT model's parameters (DeCarlo, 1998, 2010; Rouder & Lu, 2005; Rouder et al., 2007). Model parameters are the response criterion *c* representing the threshold for responding target present, and *d'* indicating the signal strength (i.e., the difference between signal and noise). Adjacency (levels: adjacent, nonadjacent) was included as fixed effect. Models were fitted with random intercepts for participants and stimulus with by-participant and by-stimulus slope adjustments for Adjacency. Table 1 shows the results of the EVSDT model. The *d'* parameter indicates that the inferred distance between noise and signal was non-different from zero. Further *d'* of adjacent compared to nonadjacent dependencies was negligibly different from zero. These results show that there was no evidence for explicit knowledge of both adjacent and nonadjacent dependencies.

Table 1

Bayesian estimates of the signal detection analysis. Criterion c and d' indicates the distance between the signal and the inferred noise distribution (i.e., the signal strength). The difference between adjacent and nonadjacent dependencies are shown for the criterion variable c and d'. The posterior distribution of the model parameters is indicated by the most probable inferred value $\hat{\mu}$ and 95% HPDIs.

Predictor	$\hat{\mu}$	HPDI[2.5%	97.5%]
-*c*	-0.04	-0.25	0.18
d'	0.04	-0.18	0.26
c (adjacent-nonadjacent)	0.15	-0.09	0.37
d' (adjacent-nonadjacent)	0.14	-0.17	0.45

Discussion

Experiment 1 tested whether adults showed dependency-specific learning for both types of dependencies concurrently and whether their knowledge could then be used intentionally. Our findings indicate that there was dependency-specific learning for both types of dependencies concurrently throughout the SRT task given that the reaction times for both dependency types were faster than the baseline during exposure and then slower specifically for the dependencies but not the baseline when the dependencies were disrupted. There was however no evidence for the knowledge obtained from the SRT task to be used intentionally after exposure. In the familiarity task, adults were not able to explicitly discriminate trained dependencies from untrained dependencies. Given that there was a dissociation between the indirect online and direct offline measures, we concluded that adults developed implicit but not explicit knowledge of both dependency types concurrently over a very brief exposure. Experiment 2 extended Experiment 1's investigation from adults to primary school-aged children.

Experiment 2

Experiment 2 investigated the same questions as Experiment 1 but in a sample of primary school-aged children. It used the same SRT task and the same familiarity task as Experiment 1. The only difference between Experiment 1 and 2 was that the two tasks were packaged as one coherent 'monster-catching' game to make the experiment child-friendly.

Method

Participants. Sixty-four children (40 girls, 24 boys) aged 10 to 11 years with a mean age of 131.15 months (*SD* = 3.30) took part. They were recruited from two different schools within Nottinghamshire, United Kingdom and volunteered after obtaining parental consents. All children spoke English as their native language, and none were reported as having special educational needs.

Materials and Design. The setup followed Experiment 1, except that the SRT task and the familiarity task were packaged and presented as one coherent 'monster-catching' game on a laptop computer using OpenSesame (version 3.0.5). That is, all shapes, response keys and other aspects of the experiment were exactly the same as Experiment 1. The only difference was that the beginning and end of the two tasks were edited to incorporate a 'monster-catching' theme. In the beginning of the SRT task, each shape was described as representing a monster by simultaneously presenting nine pairs of shapes and monsters on screen (for a link to an example file of the task, see Experiment 1; the monsters were adapted from Arciuli & Simpson, 2011). However, children were not asked to remember the correspondences between the shapes and the monsters. In fact, the monsters were irrelevant to task demands and were not presented on screen and keyboard during the task. The monsters were only shown on screen again at the end of the task to indicate they were all caught and kept in prison. In the familiarity task, the monsters were not shown at all.

Procedure. Children were tested individually in a room within their own school by an experimenter. They first completed the SRT task which lasted 6 - 8 minutes depending on reaction time. Same as the adults in Experiment 1, the principles of the task were explained verbally and visually before the tasked. However, children were told to catch each monster when its shape appeared on screen by pressing the corresponding key as fast and accurately as possible using their preferred hand. Everything else were exactly the same as Experiment 1 until the end of the task where children were praised for catching all the monsters and were verbally and visually informed that the monsters were kept at prison.

The familiarity task started by telling children that some monsters tended to appear in patterns and they were asked to help the police identify the patterns so that the police could catch them as soon as they broke out from the prison. Everything else were exactly the same as Experiment 1 until the end of the task where children were thanked for helping the police and were given the opportunity to ask questions.

Results

Children's median response accuracy was 98% (IQR = 2.14, range = 92%-100%) in the SRT task, suggesting that responses to each element by pressing the corresponding key were made with high accuracy. For the analysis of the data obtained from the SRT task, we removed incorrect responses (2.25%). Based on previous SRT studies, we also removed responses that were faster than 100 msecs and slower than 10,000 msecs (0.18%). Reaction times for the beginning elements of the dependencies were not included in the analysis (24%). Finally we excluded G-key responses where children tended to rest their hands (11%). This key was not involved in any of the dependencies. We used the same analysis tools as described in Experiment 1. For a link to the OSF repository see Experiment 1.

Serial Reaction Time. Reaction times were extracted and modelled in Bayesian linear mixed effects models in the exact same way as described in Experiment 1. Model predictors were Dependency, Adjacency, Block and all by-Block interactions of Dependency and Adjacency. The model was used to infer the parameter distributions for each Dependency and by Block after accounting variance attributed to participants, stimulus images and repetitions as described for Experiment 1. These statistically derived parameter values can be found in Figure 3.

