DOI: 10.1111/cdev.14121

EMPIRICAL ARTICLE

CHILD DEVELOPMENT

Unraveling the complex interplay between statistical learning and working memory in Chinese children with and without dyslexia across different ages

Mei Zhou[1] | Puyuan Zhang[1] | Catherine Mimeau[2] | Shelley Xiuli Tong[1]

[1]Human Communication, Learning, and Development, Faculty of Education, The University of Hong Kong, Hong Kong, China

[2]Department of Human Sciences, Arts and Communication, Université TÉLUQ, Quebec, Canada

Correspondence
Shelley Xiuli Tong, Human Communication, Learning, and Development, Faculty of Education, The University of Hong Kong, Pokfulam, Hong Kong.
Email: xltong@hku.hk

Funding information
Research Fellow Scheme, Grant/Award Number: RFS2021-7H05; General Research Fund, Grant/Award Number: 17612222 and 17616323; Research Grants Council of the Hong Kong Special Administrative Region, China

Abstract
The relation between statistical learning and working memory in children with developmental dyslexia (DD) remains unclear. This study employed a distributional and a conditional statistical learning experiment and a working memory task to examine this relation in 651 Chinese 6- to 12-year-olds with and without DD (N_{DD} = 199, 101 females; N_{woDD} = 452, 227 females; participated 2014–2019). Results showed working memory positively associated with recognizing high-predictable and familiar items in both groups, but negatively associated with recognizing unfamiliar items in the DD group only. These findings uncovered the complex interplay between statistical learning and working memory, demonstrating how different working memory abilities in children with and without DD relate to various statistical learning mechanisms at the item level.

Statistical learning is a basic cognitive process that enables individuals to detect and retain recurring distributional (e.g., frequency) and conditional (e.g., co-occurrence probability) patterns in their environment (Saffran, 2002; Thiessen, 2017). This process plays a critical role in reading acquisition, especially for children with developmental dyslexia (DD; for a review, see Lee et al., 2022), a condition characterized by persistent difficulty in learning to read despite adequate intelligence and educational opportunity (Lyon et al., 2003). An increasing number of studies have shown that children with DD exhibit deficits in learning distributional (e.g., Tong et al., 2020) and conditional (e.g., Gabay et al., 2015) regularities compared to their typically developing (TD) peers. Furthermore, statistical learning is closely related to working memory, with greater working memory capacity expanding the processing "window" for individual input, thereby enabling more efficient statistical learning processes (Janacsek & Nemeth, 2013). However, it remains unclear whether the statistical learning deficits in children with DD relate to their insufficient (Gray et al., 2019) but age-increasing working memory capacities (Alloway & Alloway, 2013). To address this question, we explored the relation between working memory and statistical learning by examining the extent to which working memory is associated with statistical learning in Chinese children with and without DD across different ages and item characteristics in distributional and conditional statistical learning tasks.

Abbreviations: 2AFC, two-alternative forced-choice; DD, developmental dyslexia; EEG, electroencephalography; SES, socioeconomic status; TD, typically developing.

Distributional statistical information, which describes the frequency and the central tendency of patterns, is particularly evident in Chinese character orthography. One typical example is the predictability of a radical's position in a character, which varies along a continuum (e.g., He & Tong, 2017; Tong et al., 2020). Some radicals consistently appear in a specific position, such as the radical 扌, which always occurs on the left in all left–right-structured characters (e.g., 扣, 扛, 抗, 扔, 抄, 扶, 托, 捫, 揚, 擴, 掃, 抑), resulting in a 100% predictability of its primary position. In contrast, other radicals exhibit variability in their position. For example, the radical 車 appears on the left in 38 of 65 characters (e.g., 軋, 軌, 軒, 斬, 軛), but it can also occur on the right (e.g., 陣, 揮, 渾, 暉, 琿), top (e.g., 擊, 輿), and bottom (e.g., 葷, 載), resulting in a 58.5% predictability of its primary (i.e., left) position (Tseng et al., 2018). In previous studies, an artificial orthography learning paradigm was utilized to simulate the quasi-regularities observed in Chinese radicals. This paradigm involved manipulating the occurrence of a radical in its primary position with varying probabilities (high: 100%, moderate: 75%, and low: 50%) among a set of pseudocharacters. The findings from these studies demonstrate that children exhibit sensitivity to these distributional statistics, as evidenced by greater recognition rates for high-predictable pseudocharacters compared to moderate-predictable or low-predictable ones (e.g., He & Tong, 2017; Tong et al., 2020).

