**No evidence for domain-general implicit learning deficits in adults with dyslexia**

* Journal: Neuropsychologia?

**Introduction**

Developmental dyslexia is a neurobiological disorder characterized by atypical reading and spelling development. Although dyslexia is most commonly thought to arise from deficits in the phonological system (Melby-Lervåg et al., 2012), domain-general nonlinguistic impairments in dyslexia, specifically in procedural learning, have also been proposed (Nicolson & Fawcett, 2007; 2011; Ullman, 2004; Ullman & Pierpont, 2005). In contrast to the declarative memory system that refers to conscious memory for facts and events, *procedural learning* refers to the implicit (i.e., unconscious) learning of skills, rules, or patterned regularities in stimuli and is often used interchangeably with other terms such as *implicit learning* and *statistical learning* (Berry & Dienes, 1993; Perruchet & Pacton, 2006; Shanks, 2005; Christiansen, 2019). In the current study, we tested whether there are domain-general procedural deficits in development dyslexia.

Early language acquisition is incidental and relies on processing probabilistic information in the speech stream (Saffran, Aslin, & Newport, 1996). Based on transitional regularities, infants develop models of within-category and across-category phonetic variations that are allowed in their language (Panneton & Newman, 2012; Saffran et al., 1996). As children begin to acquire the alphabetic script, they develop the awareness that phonemes are the building blocks for words, and interactively, they acquire the ability to manipulate phonemes and map them to letters (i.e., alphabetic principle; Kolinsky et al., 2021; Ehri, 2005). Although the alphabetic principle must be explicitly taught (Byrne, 2005), implicit learning is also important for reading development because it supports learning orthographic regularities and word meanings (Apfelbaum, Hazeltine, & McMurray, 2013; Chetail, 2017; Pacton, Perruchet, Fayol, & Cleeremans, 2001; ​​Arciuli, Monaghan, & Seva, 2010; Share, 1999, 2004). The procedural system also supports the automatization of reading, resulting in the fluent and subconscious characteristics of proficient reading.

Developmental dyslexia is characterized by inaccurate and dysfluent reading. Although individuals with dyslexia most consistently demonstrate difficulties with storing, retrieving, and/or manipulating phonological representations (Brady et al., 1983; Shankweiler et al., 1979; Snowling, 2000; Stanovich & Siegel, 1994; Vellutino et al., 1994), a range of non-linguistic deficits in the motor and perceptual domains have also been reported for this population (for reviews see Folia et al., 2008; West et al., 2020; Lum et al., 2013; Nicolson & Fawcett, 2011). Multiple theories have been developed to explain the neurocognitive underpinnings of dyslexia. The *procedural deficit hypothesis* proposes that dyslexia is caused by deficits in the procedural memory system, which is critically involved in extracting and learning the sequential or distributional regularities in language (e.g., Nicolson & Fawcett, 2007, 2011; Ullman, 2004; Ullman & Pierpont, 2005). This hypothesis is further supported by the high rates of comorbidity of dyslexia with other disorders (Ramus et al., 2003; Wimmer et al., 1999; Boada et al., 2012), suggesting a broad deficit in learning that extends beyond the language system.

Although a wide range of tasks have been used to measure procedural learning in dyslexia, common across these tasks is the requirement to extract the distributional temporal or spatial properties of sensory input. Learning from this input is indicated by shorter motor response time or higher accuracy of recognition of the learned stimuli. The tasks differ as to whether they implement auditory or visual stimuli, and in the extent to which they rely on purely perceptual or motor learning. A major limitation across studies is only one or two highly related tasks are used to establish the presence of deficit, precluding the conclusion that a broader cross-domain learning deficit is present in a particular sample (West et al., 2020).

The most common procedural task in the dyslexia literature is the serial reaction time task (Nissen and Bullemer, 1987). During this task, participants indicate in which location on the screen a stimulus appears. The locations vary based on probabilistic patterns and faster participant responses indicate learning these patterns. Through this self-paced sequential learning task, previous research has documented impaired performance on serial reaction time tasks in dyslexia (Howard et al., 2006; also see Lum et al., 2013 for meta-analyses). However, this task confounds perceptual learning of stimulus location and the automatization of motor responses, indicating the correct location (Nicolson & Fawcett, 201l; West et al., 2019; see meta-analysis by West et al., 2020).