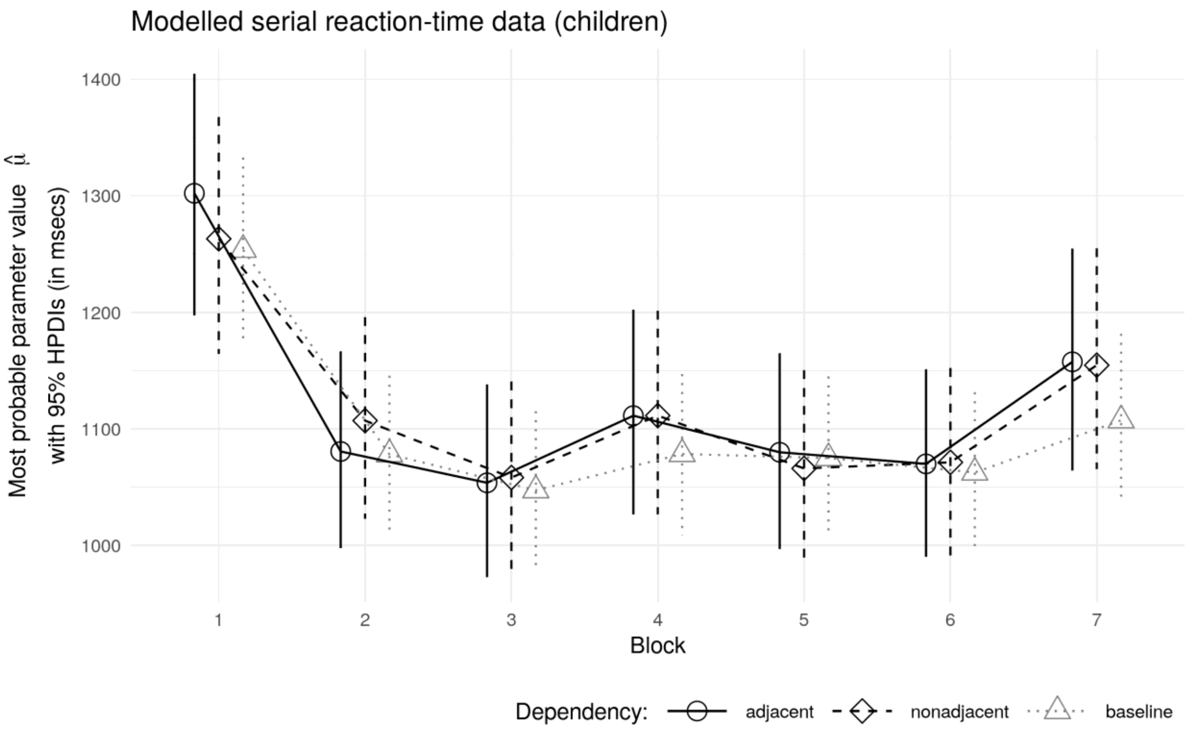

Figure 3

Summary of posterior reaction time data inferred from the Bayesian mixed effects model. Dots indicate the most probable a posteriori parameter $\hat{\mu}$, and error bars show 95% HPDIs.

Figure 3 shows that reaction times speeded up from Block 1 to 2, remained relatively stable until they slowed down from Block 6 to 7. Importantly in the transfer Block 7 we observed a slow-down of 83 msecs (95% HPDI[1, 176]) for adjacent dependencies, a slow-down of 87 msecs (95% HPDI[0, 171]) for nonadjacent dependencies, and a slow-down of 46 msecs (95% HPDI[2, 88]) for the baseline stimuli. Note that the slow-down observed for baseline stimuli was only half as large as the slow-down for both adjacent and non-adjacent dependencies. This larger slow-down for dependencies compared to baseline stimuli was weakly supported by evidence for an interaction of Dependency and Block 6-7 ($\hat{\mu}$ = -22 msecs, 95% HPDI[-55, 21]) with a posterior probability of P($\hat{\mu}$ < 0) = .84. This slow-down in Block 7 was found with a probability of P($\Delta\hat{\mu}$ > 0) = .98 for adjacent dependencies, a probability of P($\Delta\hat{\mu}$ > 0) = .97 for nonadjacent dependencies, and a probability of P($\Delta\hat{\mu}$ > 0) = .98 for baseline stimuli.

These results indicate that the disruption of dependencies in the transfer block affected responses to all stimulus images but had a larger effect on the reaction times for both types of the dependencies compared to baseline stimuli. A summary of the model's fixed effects can be found in Appendix A1.

Explicit judgement. Children's explicit judgements on trained and untrained trials in the familiarity task, based on accuracy, revealed chance-level responses for both adjacent (*M* = .51, *SD* = 0.5) and nonadjacent dependencies (*M* = .51, *SD* = 0.5). We evaluated chance-level performance in an EVSDT model (see Results section of Experiment 1). The results of the EVSDT model are shown in Table 2. The *d'* parameter indicates that the inferred distance between noise and signal was non-different from zero. Further *d'* of adjacent compared to nonadjacent dependencies showed a negligible difference. These results show that there was no evidence for explicit knowledge of both adjacent and nonadjacent dependencies.

Table 2

Bayesian estimates of the signal detection analysis. Criterion c and d' indicates the distance between the signal and the inferred noise distribution (i.e., the signal strength). The difference between adjacent and nonadjacent dependencies are shown for the criterion variable c and d'. The posterior distribution of the model parameters is indicated by the most probable inferred value $\hat{\mu}$ and 95% HPDIs.

Predictor	$\hat{\mu}$	HPDI[2.5%	97.5%]
-*c*	-0.07	-0.38	0.27
d'	0.09	-0.16	0.34
c (adjacent-nonadjacent)	0.15	-0.13	0.46
d' (adjacent-nonadjacent)	-0.04	-0.39	0.32

Discussion

Experiment 2 investigated whether children showed dependency-specific learning for both types of dependencies concurrently and whether their knowledge could then be used intentionally. Our findings indicate that children demonstrated concurrent sensitivity to both types of dependencies given that children's reaction times for both types of the dependencies decreased during exposure and increased when the dependencies were disrupted. However, whether this sensitivity was dependency-specific was not clearly reflected in the results. It was dependency-specific because the reaction times for both dependency types got slower specifically for the dependencies but not the baseline when the dependencies were disrupted.

It was also not dependency-specific because the reaction times for both dependency types got faster to the same extent as the baseline during exposure.

It is possible that the general reduction of reaction times during exposure was due to a practice effect that children were learning the mapping between shapes on screen and keys on keyboard rather than the dependencies per se. However, the slow-down observed for dependencies was two times larger than the slow-down observed for the baseline when the dependencies were disrupted. This can only be explained by a violation of the expectation regarding the ending elements of the dependencies specifically rather than the sequences in general. The slow-down for the baseline might only be a spillover effect from the dependency violations. That is, the disruptions in the dependencies may lead children to make cautious responses to other stimuli. Therefore, we argue that there was some level of dependency-specific learning evidenced for both adjacent and nonadjacent dependencies concurrently in children.

There was no evidence for children's knowledge obtained from the SRT task to be used intentionally after exposure. In the familiarity task, children were not able to explicitly discriminate trained dependencies from untrained dependencies. Given that there was a dissociation between the indirect online and direct offline measures, we concluded that children developed implicit but not explicit knowledge of both dependency types concurrently over a very brief exposure.

As such, there were similarities and differences between the findings in adults and children. Although both adults and children showed implicit but not explicit knowledge of both adjacent and nonadjacent dependencies, they performed differently throughout the SRT task. Further analyses were conducted to assess differences across adults and children on the SRT task.

Comparisons Between Adults and Children on Serial Reaction Time

Differences between adults (Experiment 1) and children (Experiment 2) observed in the SRT task were further evaluated, focusing on the dependency-specific learning effect indicated by faster reaction times for the two dependency types than the baseline during exposure and slower reaction times specifically for the dependencies but not the baseline when the dependencies were disrupted. We pooled the SRT data from both experiments and extracted the data from Blocks 1, 6 and 7.