In contrast, conditional statistical information refers to the co-occurrence probabilities of subsequent elements, or the transitional probability of a second element (e.g., English syllable /py/) given the occurrence of a first element (e.g., English syllable /hap/, as in "happy"; Saffran et al., 1996). A widely used approach to studying conditional statistics is the visual triplet learning paradigm (Arciuli & Simpson, 2011, 2012). In this paradigm, learners are exposed to a continuous stream of recurring visual triplets, where the transitional probabilities of two consecutive stimuli within and between triplets are 100% and 33%, respectively. Successful learning occurs when learners discriminate the triplets from foils in a subsequent two-alternative forced-choice (2AFC) task.

Extensive research has been conducted on both distributional and conditional statistical learning in individuals with DD, yielding a wide array of findings. In terms of distributional statistical learning, Lee and Tong (2020) found that Chinese children with DD made more radical position errors in a Chinese word dictation task than their TD peers, indicating impaired learning of radical positional distributions. Moreover, Tong et al. (2020) showed that, although children with DD's overall performance was comparable to their TD peers in learning positional regularity of artificial pseudocharacters, they exhibited lower accuracy in recognizing high- and low-predictable radical positions in left–right-structured pseudocharacters compared to top–bottom ones, suggesting item-specific difficulties. These findings highlight the complexity of statistical learning abilities in individuals with DD and the importance of considering task-specific factors.

Similarly, the results of conditional statistical learning in individuals with DD have been mixed. Some studies have shown impaired learning in individuals with DD (e.g., Sigurdardottir et al., 2017), while others have reported intact learning (e.g., van Witteloostuijn et al., 2019). Furthermore, deficits in DD have been observed in specific conditions, such as a high topological entropy condition of the artificial grammar paradigm (Schiff et al., 2017) and the delayed recall session of a serial reaction time task (Ballan et al., 2023). These findings suggest significant variations in the statistical learning abilities of individuals with DD across different contexts, potentially influenced by other cognitive functions, particularly working memory.

Working memory capacity, or the ability to temporally store and process a maximum number of items (Cowan, 1998), is posited to influence both distributional and conditional statistical learning. According to the extraction and integration framework, working memory supports conditional statistical learning by chunking co-occurrent elements (Erickson & Thiessen, 2015). Additionally, Conway (2020) argued that working memory and selective attention to certain stimuli or characteristics enable individuals to learn temporal sequences. This notion was supported by studies that found a positive association between working memory and conditional statistical learning performance when manipulating transitional statistics (e.g., Jongbloed-Pereboom et al., 2017), although some studies have reported no correlation (e.g., Frost et al., 2013; Kalra et al., 2019). Furthermore, recent research suggests that working memory is also associated with distributional statistical learning of the positions of elements (Saito et al., 2020). Based on these theories, the present study tested the connection between working memory capacity and statistical learning of distributional and conditional statistics by employing an artificial orthography learning experiment (Figure 1a) and a visual triplet learning experiment (Figure 1b), respectively, in children with and without DD.

Two hypothetical models have identified item characteristics as a potential factor influencing the relation between working memory and statistical learning (Conway, 2020; Lee et al., 2022). Conway (2020) suggested that complex or hierarchical structure requires more explicit processing, which relies on working memory. Similarly, in the statistical learning and reading model, Lee et al. (2022) postulated that the procedural and declarative memory systems interact dynamically, with working memory functioning as a buffer system regulated by input characteristics such as predictability. Predictability refers to an item's or a feature's likelihood of predicting the occurrence of another. Studies on artificial orthography learning have shown