The two landmark procedural tasks in memory literature that were used to establish the dissociation of the procedural memory system from the declarative memory system in the amnestic patient “H.M.” (Gabrieli, Corkin, Mickel, & Growdon, 1993; Heindel, Salmon, Shults,

Walicke, & Butters, 1989) are mirror tracing (Milner, 1962) and rotary pursuit (Corkin, 1968). These tasks are unique in that unlike the serial reaction time task, they require the development procedural learning without requiring the extraction of sequential regularities. Only one study to date used one of these tasks, mirror tracing, in individuals with dyslexia (Vicari et al., 2005). Italian children with and without dyslexia completed the task and a serial reaction time task. Children with dyslexia were impaired on mirror tracing, as indicated by slower tracing speed. There were no differences in number of errors produced. No other studies have implemented mirror tracing and rotary pursuit with individuals with dyslexia. Additional studies using the two tasks will allow for establishing whether dyslexia is associated with purely procedural and non-sequential learning deficits.

*Statistical learning* *(SL)*, is an implicit learning task that hinges on motor responses to a lesser extent (or not at all). Learners automatically extract the co-occurring patterns of exemplars embedded in the sensory inputs. Increased familiarity for sequences with higher transitional probability versus novel sequences is the indicator of learning. Closely related tasks are probabilistic adaptation tasks in which learning is indicated by perceptual changes in response to stimuli (e.g., selective adpatation; Eimas & Corbit, 1973). There is increasing evidence that individual differences in statistical learning and perceptual adaptation are related to reading performance (Arciuli & Simpson, 2012; Qi et al., 2019; Vandermosten et al., 2018; Arciuli & Simpson, 2012; Frost et al., 2013; Gabay et al., 2015; Spencer et al., 2015; Ozernov-Palchik et al., 2021) and that this relationship is domain specific (Qi et al., 2019; Gabay, Schiff, Vakil, 2012; Ozernov-Palchik et al., 2021). For example, better statistical learning for non-linguistic tones, but not for visual stimuli, was associated with higher reading skills (Qi et al., 2019).

There have been inconsistencies in findings of deficits in statistical learning and perceptual adaptation in individuals with dyslexia. Across several studies, individuals with dyslexia demonstrated reduced adaptation to non-linguistic auditory stimuli such as tones (Ben-Yehudah, Banai, & Ahissar, 2004; Amitay et al., 2002; Gabay et al., 2015, 2019; Gabay, Thiessen, & Holt, 2015; Ahissar et al., 2006). But some studies suggest that these deficits are the result of less reliable perception rather than adaptation deficits (Ozernov-Palchik et al., 2021; Gabay & Holt, 2021). For example, adults with dyslexia demonstrated reduced capacity for tone frequency discrimination but benefitted to the same extent in their performance as typical readers when one comparison tone was held constant across trials (Ozernov-Palchik et al., 2021). In the visual modality, there have been inconsistent findings of learning deficits with dyslexia, with some studies reporting similar learning patterns (van Witteloostuijn et al., 2021; Singh et al., 2018; Nigro et al., 2015; Howard et al., 2006) and others reporting impaired learning in dyslexia (Sigurdardottir et al., 2017; Tong et al., 2019).

The inconsistencies of findings across studies and tasks has highlighted a major drawback in the literature on implicit and procedural learning in developmental dyslexia: the specificity of each study to one particular domain (West et al., 2021). If developmental dyslexia is a result of deficits in the procedural memory system, then deficits across multiple non-linguistic domains are expected, including on purely motor skills. Accordingly, in the current study, we evaluated the performance of adults with dyslexia on four tasks measuring implicit learning in a non-linguistic domain*: rotary pursuit*, *mirror tracing*, *auditory statistical learning*, and *visual statistical learning*.

**Methods**

**Participants**

Twenty-six adults with dyslexia (16 female) and 27 typical readers (14 female) matched on age, sex ratio, and IQ (age 18-41 years, *M* = 26.6, *SD* = 6.3) participated in this study. All participants met eligibility criteria: being a native speaker of American English; born after at least 36 weeks’ gestation; no sensory or perceptual difficulties other than corrected vision; no history of head or brain injury or trauma; no neurological, neuropsychological, or developmental disorder diagnoses; no medications affecting the nervous system; nonverbal IQ standard score > 85 (Matrices subtest of the Kaufman Brief Intelligence Test/KBIT-2; Kaufman & Kaufman, 2004). Hearing tests were completed for all participants and participants with atypical hearing were excluded. The study was approved by the Committee on the Use of Humans as Experimental Subjects (COUHES) at MIT.