This analysis was largely similar to the reaction-time models reported earlier. Reaction times were modelled in Bayesian linear mixed-effects models. Model predictors were Age Group (levels: adults, children), Dependency (levels: dependency, baseline), Adjacency (levels: adjacent, nonadjacent), Learning Block (levels: 1, 6), Transfer Block (levels: 6, 7), and all by-Dependency and by-Adjacency interactions with Group and Block. All predictors were sum-coded following Schad et al. (2020). Reaction time data were fitted with a shifted-lognormal distribution, random intercepts for each stimulus image, and participants with by-participant slope adjustments for all main effects and interactions.

In addition to the parameter estimates as in the previous results sections, we used the posterior to calculate the standardised effect sizes δ, defined as $\delta = \frac{\mu}{\sigma}$, where μ is the parameter value of the effect of interest and σ is the standard deviation. The effect size was calculated to allow comparisons across age groups (Wagenmakers et al., 2010). Because the standard deviation is necessary to calculate the effect size, we specified the model with two different variance components, one for the child group and one for the adult group as the population standard deviation is likely to be different for each group (Baguley, 2009).

For the effect sizes, we determined the region of practical equivalence (henceforth, ROPE) to assess the uncertainty of the effect size (Makowski et al., 2019). The ROPE is the range of values that are practically equivalent to a null effect (e.g., Kruschke, 2010, 2011). We set the ROPE to be -0.1 and 0.1 (Kruschke, 2018) which is the range of negligible effect sizes (Cohn, 1988). The value returned is the proportion of posterior samples (of the effect size) that fall inside the ROPE. In other words, the ROPE value indicates the extent to which the posterior cannot rule out a negligible effect. A meaningful effect size should have a small proportion of posterior samples within the ROPE.

Table 3 summaries the modeling outcome. We found compelling evidence for longer reaction times for children compared to adults. Overall, adjacent dependencies were responded to faster than nonadjacent dependencies. Reaction times for Block 6 were shorter than Block 1; and reaction times for Block 7 were longer than for Block 6. Further evidence was found for two-way interactions of Age Group and Dependency, and Age Group and Learning Block. Importantly, we found evidence for a three-way interaction of Dependency by Transfer Block and Age Group. The variance estimate for the child group ($\hat{\sigma}$ = 0.37, 95% HPDI[0.36, 0.38]) was larger than the variance estimate for the adult group ($\hat{\sigma}$ = 0.32, 95% HPDI[0.31, 0.32]).

Table 3

Age-group comparison for SRT data. Estimated parameter values (in msecs) for main effects and interactions of Learning Block (levels: 1, 6) and Transfer Block (levels: 6, 7), Age Group (levels: adults, children), Dependency (levels: dependency, baseline), Adjacency (levels: adjacent, nonadjacent).

Predictor	$\hat{\mu}$	95% HPDI	$P(\hat{\mu} < 0)$
Main effects			
Age Group	-327	[-383, -270]	>.999
Dependency	3	[-30, 34]	.447
Adjacency	-23	[-41, 0]	.977
Learning Block	217	[189, 243]	<.001

Transfer Block	-35	[-54, -19]	>.999
Two-way interactions			
Age * Dependency	-64	[-88, 41]	>.999
Age * Adjacency	-12	[-30, 8]	.877
Age * Learning Block	69	[49, 92]	<.001
Age * Transfer Block	-2	[-18, 13]	.616
Dependency * Learning Block	4	[-11, 22]	.278
Dependency * Transfer Block	-16	[-40, 5]	.928
Adjacency * Learning Block	-6	[-19, 8]	.757
Adjacency * Transfer Block	-2	[-14, 19]	.415
Three-way interactions			
Dependency * Learning Block * Age	-3	[-19, 14]	.618
Dependency * Transfer Block * Age	-25	[-41, -7]	.997
Adjacency * Learning Block * Age	3	[-11, 17]	.327
Adjacency * Transfer Block * Age	-1	[-14, 15]	.499

Note. $\hat{\mu}$ indicates the most probable parameter value; 95% HPDI is the range containing 95% of the posterior probability mass; $P(\hat{\mu} < 0)$ is the probability of the true parameter value being smaller than 0; '*' indicates interactions.

Figure 4 illustrates the modeled data illustrating the speedup from Block 1 to 6 and the slowdown from Block 6 to 7 for the adult and child groups. The slowdown from Block 6 to 7 was tested within Age Group and Dependency to gain further insight into the source of the three-way interactions. Reported are the effect sizes for each comparison.

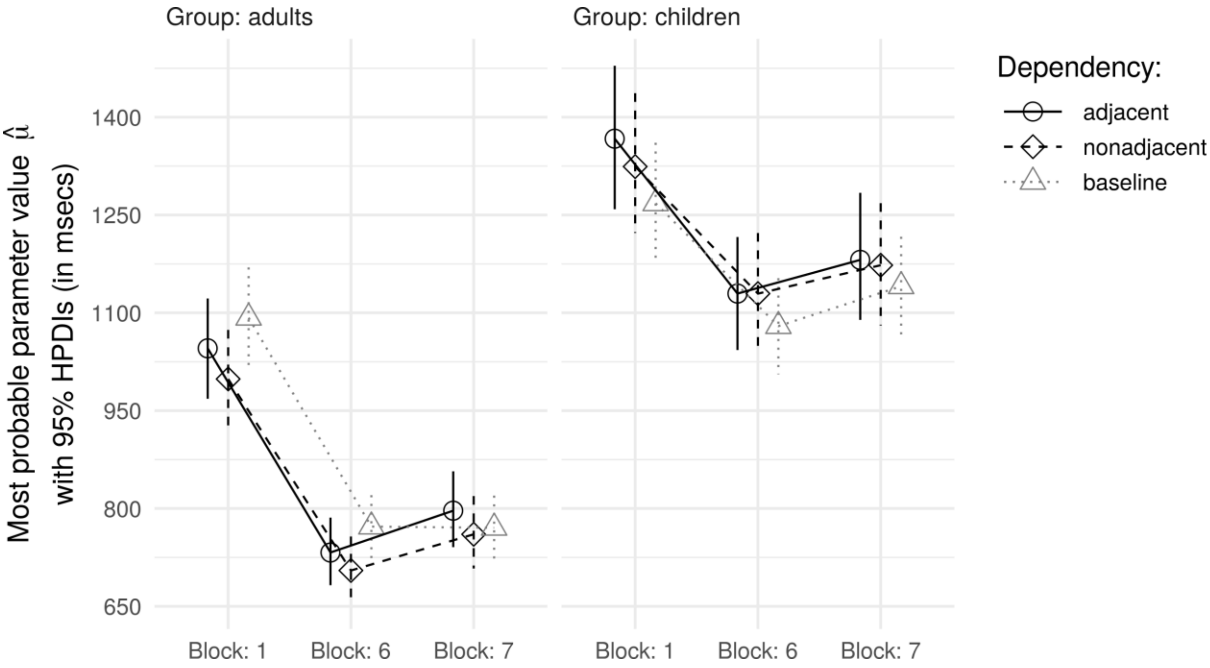

Figure 4

Modeled SRT data (group comparison). Reaction-time data are shown by age group for Block 1 compared to Block 6 and Block 6 compared to Block 7. Shown are the modelled reaction-time data with 95% HPDI (in msecs) for both dependencies (adjacent, nonadjacent) and the baseline.