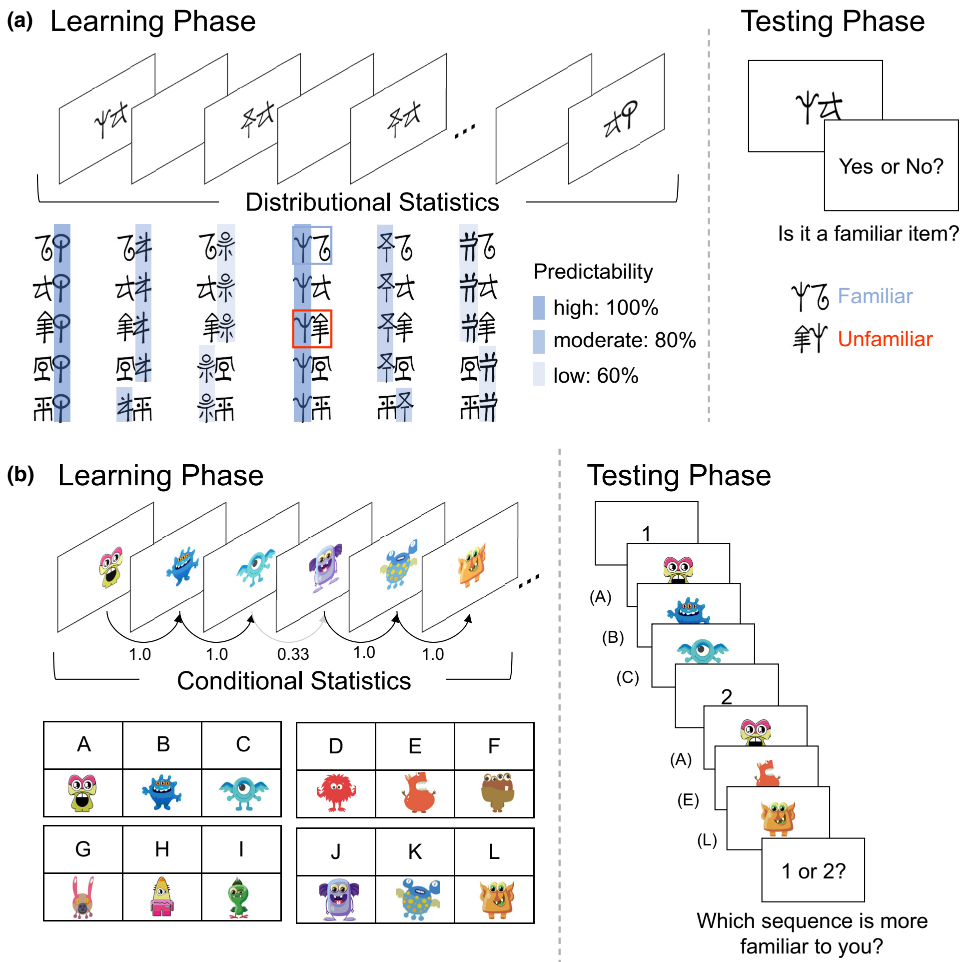

FIGURE 1 Design and procedure of the two statistical learning experiments. (a) The distributional artificial orthography learning experiment exposed children to a series of pseudocharacters carrying high (100%), moderate (80%), and low (60%) levels of positional predictability in the learning phase; children were shown familiar and unfamiliar pseudocharacters and asked to determine whether the pseudocharacter appeared in the learning phase. (b) The conditional visual triplet learning experiment presented children with a series of cartoon aliens that contained four fixed triplets. Following the learning phase, children had to choose the familiar sequence from a choice of two.

that low-predictable pseudocharacters carrying low-predictable target radicals create a higher working memory load compared to high-predictable ones. This is supported by an electroencephalography (EEG) study showing that a larger P300 component was elicited by low-predictable pseudocharacters, indicating increased cognitive demands (Tong et al., 2023). Thus, in our artificial orthography learning experiment, we expected a stronger relation between statistical learning and working memory on low-predictable items compared to high-predictable items.

Another item characteristic, familiarity, which refers to the similarity between a new input and prior experiences, is also a factor that could potentially influence the association between statistical learning and working memory. For instance, in an artificial grammar learning study, imposing a working memory load impaired the expression of the learned grammar when the items (e.g., XXVXJJ) were tested in a novel context (e.g., FFNFCC), highlighting the need for working memory resources when processing unfamiliar information (Hendricks et al., 2013). In our artificial orthography learning experiment, we examined the impact of familiarity on the association between statistical learning and working memory by comparing the recognition of familiar and unfamiliar items. Furthermore, since previous studies using auditory triplet learning experiments found a better recognition when a familiar triplet rather than an unfamiliar one was presented first during the 2AFC test (Raviv & Arnon, 2018; Shufaniya & Arnon, 2018), we investigated the impact of familiarity in our visual triplet learning experiment by comparing the recognition of items in two presenting orders: familiar–unfamiliar and unfamiliar–familiar.