**Neuropsychological characterization**

All participants completed a comprehensive battery of standardized reading, language, and cognitive assessments, as well as a background questionnaire (**Table 1**). Measures included: Sight Word Efficiency (SWE) and Phonemic Decoding Efficiency (PDE) subtests of Test of Word Reading Efficiency (Torgesen, Wagner, & Rashotte, 2012); Word ID (WID) and Word Attack (WA) subtests of Woodcock Reading Mastery Test-Revised/Normative Update (Woodcock, 2011); 2-Set subtest of Rapid Automatized Naming (RAN; Wolf & Denckla, 2005); Peabody Picture Vocabulary Test (Vocabulary; Dunn, Dunn, Lenhard, Lenhard, Suggate, 2015); Wechsler Adult Intelligence Scale (Adult-IQ; Wechsler, 2008-for adults); Elision and Nonword Repetition subtests of Comprehensive Test of Phonological Processing (Wagner, Torgesen, & Rashotte, 1999); ​​Digit Span subtest of the Wechsler Adult Intelligence Scale, 4th Edition was used to assess short term memory (WAIS-IV, Wechsler, 2008); and Gray Oral Reading Test-4 (GORT-ORI; Wiederholt and Bryant, 2001). Participants were included in the *dyslexia* group based on performance below the 25th percentile on at least two out of four standardized subtests of timed or untimed word or nonword reading (SWE, PDE, WID, and WA). Participants were included in the *typical reader* group based on performance at or above the 25th percentile on all four of the above subtests. The sample’s demographic information is available in Supplemental Table 3. A majority of participants in the DD group also reported an external diagnosis of dyslexia (*N* = 20) and a history of reading delay (*N* = 19). One participant in each group had a diagnosis of ADHD.

**Table 1**. Summary Behavioral Characterization of Participants

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | DD (N=26) | TYP (N=27) | *p* value | *Cohen’s d* |
| Age | 26.35 (6.59) | 26.68 (6.58) | 0.848 | 0.18 |
| IQ | 107.30 (14.84) | 113.68 (13.97) | 0.106 | 0.28 |
| WID | 89.44 (9.54) | 110.75 (7.40) | < 0.001 | 2.54 |
| WA | 77.59 (9.71) | 102.46 (7.97) | < 0.001 | 2.81 |
| SWE | 89.19 (9.19) | 110.25 (12.50) | < 0.001 | 1.89 |
| PDE | 83.52 (8.19) | 106.54 (7.78) | < 0.001 | 2.87 |
| Vocabulary | 106.15 (10.02) | 113.93 (8.51) | 0.004 | 0.85 |
| ORI | 83.07 (15.40) | 107.11 (8.80) | < 0.001 | 1.91 |
| Elision | 8.22 (2.31) | 9.89 (1.91) | 0.005 | 0.8 |
| Nonword | 6.41 (1.69) | 8.89 (2.10) | < 0.001 | 1.29 |
| RAN | 102.04 (8.46) | 114.82 (8.94) | < 0.001 | 1.44 |
| DigitsForward | 8.41 (2.52) | 11.11 (2.64) | < 0.001 | 1.09 |
| DigitsBackward | 9.04 (2.56) | 10.36 (2.80) | 0.074 | 0.53 |

**Tasks**

***Mirror Tracing***

Participants *(N:* DD = 26, TYP = 27) watched their hands in the mirror while tracing the outline of a six-sided star (Gabrieli et al., 1993; Milner, 1962). A Lafayette Instruments Auto-Scoring Mirror Tracer, a device that includes a metal stylus and a metal test plate (except for the star pattern), was used for this task (Model 58024A⁄C). When the stylus goes off the star and touches the metal plate, it completes an electrical circuit and an error is recorded. Participants were instructed to trace as quickly and accurately as possible, while staying inside the outline of the star. Participants first completed a practice trial and then traced four times. After 30 min of performing other tasks, participants traced five more times. Completion time and number of errors per trial were the dependent variables.

***Rotary Pursuit***

Participants *(N:* DD = 26, TYP = 27) used a Lafayette photoelectric pursuit rotor (Lafayette Instruments, Model 30014C⁄C) to maintain contact with a photoelectric target that rotated in the shape of a rectangle with truncated corners using a stylus. Participants first completed a 20-s practice trial to establish baseline speed (15, 30, 45, or 60 rotations per minute). The speed at which a participant’s time-on-target was closest to 5 sec was selected as the baseline and used for all subsequent trials. Participants then completed four 20-sec trials, took a break for 1 min, and then completed four more 20-sec trials. After 30 min of performing other tasks, participants completed eight more 20-sec trials, taking a 1- min break after the first four trials as before. The dependent measure was the proportion of time-on-target divided by time-off-target per trial.