We observed an interaction of Dependency, Transfer Block, and Age Group. For the adult group, we observed a small effect for a slowdown from Block 6 to 7 for the dependencies ($\hat{\delta}$ = 0.26, 95% HPDI[0.11, 0.39], ROPE = 0) but negligible evidence for a slowdown for the baseline ($\hat{\delta}$ = -0.01, 95% HPDI[-0.08, 0.07], ROPE = 100%). For the child group, we observed negligible to small slowdown effects for both types of dependencies ($\hat{\delta}$ = 0.11, 95% HPDI[-0.02, 0.25], ROPE = 42%) and for the baseline ($\hat{\delta}$ = 0.14, 95% HPDI[0.07, 0.23], ROPE = 6%). Therefore, when we compared across the two groups adults showed a dependency-specific learning effect after exposure while children did not.

General Discussion

The present study is the first to investigate concurrent learning of adjacent and nonadjacent dependencies in both adults and children, using both indirect online and direct offline measures. Our findings suggested that both adults and children developed implicit but not explicit knowledge of both adjacent and nonadjacent dependencies over a very brief exposure. Moreover, both adjacent and nonadjacent dependencies were learnt implicitly and concurrently to the same extent. However, adults showed a rapid, sustainable and dependency-specific sensitivity throughout the indirect online measure while children only showed a dependency-specific sensitivity to violations of expectations after exposure. When the two groups were statistically compared, only adults showed a dependency-specific learning effect after exposure. Therefore, our results revealed an age-related difference in implicit concurrent learning of both adjacent and nonadjacent dependencies.

First, adults showed faster reaction times for both dependencies compared to those for the baseline during exposure. However, children's reaction times for the dependencies were getting faster to the same extent as the baseline as exposure proceeded. Second, adults showed slower reaction times for both dependency types but not for the baseline when the dependencies were disrupted after exposure. Children, in contrast, showed a slow-down for the baseline that was half the magnitude of the slow-down observed for both dependency types when the dependencies were disrupted after exposure. These findings corresponded to Thomas et al.'s (2004) findings that adults showed significant dependency-specific learning for adjacent dependencies throughout exposure whereas children did not show the same until later part of the exposure, in which the magnitude of their learning effect was still significantly smaller than adults'. Both studies suggest that adults implicitly and concurrently learnt dependencies quicker and better than children. This is in line with the age-related improvement model and other evidence that suggests dependency learning improves with age (e.g., Arciuli & Simpson, 2011; Gómez & Maye, 2005). In addition to the consistency with previous literature, the current study importantly extended previous literature by providing the first evidence for an age-related improvement in implicit concurrent learning of adjacent and nonadjacent dependencies.

The age-related improvement observed in the current study was however inconsistent with the age invariance model (e.g., Ferdinand & Kray, 2017; Meulemans et al., 1998) and the age-related decline model (e.g., Janacsek et al., 2012; Juhasz et al., 2019). This inconsistency may be explained by the difference between the paradigms used in these studies. For example, the use of indirect online measures that assess dependency-specific learning could be one possible factor that captured age-related differences in dependency learning. This possibility was suggested based on the fact that both Thomas et al. (2004) and the current study used indirect online measures that measure sequence- or dependency-specific learning and identified similar age-related improvement during exposure. By comparing reaction times for training sequence or dependencies against those for random sequences or baseline stimuli, both studies suggested that children may take longer than adults to practice the mapping between response keys and stimuli before their reaction times for training sequence or dependencies may become faster than those for random sequences or baseline stimuli to indicate sequence-/dependency-specific learning during exposure. The same comparison of reaction times further allowed the current study to reveal an age-related improvement in dependency-specific learning after exposure. That is, the disruption of the dependencies affected children's reaction time to baseline stimuli but not adults'. Without a baseline of reaction times, Ferdinand and Kray (2017)'s indirect online measure did not examine sequence- or dependency-specific learning so they were not able to capture the age differences in sequence- or dependency-specific learning that took place during and after exposure.

Moreover, a SRT task that aggregates reaction times across training trials (earlier part of the task) versus violation trials (later part of the task) and compares the aggregations, such as the one in Ferdinand and Kray's (2017) study, lacks reliability (Stark-Inbar, Raza, Taylor, & Ivry, 2017; West, Vadillo, Shanks, & Hulme, 2018). In contrast, the Alternating SRT task that compares high-frequency triplets (i.e., the third element was highly predictable from the first element, e.g., 1 frequently followed 2x where x represented a random element) versus low-frequency triplets (i.e., the third element was rarely predictable from the first element, e.g., 3 infrequently followed 2x) on a continuous basis throughout the task, such as the one in Janacsek et al. (2012), has good reliability (Stark-Inbar et al., 2017). Given that our measure of dependency-specific learning was a continuous comparison between dependency and baseline throughout the task and our data were not aggregated, it may not have a similar reliability issue as the traditional SRT task. These contrasts between the studies highlighted the importance of including a measure of sequence- or dependency-specific learning throughout a task to understand how individuals across ages learn dependencies over time, supporting the recent emphases proposed by Janacsek et al., Lammertink et al. (2019) and van Witteloostuijn et al. (2019).

Using the Alternating SRT task which claimed to reveal 'pure' sequence-specific learning if reaction times for high-frequency triplets were faster than low-frequency triplets (Song, Howard, & Howard, 2007), Janacsek et al. (2012), Juhasz et al. (2019) and Nemeth et al. (2013) reported an age-related decline in their findings. However, the sequence-specific learning (or 'maximised learning' and 'pure statistical learning' in Nemeth et al., 2013) they observed confounded with reactive inhibition which was defined as accumulative performance deterioration during a continuous task (Hull, 1943; Pan & Pickard, 2015; Török, Janacsek, Nagy, Orbán, & Nemeth, 2017). Given that fatigue would have been controlled by comparing reaction times for high- versus low-frequency triplets, reactive inhibition may not be simply equivalent to fatigue. It is also an inhibition that occurs after a response set has been initiated (for a review, see Meyer & Bucci, 2016). This inhibition of a response set should have a larger effect on the reaction times for low- than high-frequency triplets given that the third elements in the low-frequency triplets are unexpected (e.g., 3 is unexpected after 2x). Therefore, reactive inhibition may affect the magnitude of the sequence-specific learning reported in Janacsek et al.'s, Juhasz et al.'s and Nemeth et al.'s studies. Moreover, reactive inhibition is known to improve with age (e.g., Shen et al., 2020; van de Laar et al., 2011; Williams et al., 1999). Adults are more capable in response inhibition such that their reaction times for low-frequency triplets may not be affected as much as children's, leading to a larger magnitude of sequence-specific learning reported in children than adults in their studies. In contrast, reactive inhibition may not have a larger effect on the reaction times for the baseline stimuli than those for the dependencies in the current study given that the baseline stimuli and the dependencies were of different sets of stimuli (i.e., the predictability of the stimuli involved in the dependencies did not affect the reaction times for the baseline stimuli). Thus, if there was reactive inhibition, it may have been controlled by comparing the reaction times for the baseline stimuli against those for the dependencies in the current study. However, further investigation is required to verify this speculation.