In addition to item characteristics such as predictability and familiarity, another factor that may contribute to the mixed results between working memory and statistical learning in individuals with DD is their adoption of an exploration strategy. This strategy involves allocating more resources to unknown information, in contrast to an exploitation strategy that prioritizes the application of acquired knowledge (Taylor & Vestergaard, 2022; Zhang et al., 2023). Together, both strategies form a dual mechanism that manages the utilization of limited resources, balancing the search for new information (i.e., exploration) and the application of established knowledge (i.e., exploitation; Frankenhuis & Gopnik, 2023; Gopnik, 2020). Due to their low working memory capacity, individuals with DD struggle to effectively utilize (exploit) local cues to retrieve information from long-term memory. Instead, they adopt an exploration strategy by seeking relevance and connections among items (Taylor & Vestergaard, 2022). For instance, an EEG study indicated that individuals with DD allocated attentional resources not only to high-predictable items but also to low-predictable items, suggesting a broader focus compared to TD individuals (Singh et al., 2018). Based on this theoretical and empirical evidence, we hypothesized that children with DD would allocate more resources to unknown information (e.g., low-predictable and unfamiliar items) relative to acquired knowledge (e.g., high-predictable and familiar items).

Moreover, the influence of age on the relation between statistical learning and working memory in children with DD remains unexplored. Statistical learning is recognized as a multi-faceted construct that relies on implicit and explicit mechanisms (e.g., Batterink et al., 2015; Weinberger & Green, 2022), which change throughout the lifespan to optimize learning at different developmental stages (e.g., Bo et al., 2012; Daltrozzo & Conway, 2014; but see Siegelman & Frost, 2015). Notably, statistical learning occurs in infancy when working memory ability is still immature (Emberson et al., 2019; Kirkham et al., 2002; Saffran et al., 1996), suggesting that it may rely more on bottom-up mechanisms rather than working memory at an early age. As the attention-dependent system of working memory develops, statistical learning may increasingly involve more explicit memory resources and top-down mechanisms. This is supported by evidence showing that age-related declines in working memory are associated with a reduction in chunk length during sequence learning (Bo et al., 2009).

As indicated by the exploration and exploitation mechanisms, the item-regulated association between statistical learning and working memory may also vary with age. As high working memory capacity facilitates the employment of an exploitation strategy, age-increased working memory may lead to a developmental shift from an exploration to an exploitation strategy (Frankenhuis & Gopnik, 2023; Taylor & Vestergaard, 2022). This means that children may allocate more working memory resources to high-predictable and familiar items than to low-predictable and unfamiliar items as they grow older. However, this pattern may be less apparent in children with DD due to their preference for an exploration strategy and their consistently lower working memory. To test this assumption, we investigated the allocation of working memory resources to specific items in children with DD from 6 to 12 years old, a period of steady development of working memory capacities.

In sum, this cross-sectional study addressed two research questions by conducting two statistical learning experiments on and administering one working memory task to 6- to 12-year-old Chinese children with DD and their TD peers. The first question investigated whether children with DD exhibited a different relation between statistical learning and working memory than their TD peers. This relation was assessed by examining the predictive effect of working memory capacity on statistical learning using generalized linear mixed models (confirmatory analyses). Given the lower working memory capacity of children with DD and its potential role in their statistical learning deficits (Lee et al., 2022), we expected a stronger effect of working memory in children with DD compared to their TD peers.

Our second research question examined the strength of the association between statistical learning and working memory among children with DD compared to their TD peers, specifically in relation to item characteristics (i.e., predictability and familiarity) and age. This was assessed by analyzing the possible interaction effects of working memory with item characteristics, age, and reading status (confirmatory analyses). Based on previous evidence indicating the role of working memory in statistical learning under different contexts (Daltrozzo & Conway, 2014; Hendricks et al., 2013; Tong et al., 2023), we predicted for both groups a stronger association between working memory and statistical learning for (1) low-predictable items compared to high-predictable items, and unfamiliar items compared to familiar items, in the recognition task of the distributional statistical learning experiment; and (2) unfamiliar–familiar items compared to familiar–unfamiliar items in the 2AFC task of the conditional statistical learning experiment. Furthermore, as age increased, we hypothesized that these interaction effects would decrease. Similar effects of item characteristics would be observed for children with DD, but without an age-related decrease.