***Statistical Learning (SL)***

Participants (*N*: DD = 17, TYP = 24) completed one visual SL (VSL) task and one auditory SL (ASL) task, hosted on a secured website (<https://www.cogscigame.co>). The detailed design and the procedure of the SL tasks have been previously described in Qi et al. (2019). For each SL task, a familiarization phase, in which participants performed a target detection task for about 5 minutes, was immediately followed by a test phase, in which a two-alternative forced-choice (2-AFC) test was given.

In the familiarization phase, stimuli were presented in a continuous stream according to an embedded pattern of four unique triplets. In VSL, 12 unique alien images formed four target triplets. Each of the target triplets was repeated 24 times for a total of 96 triplets. Each image was presented one at a time at the center of the screen for 800 ms with 200 ms of inter-stimulus interval (stimulus onset asynchrony (SOA) = 1000 ms), lasting 4 min 48 s. Participants were instructed to press the spacebar as quickly as possible whenever the target alien appeared on the screen. The target alien image was always the third alien of one of the four base triplets so that online learning could be measured via response time acceleration over 24 target trials during exposure. In ASL, 12 unique monotones of the same duration (328 ms) (Tone Task) formed four target triplets. Each triplet was repeated 48 times for a total of 192 triplets. The SOA was 480 ms, with the familiarization phase lasting 4 min and 36 s. Presentation speed was faster in the auditory than visual tasks due to differences in perceptual preference (Conway & Christiansen, 2009; Emberson, Conway, & Christiansen, 2011). The procedure was identical to that of VSL except that the target tones used in the target-detection task during familiarization were constrained to only the lowest and highest notes of the final tones of the four triplets to facilitate identification. Two practice trials before the continuous stream of tones ensured that participants could distinguish the target tone. Response time was measured over 48 target trials. This approach of measuring online learning has been validated in our previous work in adult learners who accelerated more quickly in their responses to target stimuli in structured sequences similar to the ones used here than in random sequences where no triplets were formed and the same stimuli were displayed in a random order (Kozloff et al., 2018; Schneider et al., 2020). RT acceleration was also observed in children using a tablet to respond to target stimuli at the final position of a triplet, but not for target stimuli at the start position of a triplet (Zinszer et al., 2020).

The 2AFC test phase was introduced after the familiarization phase and was composed of 32 questions. For each question, two options were included: a target triplet from the familiarization phase and a foil triplet that was novel to the participant. Foil triplets were constructed so that the relative position of each image in the foil triplet was the same as the target triplet. The test phase consisted of 32 (4 target triplets x 4 foil triplets x 2 repetitions) randomly ordered trials. The images and sounds within each triplet were presented one at a time at the same presentation rate as the familiarization phase with a 1000 ms pause between the target and the foil triplets. Participants were asked to identify which of the two triplets (embedded versus foil) seemed more like what they saw during the familiarization phase. There were no time constraints for responses and no feedback on the accuracy of answers. The order of test trials was randomized across participants.

1 participant in the DDgroup was removed from the ASL analyses and 1 participant in the TYP group was removed from the VSL analyses because their data were not recorded due to technical issues. Task order (VSL or ASL first) was counterbalanced across participants.

**Statistical Analysis**

For all four experiments, outcome variables were analyzed in R v3.5.0 ( R Core Team, 2013), using identical statistical thresholds (*p* < 0.05*)*, and maximal random effect structures (Barr et al., 2013), using the package *lme4* (Bates et al., 2007). The significance of fixed effects in the linear mixed-level models was tested in an ANOVA (using Satterthwaite approximations) and fit with restricted maximum likelihood (REML) using the package *lmerTest* (Kuznetsova et al., 2016). The significance of fixed effects in the generalized linear mixed-level models was estimated using personalized quasi-likelihood using the *glmmPQL* function of the package *MASS* to improve model convergence. The effect sizes for the fixed effects in these linear mixed models were computed via the coefficient of determination (R squared) using *r2beta* function of the package *r2glmm* (Edwards et al., 2008; Nakagawa & Schiezeth, 2013; Jaeger et al., 2016). In all models, Sex, Age, and nonverbal IQ were included as nuisance covariates.