Sequence complexity, as mentioned earlier, could be another possible factor in explaining why age-related differences emerged in some studies but not others. Given that the sequence used in the current study concurrently included both adjacent and nonadjacent dependencies as well as five intervening elements, it could arguably be the most complex compared to previous studies. In this sense, however, it is not clear whether the age-related improvement revealed in the current study was an effect of sequence complexity or concurrent adjacent and nonadjacent dependency learning or both. This uncertainty could be a limitation of the current study and may require further research to clarify although sequence complexity may inevitably confound in concurrent adjacent and nonadjacent dependency learning. Given that sequences in natural languages and various skills across domains are very likely to be complex and involving both adjacent and nonadjacent dependencies concurrently, we suggest that the sequences used and the age-related improvement reported in the current study were ecologically consistent. It is likely that individuals may become more rapidly sensitive to both adjacent and nonadjacent dependencies in a given sequence and their concurrent learning effect for both dependency types may become larger and more specific as they develop from childhood to adulthood. However, this developmental improvement may be related to other perceptual, motor and/or cognitive development such as stimulus-response mapping which has been suggested to elicit similar age-related changes in brain activity as dependency learning (Casey, Thomas, Davidson, Kunz, & Franzen, 2002; Thomas et al., 2004). This possibility and the nature of the developmental improvement shown in the current study were not tested and could be considered as further limitations of the study.

Findings of the current study also showed that both adults and children implicitly and concurrently learnt both adjacent and nonadjacent dependencies to the same extent. This is in line with Vuong et al.'s (2016) findings that nonadjacent dependencies were learnt as well as adjacent dependencies in adults. Both studies thus reported findings that are inconsistent with Deocampo et al.'s (2019) findings that adjacent dependencies were learnt better than nonadjacent dependencies. They also go against previous suggestions that there is a default of tracking adjacent over nonadjacent dependencies (i.e., the adjacency bias; Gómez, 2002) and that the two types of dependency learning were in competition (e.g., Creel et al., 2004; Gebhart et al., 2009; Grama et al., 2016; Newport & Aslin, 2004). A possible explanation for this inconsistency is that the two sets of studies used different ways to create their adjacent and nonadjacent dependencies. Both Vuong et al.'s study and the current study used two different sets of stimuli to form their adjacent and nonadjacent dependencies, allowing for an observation of nonadjacent dependency learning that is independent of the observation of adjacent dependency learning. Other studies such as Deocampo et al.'s used the same set of stimuli for both types of dependencies so that the two dependency types shared the same beginning element or the same ending element, leading to an inevitable competition between the two types of dependency learning. Given the inherent advantage of the adjacent dependencies, they were more easily learnt than the nonadjacent dependencies and their learning mediated the learning of the nonadjacent dependencies. Therefore, the nonadjacent dependency learning observed may be a less clear-cut observation than those in Vuong et al.'s study and the current study. However, it was more ecologically consistent according to Deocampo et al. Taken together, it is likely that, in principle, both adjacent and nonadjacent dependencies can be concurrently learnt to the same extent, but in practice when the two dependency types are intertwined, adjacent dependencies are learnt better than nonadjacent dependencies.

The current study also tested whether adults and children could intentionally use their concurrent knowledge of both adjacent and nonadjacent dependencies after exposure. There was no evidence for both adults' and children's knowledge obtained from the indirect online measure to be shown in the direct offline measure. This is consistent with previous findings that learning of adjacent or nonadjacent dependencies via indirect online measures may not be demonstrated in direct offline measures in both adults and children (Lammertink et al., 2019; Meulemans et al., 1998; Thomas & Nelson, 2001; van Witteloostuijn et al., 2019). However, this is inconsistent with Vuong et al.'s (2016) finding that adults showed concurrent learning of both adjacent and nonadjacent dependencies in both indirect online and direct offline measures. This inconsistency could be due to the difference in exposure length between the studies. According to Cleeremans' (2006) model of unconscious cognition, learning a sequence from exposure may first be implicit and then become explicit when the strength of the relevant representation gradually emerges into a high-quality representation over time through practice. Therefore, it is possible that Vuong et al.'s extended and multiple learning sessions allowed concurrent adjacent and nonadjacent dependency learning to be observed via direct offline measures whereas the current brief learning session did not. However, further investigation on the effect of exposure length on concurrent adjacent and nonadjacent dependency learning via both indirect online and direct offline measures is needed to test this explanation and further verify Cleeremans' model.

It is also possible that the direct offline measure used in the current study, that is, the familiarity task, might not be sensitive to the knowledge adults and children obtained via indirect online measures although it has been widely used in this area of research (Conway & Christiansen, 2005; Ferdinand & Kray, 2017; Gómez, 2002; Iao et al., 2017; Meulemans et al., 1998; Misyak et al., 2010; Vuong et al., 2016). There is evidence indicating that recognition of sequence fragments may not be as sensitive as other measures, such as free recall, in assessing the explicit knowledge obtained in the context of the SRT task in adults (Rauch et al., 1995; for reviews, see Cleeremans, Destrebecqz, & Boyer, 1998; Destrebecqz & Peigneux 2005). Other studies, using non-SRT tasks and different direct offline measures, such as forced-choice completion, found that adjacent dependency learning was accessible via both indirect and direct measures in both adults and children after relatively short exposure (Batterink, Reber, Neville, & Paller, 2015; Bertels et al., 2015). However, there are also evidence of dissociations found between non-SRT tasks and different direct offline measures when examining either adjacent or nonadjacent dependency learning, suggesting implicit knowledge only, in both adults and children (Kim, Seitz, Feenstra, & Shams, 2009; Lammertink et al., 2019; van Witteloostuijn et al., 2019). Therefore, the tasks per se may not sufficiently explain the findings. Rather, it is still crucial to understand how individuals across the lifespan learn both adjacent and nonadjacent dependencies during exposure and how their knowledge can be used intentionally after exposure.

To conclude, this study provides the first evidence that shows dependency-specific learning for both adjacent and nonadjacent dependencies concurrently in both adults as well as children. Both groups demonstrated implicit but not explicit knowledge of the dependencies. Children may become more rapidly sensitive to both dependency types in a given sequence and their concurrent learning effect may become larger and more specific as they develop into adulthood. Although further cross-sectional and longitudinal investigations across ages are warranted to address the limitations of the current study, the current findings contribute to our understanding of statistical learning in terms of its development and generality across indirect and direct measures. This broader view of statistical learning may help in explaining language acquisition and various skill learning as well as proficiency development across domains and the lifespan.