METHOD

Participants

Participants were 651 native Cantonese-speaking 6- to 12-year-old children in Grades 1 to 6: 199 children with DD (M_{age} = 9 years 0 months; 101 females, 98 males) and 452 TD children (M_{age} = 8 years 11 months; 227 females, 225 males). All children studied in Hong Kong mainstream primary schools and learned to read in Chinese and English simultaneously. Written informed consent was obtained from parents or caregivers, and the children received a compensation fee for participating. The data collection process was conducted from 2014 to 2019. This study was approved by the Human Research Ethics Committee at the corresponding author's university.

The inclusion criteria for the DD group were that children (1) were formally diagnosed by a clinical or educational psychologist, and (2) did not have comorbidity with another learning difficulty or disorder, such as attention deficit hyperactivity disorder, autistic spectrum disorder, hearing loss, or developmental language disorder. During the time of this study, Hong Kong's diagnosis of Chinese children with dyslexia was based on three literacy tests (i.e., Chinese character recognition, one-minute word reading, and word dictation) and four cognitive tests (i.e., rapid digit naming, phonological awareness, phonological memory, and orthographic knowledge) from the Hong Kong Test of Specific Learning Difficulties in Reading and Writing for Primary School Students–second edition (Ho et al., 2007). Students were identified as having dyslexia when they had an IQ of 85 or above but scored one standard deviation below average for both the literacy composite score and at least one of the cognitive composite scores. Our exclusion criteria for TD children involved a history of learning or developmental disorders, significant reading difficulties, cognitive impairments, non-fluency in Chinese, and significant physical or sensory impairments.

Participants' nonverbal intelligence was assessed using the 30-item standardized matrix reasoning test from the Wechsler Abbreviated Scale of Intelligence–second edition (Wechsler, 2011). Socioeconomic status (SES) was measured using parent-reported family income per month on a 6-point scale (1 = less than 10,000; 2 = 10,000–20,000; 3 = 20,000–30,000; 4 = 30,000–40,000; 5 = 40,000–50,000; and 6 = more than 50,000 Hong Kong dollars) and parents' education level on an 8-point scale (1 = no primary school education; 2 = some primary school education; 3 = primary school graduate; 4 = middle school graduate; 5 = high school graduate; 6 = associate degree; 7 = university or college degree; and 8 = graduate or professional degree). Based on the reported data from 174 parents of DD children and 425 parents of TD children, the mean SES (an average score of family income and parents' education level) was lower for children with DD (M = 5.19, SD = 1.21) than for their TD peers (M = 5.48, SD = 1.21), p = .008. However, the effect size was small (Cohen's d = 0.24), indicating the negligible difference in SES between the two groups.

Artificial orthography learning experiment

Stimuli

The learning stimuli consisted of 60 pseudocharacters, divided into 30 left–right-structured and 30 top–bottom-structured sets. These pseudocharacters were created by combining six target radicals and five control radicals from pictographic Dongba and syllabic Geba scripts that were used by the Naxi minority in western China during the 17th century and were unfamiliar to the participants. Each pseudocharacter comprised one target radical and one control radical. Children were randomly assigned to either the left–right or the top–bottom structure condition.

Within each structure (left–right or top–bottom), the target radical was at a dominant position (half were on the left or top and the other half on the right or bottom) with high (100%), moderate (80%), or low (60%) probabilities in five, four, and three pseudocharacters among all five pseudocharacters containing that target radical, respectively. Each control radical occurred equally at both positions across the 30 pseudocharacters.

Thirty testing pseudocharacters for each structure were generated for the recognition test. Half (i.e., 15) were familiar items selected from the 30 learning pseudocharacters, and the other half were unfamiliar items created by switching the positions of the two radicals of the learning items. As a result, participants were tested on two levels of familiarity for the items—familiar and unfamiliar. Furthermore, testing items were equally selected from the three predictability levels (i.e., high, moderate, and low).

Procedure

Children completed consecutive learning and testing phases (Figure 1a). During the learning phase, children were shown pseudocharacters one at a time, with each one displayed for 1,200 ms with a 200 ms interstimulus interval. Each pseudocharacter was repeated 20 times, resulting in 600 trials. Children were not informed of any patterns or regularities in the stimuli and performed a cover task by pressing the SPACEBAR as quickly as possible when a pseudocharacter appeared twice in a row. In the testing phase, they identified whether a pseudocharacter was previously shown in the learning phase by pressing the "yes" or "no" key. The presentation order was pseudorandomized, and there was no response time limit.