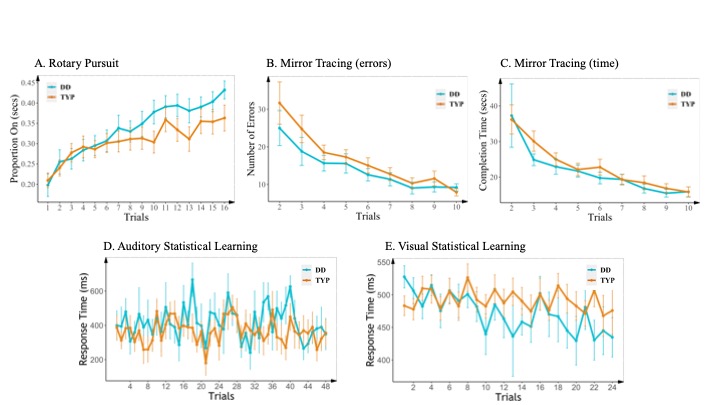
***​​Pairwise Correlation Analysis***

Pearson pairwise correlations were computed to test for cross-task associations using the Hmisc package in R (Harrell Jr & Harrell Jr, 2019). Bayesian correlations were also computed using the BayesFactor package (Morey et al., 2015) with default priors comparing a null model of no correlation with the alternative model of correlation. Bayesian models provide good precision even in smaller data sets (Lee & Song, 2004). Importantly, Bayes factors provide a measure of how likely the data is under the null versus alternative hypothesis, allowing us to quantify and compare relative support for the existence of a relationship between each pair of variables. Based on the previous work, Bayes factors larger than 1 were considered to provide positive evidence (albeit weak if under 3) in favor of the alternative hypothesis that two variables are correlated (Jeffreys, 1998; Wetzels et al., 2011).

**Results**

**Rotary Pursuit**

The group performance by trial is shown in **Figure 1A**. There were no significant group differences in the baseline speed (*t*(45.39) = -0.06, *p* = 0.95, d = 0.02). Both groups showed substantial improvement across trials. To test for group differences in time of target, a linear mixed-effects model was conducted with *proportion on* (computed as time on target/ (time on target + time off target)) as the dependent variable. Fixed factors in the model included *trial number* and *group* (DD vs. TYP); the model’s random effects structure included random intercepts by participants and by-participant random slopes on trial (Jaeger, 2008). The main effect of *trial* was significant (*b* = 0.01, SE = 0.002, *t* = 7.13, *p* < 0.001, 𝑅2𝑚 = 0.072), with an increase in the proportion of time on target across trials for both groups. The main effect of *group* was not significant (*b* = 0.01, SE = 0.04, *t* = 0.30, *p* = 0.77, 𝑅2𝑚 = 0). The interaction between *trial number* and *group* was marginal (*b* = -0.005, SE = 0.002, *t* = -1.86, *p* = 0.069, 𝑅2𝑚 = 0.005), suggesting a marginally steeper slope in the Dyslexic group (*b* = 0.01, SE = 0.001, *t* = 8.37, *p* < 0.001) than the Typical Reader group (*b* = 0.008, SE = 0.002, *t* = 4.15, *p* < 0.001).



**Figure 1**. The time courses of procedural learning and implicit SL during the familiarization phase. Shown is performance for typical readers (TYP, orange) and adults with dyslexia (DD, teal) plotted across trials for rotary pursuit (A), mirror tracing (B, C), the familiarization phase of auditory statistical learning (D), and the familiarization phase of visual statistical learning (E).

**Mirror Tracing**

The completion time by trial and the number of errors by trial are shown in **Figure 1B** and **Figure 1C** respectively. There were no significant group differences in the baseline time (*t*(24.60) = -0.46, *p* = 0.65, d = 0.14) or error (*t*(38.23) = -0.05, *p* = 0.96, d = 0.01) during the first practice trial. Both groups showed substantial reduction in completion time and number of errors across trials. To test for group differences, two linear mixed-effects models were conducted with completion *time* and total number of *errors* as the dependent variables. Fixed effects in the models included *trial number* (Trials 2–10 because the first trial was practice) and *group* (DD vs. TYP); the model’s random effects structure included random intercepts by participants and by-participant random slopes on trial. Age, IQ, and Sex were include as nusisance covariates.

For the *time* model, the main effect of *trial* was significant (*b* = -2.43, SE = 0.68, *t* = -3.59, *p* = 0.001, 𝑅2𝑚 = 0.043), with a reduced time on task across trials. Neither the main effect of *group* (*b* = -3.21, SE = 10.20, *t* = -0.32, *p* = 0.75, 𝑅2𝑚 = 0.001) nor the interaction between *trial* and *group* (*b* = 0.14, SE = 0.94, *t* = 0.15, *p* = 0.88, 𝑅2𝑚 = 0) weresignificant . Similar results were revealed for the *error* model. The main effect of *trial* was significant (*b* = -1.88, SE = 0.43, *t* = -4.32, *p* < 0.001, 𝑅2𝑚 = 0.058), with a reduced number of errors across trials. Neither the main effect of *group* (*b* = 6.98, SE = 5.84, *t* = 1.20, *p* = 0.24, 𝑅2𝑚 = 0.01) nor the interaction between *trial* and *group* (*b* = -0.71, SE = 0.61, *t* = -1.17, *p* = 0.25, 𝑅2𝑚 = 0.004) weresignificant.