References

Arciuli, J., & Simpson, I. C. (2011). Statistical learning in typically developing children: The role of age and speed of stimulus presentation. *Developmental Science, 14*, 464-473. http://doi.org/10.1111/j.1467-7687.2009.00937.x

Aslin, R. N., Saffran, J. R., & Newport, E. L. (1998). Computation of conditional probability statistics by 8-month-old infants. *Psychological Science, 9*, 321-324.

Baguley, T. (2009). Standardized or simple effect size: What should be reported? *British Journal of Psychology*, *100* (3), 603-617. http://doi.org/0.1348/000712608X377117

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255-278.

Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2015). Parsimonious mixed models. *arXiv preprint arXiv:1506.04967*.

Batterink, L. J., Reber, P. J., Neville, H. J., & Paller, K. A. (2015). Implicit and explicit contributions to statistical learning. *Journal of Memory and Language, 83*, 62-78. http://doi.org/10.1016/j.jml.2015.04.004

Bertels, J., Boursain, E., Destrebecqz, A., & Gaillard, V. (2015). Visual statistical learning in children and young adults: How implicit? *Frontiers in Psychology, 5*, 1541. http://doi.org/10.3389/fpsyg.2014.01541

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

Bürkner, P.-C. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software, 80*(1), 1-28. doi:10.18637/jss.v080.i01

Bürkner, P.-C. (2018). Advanced Bayesian multilevel modeling with the R package brms. *The R Journal, 10*(1), 395-411. doi:10.32614/RJ-2018-017

Campbell, K. L., Zimerman, S., Healey, M. K., Lee, M. M. S., & Hasher, L. (2012). Age differences in visual statistical learning. *Psychology and Aging*, *27*, 650-656. http://doi.org/10.1037/a0026780

Carpenter, B., Gelman, A., Hoffman, M., Lee, D., Goodrich, B., Betancourt, M., . . . Riddell, A. (2016). Stan: A probabilistic programming language. *Journal of Statistical Software*, *20*.

Casey, B. J., Thomas, K. M., Davidson, M. C., Kunz, K., & Franzen, P. L. (2002). Dissociating striatal and hippocampal function developmentally with a stimulus-response compatibility task. *Journal of Neuroscience*, *22*(19), 8647-8652.

Cleeremans, A. X. (2006). Conscious and unconscious cognition: A graded, dynamic perspective. In Q. Jing, M. Rosenzweig, G. d'Ydewalle, H. Zhang, H.-C. Chen, & K. Zhang (Eds.), *Progress in psychological science around the world: Vol. 1. Neural, cognitive, and developmental issues* (pp. 401-418). Hove, England: Psychology Press.

Cleeremans, A., Destrebecqz, A., & Boyer, M. (1998). Implicit learning: News from the front. *Trends in Cognitive Sciences, 2*, 406-416. http://www.ncbi.nlm.nih.gov/pubmed/21227256

Cohn, J. (1988). *Statistical power analysis for the behavioral sciences*. Lawrence Earlbam Associates.

Conway, C. M. (2020). How does the brain learn environmental structure? Ten core principles for understanding the neurocognitive mechanisms of statistical learning.

Neuroscience and Biobehavioral Reviews, *112*, 279-299. https://doi.org/10.1016/j.neubiorev.2020.01.032

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*, 24–39. http://doi.org/10.1037/0278-7393.31.1.24

Conway, C. M., Eghbalzad, L., Deocampo, J. A., Smith, G. N. L., Na, S., & King, T. Z. (2020). Distinct neural networks for detecting violations of adjacent versus nonadjacent sequential dependencies: An fMRI study. *Neurobiology of Learning and Memory*, *169*, 107175. https://doi.org/10.1016/j.nlm.2020.107175

Creel, S. C., Newport, E. L., & Aslin, R. N. (2004). Distant Melodies: Statistical Learning of Nonadjacent Dependencies in Tone Sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*, 1119-1130. http://doi.org/10.1037/0278-7393.30.5.1119

DeCarlo, L. T. (1998). Signal detection theory and generalized linear models. *Psychological Methods*, *3*(2), 186-205.

DeCarlo, L. T. (2010). On the statistical and theoretical basis of signal detection theory and extensions: Unequal variance, random coefficient, and mixture models. *Journal of Mathematical Psychology*, *54*(3), 304-313.

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. http://doi.org/10.3389/fpsyg.2019.01107

Destrebecqz, A., & Peigneux, P. (2005). Methods for studying unconscious learning. *Progress in Brain Research*, *150*, 69-80. https://doi.org/10.1016/S0079-6123(05)50006-2

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. *The Quarterly Journal of Experimental Psychology*, *64*(5), 1021-1040. http://doi.org/10.1080/17470218.2010.538972

Ferdinand, N. K., & Kray, J. (2017). Does language help regularity learning? The influence of verbalizations on implicit sequential regularity learning and the emergence of explicit knowledge in children, younger and older adults. *Developmental Psychology, 53*, 597-610. https://doi.org/10.1037/dev0000262

Fiser, J., & Aslin, R. N. (2001). Unsupervised statistical learning of higher-order spatial structures from visual scenes. *Psychological Science, 12*, 499-504. http://doi.org/10.1111/1467-9280.00392

Gebhart, A. L., Newport, E. L., & Aslin, R. N. (2009). Statistical learning of adjacent and nonadjacent dependencies among nonlinguistic sounds. *Psychonomic Bulletin and Review*, *16*(3), 486-490. http://doi.org/10.3758/PBR.16.3.486

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2014). *Bayesian data analysis* (3rd ed.). Chapman and Hall/CRC.

Gelman, A., & Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical science*, *7*(4), 457-472.

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

Gómez, R. L., & Maye, J. (2005). The developmental trajectory of nonadjacent dependency learning. *Infancy*, *7*, 183-206. http://doi.org/10.1207/s15327078in0702_4

Grama, I. C., Kerkhoff, A., & Wijnen, F. (2016). Gleaning structure from sound: The role of prosodic contrast in learning non-adjacent dependencies. *Journal of Psycholinguistic Research, 45*(6), 1427-1449. http://doi.org/10.1007/s10936-016-9412-8

Heathcote, A., Brown, S., & Cousineau, D. (2004). QMPE: Estimating Lognormal, Wald, and Weibull RT distributions with a parameter-dependent lower bound. *Behavior Research Methods, Instruments, & Computers, 36*(2), 277-290.

Heathcote, A., Popiel, S. J., & Mewhort, D. J. (1991). Analysis of response time distributions: An example using the Stroop task. *Psychological Bulletin, 109*(2), 340-347.

Hoffman, M. D., & Gelman, A. (2014). The No-U-Turn sampler: Adaptively setting path lengths in Hamiltonian Monte Carlo. *Journal of Machine Learning Research*, *15*(1), 1593-1623.