Visual triplet learning experiment

Stimuli

Employing a modified visual triplet learning paradigm (Arciuli & Simpson, 2012), 18 novel cartoon figures of colored aliens that did not resemble any real objects were created to eliminate verbal memorization and familiarity effects. Twelve aliens (A-L) constituted four triplets (ABC, DEF, GHI, and JKL), and the other six aliens were used for practice trials. Four new triplets

(AEI, DHL, GKC, and JBF) were added as foils in the 2AFC test.

Procedure

The experiment was divided into two phases, a learning phase and a 2AFC test (Figure 1b). During the learning phase, children were shown a series of aliens, each appearing for 400 ms with a 200 ms interstimulus interval. A triplet consisted of three aliens presented in succession, with the order of the triplets randomized. This led to a high transitional probability within triplets (100%) and a low transitional probability between them (33%). Each triplet was repeated 24 times, resulting in 96 triplets. In six repetitions of each triplet, one alien was displayed twice consecutively. Children performed a cover task by pressing a designated key to detect each recurring alien. During the 2AFC test, children were presented with both a real triplet and a foil triplet and were asked to choose the one that was more familiar to them. The 2AFC test consisted of 64 randomized test trials, following two practice trials. The presentation times and interstimulus intervals were consistent with the learning phase.

Working memory task

Working memory capacity was assessed using an 18-trial backward digit span task. Children repeated an orally presented series of digits in reverse order. Every two trials, the series ascended in length from 1 to 9 digits. The task was terminated when children failed both trials of the same level. The highest digit length achieved represented their working memory capacity.

General procedure

Participants were individually tested by a native Cantonese-speaking experimenter in a sound-attenuated University laboratory after obtaining written informed consent from their parents or caregivers. Both the artificial orthography learning experiment and the visual triplet learning experiment were conducted using E-prime 2.0 (Schneider et al., 2002). The experimenter verbally explained the instructions while simultaneously displaying them on the computer screen. The order of these two experiments was counterbalanced across participants. Prior to the statistical learning experiments, the working memory task was administered along with other pencil-and-paper measures of Chinese literacy skills, which are not reported here. Participants completed the tasks over two sessions, each lasting approximately 1 h with breaks in between.

Data analytical approach

We performed generalized linear mixed models using lme4 (Version 1.1-30) and lmerTest (Version 3.1-3) packages in R (Version 4.2.1, R Core Team, https://www.r-project.org/). Continuous variables, including working memory capacity, age, and nonverbal intelligence, were centered to reduce multicollinearity and make coefficients more interpretable (Kraemer & Blasey, 2004). Categorical variables were coded using a simple contrast coding approach with the contr.simple function from the YawMMF package (Zhang, 2022) to obtain reliable estimates of the main effects from interaction terms. To probe interactions, simple effects were evaluated by post hoc analyses using the emmeans package (Version 1.8.5). The significance threshold of all tests was $p < .05$.

To address the question of how item characteristics and age influence the relation between statistical learning and working memory in children with and without DD, one generalized linear mixed model for distributional statistical learning and one for conditional statistical learning were conducted, with responses in a given trial as the dependent variable, and children and stimuli as random intercepts. Both models evaluated the main effects of working memory capacity, item characteristics, age, and reading status (contrast coding: −1/2 for TD and 1/2 for DD), as well as interactions involving working memory capacity. Children's nonverbal intelligence and sex (contrast coding: −1/2 for male and 1/2 for female) were included as covariates.

In the distributional statistical learning model, the pseudocharacter structure (contrast coding: −1/2 for top–bottom and 1/2 for left–right) was included as an additional covariate. The item characteristics included predictability (contrast coding: for high predictable vs. moderate predictable: −1/3 for high predictable, 2/3 for moderate predictable, and −1/3 for low predictable; for high predictable vs. low predictable: −1/3 for high predictable, −1/3 for moderate predictable, and 2/3 for low predictable) and familiarity (contrast coding: −1/2 for unfamiliar and 1/2 for familiar). In the conditional statistical learning model, familiarity (contrast coding: −1/2 for unfamiliar–familiar and 1/2 for familiar–unfamiliar) was the only item characteristic. The sample size of this model was 645 after excluding 6 children who had missing data.

RESULTS

Descriptive statistics

Our initial analyses aimed to verify the literacy profile of the two groups. As displayed in Table 1, children with DD showed significantly lower scores than their

TABLE 1 Literacy and cognitive performances of children with developmental dyslexia (DD) and their typically developing (TD) peers.