**Auditory Statistical Learning**

Familiarization Phase: Valid button presses for the target tones were defined from the onset of the previous stimulus to the onset of the next two stimuli (-480 ms - 960 ms) in order to allow for anticipatory button presses as well as delayed button presses for the target tones. Eight participants (2 in the DD group and 6 in the TYP group) were removed from the analyses due to fewer than 25% of the valid responses to a total of 48 target stimuli presentations. As a result, 14 participants in the DD group and 18 participants in the TYP group are included in the reaction time analyses. The two groups were not significantly different in their baseline response time (*t*(25.9) = 1.52, *p* = 0.14, *d* = 0.57) or hit rate (*t*(30.0) = -1.38, *p* = 0.18, *d* = 0.49; Supplementary Table 1).

The group performance by trial is shown in **Figure 1D**. Neither group showed a trend of acceleration over the course of familiarization. To test for group differences, a linear mixed-effect model was conducted with reaction *time* as the dependent variables. Fixed effects in the models included *trial number* (Trials 1–48), *group* (DD vs. TYP), and their interaction. Age, gender, and non-verbal IQ were included as covariates. The model’s random effects structure included random intercepts by participants and by-participant random slopes on *trial number*. There was no significant effect of *trial number* (*b* = -0.20, SE = 1.48, *t* = -0.14, *p* = 0.89, 𝑅2𝑚 = 0), *group* (*b* = -67.22, SE = 60.88, *t* = -1.10, *p* = 0.28, 𝑅2𝑚 = 0.004), or their interaction (*b* = 0.24, SE = 1.92, *t* = 0.13, *p* = 0.90, 𝑅2𝑚 = 0) on reaction time.

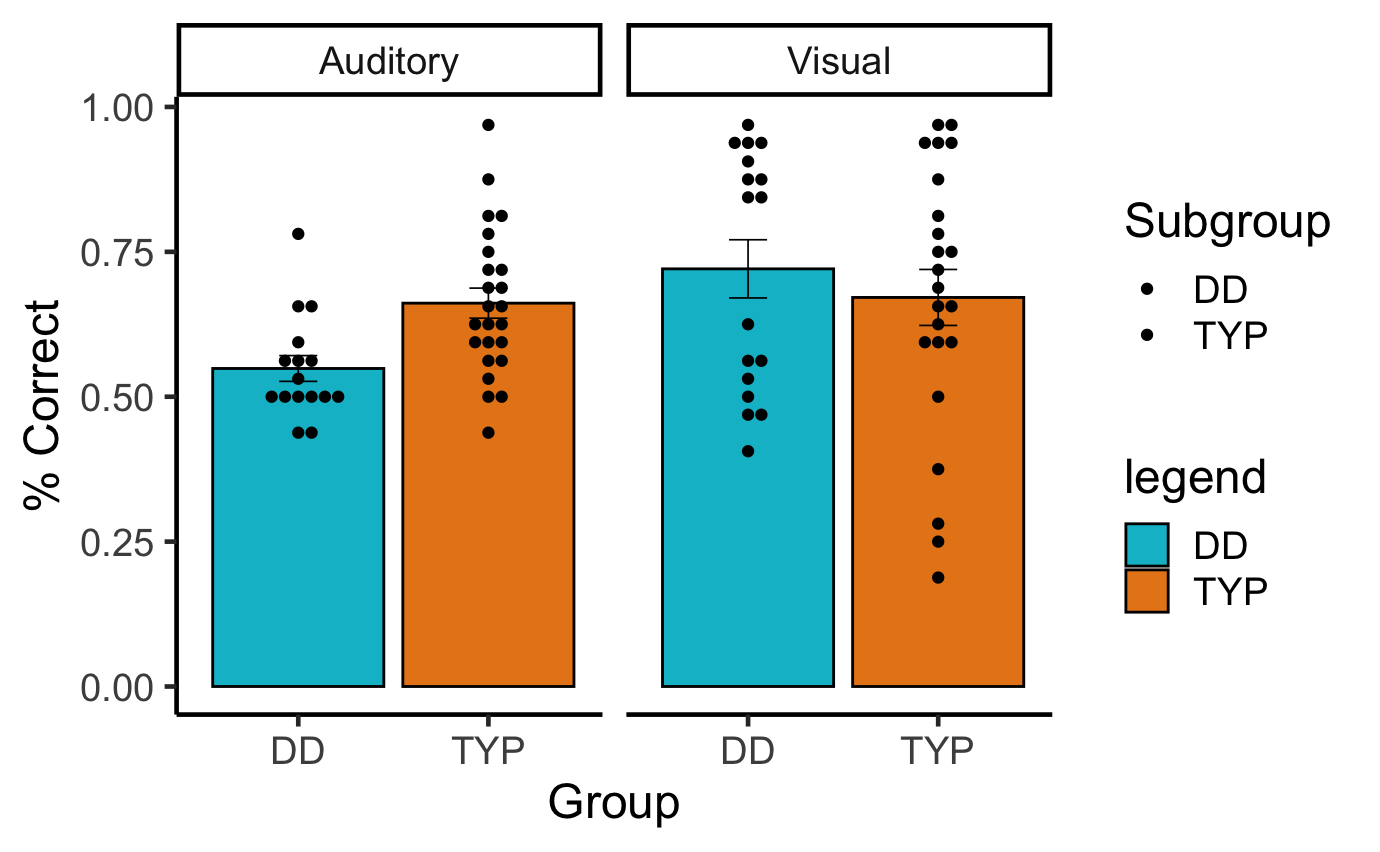
Test Phase: Participants’ overall accuracy in the 32-trial 2AFC test was depicted in **Figure 2**. Both groups performed significantly above 50% chance level (MeanDD = 0.55, SDDD= 0.09, *t*(15) = 2.19, *p* = 0.02; MeanTYP = 0.66, SDTYP= 0.13, *t*(23) = 6.22, *p* < 0.001). We compared the two groups using a generalized linear mixed model fit by maximum likelihood (Laplace Approximation). The dependent variables are each participants’ trial-by-trial binomial accuracy data. The fixed effect included *group* (DD vs. TYP). Age, gender, and non-verbal IQ were included as covariates. The model’s random effects structure included random intercepts by participants and by trial. There was a significant main effect of the group. The Dyslexic group performed significantly worse compared to the Typical Reader group (*b* = 2.43, SE = 0.65, *z* = 3.71, *p* < 0.001, 𝑅2𝑚 = 0.02).