Hull, C. (1943). *Principles of Behavior: An Introduction to Behavior Theory* (p. 278). Oxford, UK: Appleton-Century-Crofts.

Iao, L.-S., Ng, L. Y., Wong, M. Y., & Lee, O. T. (2017). Non-adjacent dependency learning in Cantonese-speaking children with and without a history of specific language impairment. *Journal of Speech, Language, and Hearing Research, 60,* 694-700. http://doi.org/10.1044/2016_JSLHR-L-15-0232

Iao, L.-S., Roeser, J., Justice, L., & Jones, G. (2020, December 29). Concurrent Visual Learning of Adjacent and Nonadjacent Dependencies in Adults and Children. https://doi.org/10.17605/OSF.IO/BCYVR

Janacsek, K., Fiser, J., & Nemeth, D. (2012). The best time to acquire new skills: Age-related differences in implicit sequence learning across the human lifespan. *Developmental Science, 15*, 496-505. http://dx.doi.org/ 10.1111/j.1467-7687.2012.01150.x

Juhasz, D., Nemeth, D., & Janacsek, K. (2019). Is there more room to improve? The lifespan trajectory of procedural learning and its relationship to the between-and within-group differences in average response times. *PloS one*, *14*(7), e0215116.

Kim, R., Seitz, A., Feenstra, H., & Shams, L. (2009). Testing assumptions of statistical learning: Is it long-term and implicit? *Neuroscience Letters, 461*, 145-149. http://doi.org/10.1016/j.neulet.2009.06.030

Kóbor, A., Takács, Á., Kardos, Z., Janacsek, K., Horváth, K., Csépe, V., & Nemeth, D. (2018). ERPs differentiate the sensitivity to statistical probabilities and the learning of sequential structures during procedural learning. *Biological Psychology, 135*, 180-193. https://doi.org/10.1016/j.biopsycho.2018.04.001

Kruschke, J. K. (2010). What to believe: Bayesian methods for data analysis. *Trends in Cognitive Sciences*, *14*(7), 293-300.

Kruschke, J. K. (2011). Bayesian assessment of null values via parameter estimation and model comparison. *Perspectives on Psychological Science*, *6*(3), 299-312.

Kruschke, J. K. (2014). *Doing bayesian data analysis: A tutorial with R, JAGS, and Stan* (2nd ed.). Academic Press.

Kruschke, J. K. (2018). Rejecting or accepting parameter values in Bayesian estimation. *Advances in Methods and Practices in Psychological Science, 1*(2), 270-280.

Kruschke, J. K., Aguinis, H., & Joo, H. (2012). The time has come: Bayesian methods for data analysis in the organizational sciences. *Organizational Research Methods, 15*(4), 722-752.

Lammertink, I., Van Witteloostuijn, M., Boersma, P., Wijnen, F., & Rispens, J. (2019). Auditory statistical learning in children: Novel insights from an online measure. *Applied Psycholinguistics*, *40*(2), 279-302.

Makowski, D., Ben-Shachar, M. S., & Lüdecke, D. (2019). bayestestR: Describing effects and their uncertainty, existence and significance within the Bayesian framework. *Journal of Open Source Software*, *4*(40), 1541.

McElreath, R. (2016). *Statistical rethinking: A bayesian course with examples in R and Stan*. CRC Press.

Meulemans, T., Van der Linden, M., & Perruchet, P. (1998). Implicit Sequence Learning in Children. *Journal of Experimental Child Psychology*, *69*, 199-221. http://doi.org/10.1006/jecp.1998.2442

Meyer, H. C., & Bucci, D. J. (2016). Neural and behavioral mechanisms of proactive and reactive inhibition. *Learning & Memory*, *23*(10), 504-514.

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. http://doi.org/10.3389/fpsyg.2010.00031

Nemeth, D., Janacsek, K., & Fiser, J. (2013). Age-dependent and coordinated shift in performance between implicit and explicit skill learning. *Frontiers in computational neuroscience*, *7*, 147. https://doi.org/10.3389/fncom.2013.00147

Newport, E. L., & Aslin, R. N. (2004). Learning at a distance I. Statistical learning of non-adjacent dependencies. *Cognitive Psychology, 48*, 127-62. http://www.ncbi.nlm.nih.gov/pubmed/14732409

Nissen, M. J., & Bullemer, P. T. (1987). Attentional requirements for learning: Evidence from performance measures. *Cognitive Psychology*, *19*, 1-32.

Onnis, L., Monaghan, P., Richmond, K., & Chater, N. (2005). Phonology impacts segmentation in online speech processing. *Journal of Memory and Language*, *53*(2), 225-237.

Pan, S. C., & Rickard, T. C. (2015). Sleep and motor learning: Is there room for consolidation? *Psychological Bulletin*, *141*, 812–834. http://dx.doi.org/10.1037/bul0000009

Rauch, S. L., Savage, C. R., Brown, H. D., Curran, T., Alpert, N. M., Kendrick, A., ... & Kosslyn, S. M. (1995). A PET investigation of implicit and explicit sequence learning. *Human Brain Mapping*, *3*, 271-286.

Remillard, G. (2008). Implicit learning of second-, third-, and fourth-order adjacent and nonadjacent sequential dependencies. *Quarterly Journal of Experimental Psychology*, *61*(3), 400-424. https://doi.org/10.1080/17470210701210999

Remillard, G. (2010). Implicit learning of fifth-and sixth-order sequential probabilities. *Memory & Cognition*, *38*(7), 905-915. https://doi.org/10.3758/MC.38.7.905

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

Rouder, J. N., & Lu, J. (2005). An introduction to Bayesian hierarchical models with an application in the theory of signal detection. *Psychonomic Bulletin & Review*, *12*(4), 573-604.

Rouder, J. N., Lu, J., Sun, D., Speckman, P., Morey, R., & Naveh-Benjamin, M. (2007). Signal detection models with random participant and item effects. *Psychometrika*, *72*(4), 621.

Saffran, J. R. (2002). Constraints on statistical language learning. *Journal of Memory and Language*, *47,* 172-196.

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

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.

Saffran, J. R., Newport, E. L., Aslin, R. N., Tunick, R. A., & Barrueco, S. (1997). Incidental language learning: Listening (and learning) out of the corner of your ear. *Psychological Science*, *8*(2), 101-105.

Saffran, J. R., & Kirkham, N. Z. (2018). Infant statistical learning. *Annual Review of Psychology*, *69*, 181-203. http://doi.org/10.1146/annurev-psych-122216-011805

Schad, D. J., Vasishth, S., Hohenstein, S., & Kliegl, R. (2020). How to capitalize on a priori contrasts in linear (mixed) models: A tutorial. *Journal of Memory and Language, 110*, 104038.