**Visual Statistical Learning**

Familiarization Phase: Valid button presses for the target tones were defined from the onset of the previous stimulus to the onset of the next two stimuli (-1000 ms - 2000 ms) in order to allow for anticipatory button presses as well as delayed button presses for the target tones. Participants performed the target detection task with high hit rates. As a result, all participants were kept in this analysis. The two groups are not significantly different in their baseline response time (*t*(34.5) = -0.70, *p* = 0.49, *d* = 0.23) or hit rate (*t*(21.7) = -0.70, *p* = 0.13, *d* = 0.57; Supplementary Table 1).

The group performance by trial is shown in **Figure 1E**. To test for group differences in response time changes over the course of familiarization, a linear mixed-effect model was conducted with reaction *time* as the dependent variables. Fixed effects in the models included *trial number* (Trials 1–24), *group* (DD vs. TYP), and their interaction. The model’s random effects structure included random intercepts by participants and by-participant random slopes on *trial number*. There was a significant effect of *trial number* (*b* = -2.56, SE = 0.91, *t* = -2.81, *p* = 0.008, 𝑅2𝑚 = 0.01) and a marginal interaction between *group* and *trial number* (*b* = 2.10, SE = 1.19, *t* = 1.76, *p* = 0.087, 𝑅2𝑚 = 0.004) on reaction time. As expected, the *group* difference on reaction time was not significant(*b* = -8.14, SE = 24.39, *t* = -0.33, *p* = 0.74, 𝑅2𝑚 = 0) . Post-hoc within-group analyses suggest only the Dyslexic group showed a significant acceleration over the course of learning (*b* = -2.50, SE = 1.08, *t* = -2.31, *p* = 0.035, 𝑅2𝑚 = 0.02), but the Typical Reader group (*b* = -0.46, SE = 0.65, *t* = -0.71, *p* = 0.49, 𝑅2𝑚 = 0.005) did not show significant acceleration during the familiarization phase.

Test Phase: Participants’ overall accuracy in the 32-trial 2AFC test was depicted in **Figure 2**. Both groups performed significantly above 50% chance level (MeanDD = 0.72, SDDD= 0.21, *t*(16) = 4.39, *p* < 0.001; MeanTYP = 0.67, SDTYP= 0.23, *t*(22) = 3.55, *p* < 0.001). We compared the two groups using a generalized linear mixed model. The dependent variables are each participants’ trial-by-trial binomial accuracy data. The fixed effect included *group* (DD vs. TYP). Age, gender, and non-verbal IQ were included as covariates. The model’s random effects structure included random intercepts by participants and by trial. The two groups were not significantly different on the 2AFC performance (*b* = -0.75, SE = 1.67, *t* = -0.45, *p* = 0.65, 𝑅2𝑚 = 0.001). A generalized linear mixed model comparing the group differences across the two SL tasks suggested a significantly reduced group difference in VSL compared to ASL (*b* = -3.08, SE = 0.86, *t* = -3.59, *p* < 0.001, 𝑅2𝑚 = 0.005).



**\*\*\***

**\*\*\***

**Figure 2**. Implicit SL performance during the test phase. Mean proportion of correct trials in adults with dyslexia (DD, teal) and typical readers (TYP, orange) for auditory statistical learning and visual statistical learning. \*\*\*, *p* < 0.001.

**Supplementary Table 1**: Performance during the SL familiarization phase.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | ASL | | VSL | |
| Mean (SD) | Mean RT (ms) | Hit rate | Mean RT (ms) | Hit rate |
| DD | 408.6 (102.6) | 0.55 (0.20) | 475.4 (70.8) | 0.96 (0.06) |
| TYP | 356.0 (88.76) | 0.66 (0.25) | 491.2 (70.1) | 0.99 (0.03) |

**Supplementary Table 2**: Cross-task correlations

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | 1 | 2 | 3 | 4 | 5 | 6 |
| 1. RP Mean Prop On | 1 |  |  |  |  |  |
| 2. MT Time | -0.067  (BF = 0.346) | 1 |  |  |  |  |
| 3. MT Error | 0.032  (BF = 0..317) | 0.623\*\*\*  BF= 18158.17 | 1 |  |  |  |
| 4. TL Accuracy | -0.014  (BF= 0..366) | -0.052  BF= 0.399 | 0.078  BF= 0.4 | 1 |  |  |
| 5. VL Accuracy | -0.01  BF= 0.365 | -0.045  BF= 0.379 | 0.051  BF= 0.379 | -0.116  BF= 0.443 | 1 |  |
| 6. TL Slope | -0.032  BF= 0.405 | 0.223  BF= 0.434 | -0.083  BF= 0.434 | -0.162  BF= 0.544 | 0.082  BF= 0.428 | 1 |
| 7. VL Slope | 0.177  BF= 0.59 | -0.129  BF= 0.627 | -0.188  BF= 0.628 | 0.072  BF= 0.387 | -0.351\*  BF= 3.106 | -0.182  BF= 0.595 |

**Supplementary Table 3**= Demographic information

|  |  |  |
| --- | --- | --- |
|  | DD (N=26) | TYP (N=27) |
| History of Reading Delay | 19 (73.1%) | 1 (3.7%) |
| History of Language Delay | 2 (7.7%) | 0 (0.0%) |
| Diagnosis Dyslexia | 20 (76.9%) | 0 (0.0%) |
| Diagnosis ADHD | 1 (3.8%) | 1 (3.7%) |
| Race |  |  |
| White | 22 (84.6%) | 24 (88.9%) |
| Black/African American | 4 (15.4%) | 3 (11.1%) |
| Hispanic | 0 (0.0%) | 1 (3.7%) |
| Income |  |  |
| <30k | 6 (24.0%) | 3 (11.5%) |
| 30-60k | 11 (44.0%) | 9 (34.6%) |
| 60-100k | 4 (16.0%) | 3 (11.5%) |
| >100k | 4 (16.0%) | 11 (42.3%) |
| Education |  |  |
| Less7Grd | 0 (0.0%) | 0 (0.0%) |
| Junior High | 0 (0.0%) | 0 (0.0%) |
| High School | 1 (3.8%) | 0 (0.0%) |
| Partial College | 2 (7.7%) | 0 (0.0%) |
| College | 7 (26.9%) | 12 (44.4%) |
| Masters | 14 (53.8%) | 10 (37.0%) |
| Doctorate | 2 (7.7%) | 5 (18.5%) |

**Supplementary Table 4**: Cronbach’s alpha for all tasks.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Procedural Learning** | | | **Statistical Learning** | | | |
|  | Rotary Pursuit | Mirror Tracing - Error | Mirror Tracing - Time | VSL - RT | VSL - Accuracy | ASL - RT | ASL - Accuracy |
| Cronbach’s Alpha | 0.98 | 0.68 | 0.56 | 0.94 | 0.89 | 0.89 | 0.65 |