Shen, Y., Zhao, H., Zhu, J., He, Y., Zhang, X., Liu, S., & Chen, J. (2020). Comparison of Intentional Inhibition and Reactive Inhibition in Adolescents and Adults: An ERP

Study. *Developmental Neuropsychology*, *45*(2), 66-78. https://doi.org/10.1080/87565641.2020.1730376

Simor, P., Zavecz, Z., Horváth, K., Éltető, N., Török, C., Pesthy, O., ... & Nemeth, D. (2019). Deconstructing procedural memory: Different learning trajectories and consolidation of sequence and statistical learning. *Frontiers in Psychology, 9*, 2708. https://doi.org/10.3389/fpsyg.2018.02708

Song, S., Howard, J.H. Jr, & Howard, D.V. (2007). Sleep does not benefit probabilistic motor sequence learning. *Journal of Neuroscience, 27*(46), 12475-12483. https://doi.org/10.1523/JNEUROSCI.2062-07.2007

Sorensen, T., Hohenstein, S., & Vasishth, S. (2016). Bayesian linear mixed models using stan: A tutorial for psychologists, linguists, and cognitive scientists. *Quantitative Methods for Psychology*, *12*(3), 175-200.

Stan Development Team. (2015). *Stan: A C++ library for probability and sampling.* http://mc-stan.org/. (R package version 2.8.0.)

Stark-Inbar, A., Raza, M., Taylor, J. A., & Ivry, R. B. (2017). Individual differences in implicit motor learning: Task specificity in sensorimotor adaptation and sequence learning. *Journal of Neurophysiology, 117(1)*, 412-428. http://doi.org/10.1152/jn.01141.2015

Thomas, K. M., Hunt, R. H., Vizueta, N., Sommer, T., Durston, S., Yang, Y., & Worden, M. S. (2004). Evidence of developmental differences in implicit sequence learning: An fMRI study of children and adults. *Journal of Cognitive Neuroscience, 16*, 1339-1351. https://doi.org/10.1162/0898929042304688

Thomas, K. M., & Nelson, C. A. (2001). Serial reaction time learning in preschool- and school-age children. *Journal of Experimental Child Psychology, 79*, 364-387. http://doi.org/10.1006/jecp.2000.2613

Török, B., Janacsek, K., Nagy, D. G., Orbán, G., & Nemeth, D. (2017). Measuring and filtering reactive inhibition is essential for assessing serial decision making and learning. *Journal of Experimental Psychology: General*, *146*(4), 529-542. https://doi.org/10.1037/xge0000288

Toth-Faber, E., Janacsek, K., Szollosi, A., Keri, S., & Nemeth, D. (2020). Procedural learning under stress: Boosted statistical learning but unaffected sequence learning. *bioRxiv*. https://doi.org/10.1101/2020.05.13.092726

van De Laar, M. C., Van Den Wildenberg, W. P., van Boxtel, G., & van der Molen, M. (2011). Lifespan changes in global and selective stopping and performance adjustments. *Frontiers in Psychology*, *2*, 357. https://doi.org/10.3389/fpsyg.2011.00357

van Witteloostuijn, M., Lammertink, I., Boersma, P., Wijnen, F., & Rispens, J. (2019). Assessing visual statistical learning in early-school-aged children: The usefulness of an online reaction time measure. *Frontiers in Psychology*, *10*, 2051. https://doi.org/10.3389/fpsyg.2019.02051

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

Wagenmakers, E.-J., & Brown, S. (2007). On the linear relation between the mean and the standard deviation of a response time distribution. *Psychological Review, 114*(3), 830-841.

Wagenmakers, E.-J., Lodewyckx, T., Kuriyal, H., & Grasman, R. (2010). Bayesian hypothesis testing for psychologists: A tutorial on the Savage-Dickey method. *Cognitive Psychology*, *60*(3), 158-189.

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. http://doi.org/10.1111/desc.12552

Williams, B. R., Ponesse, J. S., Schachar, R. J., Logan, G. D., & Tannock, R. (1999). Development of inhibitory control across the life span. *Developmental Psychology*, *35*(1), 205-213. https://doi.org/10.1037/0012-1649.35.1.205

Zwart, F. S., Vissers, C. T. W., Kessels, R. P., & Maes, J. H. (2019). Procedural learning across the lifespan: A systematic review with implications for atypical development. *Journal of Neuropsychology*, *13*(2), 149-182. http://doi.org/10.1111/jnp.12139

Appendix

Table A1

Fixed effects summary of the models fitted for the reaction-time data of Experiment 1 and Experiment 2. Shown are the estimated effects for Adjacency (levels: adjacent, nonadjacent), Dependency (levels: dependency, baseline), Block (levels: 1-7), and all by-Block 2-way interactions with Dependency and Adjacency. Effects summarised as the most probable parameter value $\hat{\mu}$ with 95% HPDIs are shown in msecs.

	Experiment 1: Adults			Experiment 2: Children		
Predictor	$\hat{\mu}$	95% HPDI	$P(\hat{\mu} < 0)$	$\hat{\mu}$	95% HPDI	$P(\hat{\mu} < 0)$
Main effects						
Dependency	-101	[-181, -19]	.992	76	[-65, 211]	.148
Adjacency	-43	[-83, -10]	.989	-4	[-84, 77]	.548
Block 1-2	147	[122, 167]	<.001	140	[109, 170]	<.001
Block 2-3	32	[18, 49]	<.001	29	[3, 55]	.015
Block 3-4	33	[18, 50]	<.001	-38	[-61, -12]	.998
Block 4-5	9	[-7, 25]	.151	21	[-5, 47]	.057
Block 5-6	9	[-7, 23]	.168	2	[-23, 31]	.369
Block 6-7	-27	[-46, -4]	.991	-53	[-86, -21]	.998
Interactions						
Dependency * Block 1-2	47	[32, 62]	<.001	48	[23, 72]	<.001
Adjacency * Block 1-2	0	[-16, 14]	.547	-16	[-40, 9]	.89
Dependency * Block 2-3	3	[-14, 18]	.416	3	[-25, 30]	.406
Adjacency * Block 2-3	-6	[-21, 8]	.799	6	[-18, 30]	.323
Dependency * Block 3-4	0	[-18, 15]	.548	-10	[-40, 14]	.807
Adjacency * Block 3-4	9	[-4, 23]	.094	4	[-21, 24]	.453
Dependency * Block 4-5	-1	[-17, 15]	.564	17	[-11, 46]	.102
Adjacency * Block 4-5	-15	[-28, -2]	.984	3	[-18, 26]	.37
Dependency * Block 5-6	1	[-16, 17]	.444	-3	[-33, 23]	.656
Adjacency * Block 5-6	10	[-2, 24]	.052	-4	[-26, 19]	.641
Dependency * Block 6-7	-22	[-47, 4]	.954	-22	[-55, 21]	.836
Adjacency * Block 6-7	0	[-15, 20]	.423	-1	[-26, 30]	.464

Note. $\hat{\mu}$ indicates the most probable *a posteriori* parameter value; 95% HPDI is the range containing 95% of the posterior probability mass; $P(\hat{\mu} < 0)$ is the posterior probability that the true parameter value is smaller than 0; '*' indicates interactions.