Measure (maximum)	Mean (SD) DD	Mean (SD) TD	*t*	*p*(>|*t*|)	Cohen's *d*
Sex (female/male)	101/98	227/225			
Age in years	8.96 (1.46)	8.88 (1.56)	−0.64	.523	−0.05
Nonverbal intelligence					
Matrix reasoning (30)	15.91 (4.74)	17.74 (4.64)	4.60	<.001***	0.39
Chinese word reading (150)	66.33 (35.91)	103.60 (30.66)	12.74	<.001***	1.12
Chinese word reading fluency	47.76 (17.94)	69.30 (17.42)	14.40	<.001***	1.22
Chinese word dictation (96)	30.84 (16.94)	56.17 (22.38)	15.86	<.001***	1.28
Working memory (18)	4.80 (2.35)	6.42 (3.22)	7.21	<.001***	0.58
Distributional statistical learning					
Artificial orthography learning (%)	0.58 (0.09)	0.58 (0.09)	0.16	.876	0.01
Predictability					
High	0.63 (0.15)	0.64 (0.16)	1.27	.204	0.11
Moderate	0.56 (0.13)	0.55 (0.14)	−0.60	.551	−0.05
Low	0.57 (0.15)	0.56 (0.14)	−0.51	.610	−0.04
Familiarity					
Familiar	0.74 (0.19)	0.77 (0.17)	1.89	.059	0.16
Unfamiliar	0.43 (0.19)	0.40 (0.18)	−1.64	.102	−0.14
Conditional statistical learning					
Visual triplet learning (%)	0.52 (0.09)	0.54 (0.10)	2.12	.035*	0.18
Familiarity					
Familiar–unfamiliar	0.51 (0.13)	0.55 (0.13)	2.79	.005**	0.24
Unfamiliar–familiar	0.54 (0.13)	0.54 (0.13)	0.31	.757	0.03

*$p<.05$; **$p<.01$; ***$p<.001$.

age-matched TD peers in Chinese literacy tasks, including word reading accuracy and fluency, as well as word dictation ($ps<.001$).

In terms of cognitive profile, children with DD showed lower working memory capacity ($p<.001$) and lower performance in conditional statistical learning ($p=.035$) compared to TD children, but no difference in distributional statistical learning ($p=.876$). However, both DD and TD groups performed above chance (50%) for distributional ($t_{DD}=12.94$, $p<.001$; $t_{TD}=19.54$, $p<.001$) and conditional statistical learning ($t_{DD}=3.64$, $p<.001$; $t_{TD}=8.52$, $p<.001$). Specifically, 75.9% of children with DD and 75.0% of TD children demonstrated successful distributional statistical learning, while 48.2% of children with DD and 60.0% of TD children showed successful conditional statistical learning. In terms of performance across item characteristics, children with DD showed equivalent performance as TD children across item predictability and familiarity conditions ($ps>.05$) during distributional statistical learning, but they showed significantly lower accuracy than TD children for familiar–unfamiliar items during conditional statistical learning ($p=.005$).

Distributional statistical learning: Working memory is positively associated with the recognition of high-predictable and familiar items regardless of reading status, but negatively associated with the recognition of unfamiliar items in older children with DD only

To examine the influence of item characteristics, age, and reading status on children's distributional statistical learning performance in the artificial orthography learning experiment, a generalized linear mixed model was conducted. The coefficient estimates are displayed in Figure 2a (see Table S1 for details). For the covariate variables, nonverbal intelligence ($\beta=0.06$, 95% CI [0.02, 0.10], $z=3.10$, $p=.002$) and the pseudocharacter structure ($\beta=-0.28$, 95% CI [−0.34, −0.21], $z=-8.32$, $p<.001$) but not sex ($p=.709$) yielded significant effects. Compared to high-predictable items, children showed lower performance on moderate-predictable ($\beta=-0.49$, 95% CI [−0.62, −0.36], $z=-7.40$, $p<.001$) and low-predictable items ($\beta=-0.34$, 95% CI [−0.46, −0.21], $z=-5.23$, $p<.001$). They also showed higher accuracy in recognizing familiar compared to unfamiliar items ($\beta=1.71$, 95% CI [1.60, 1.82], $z=31.05$, $p<.001$). No significant main effects were

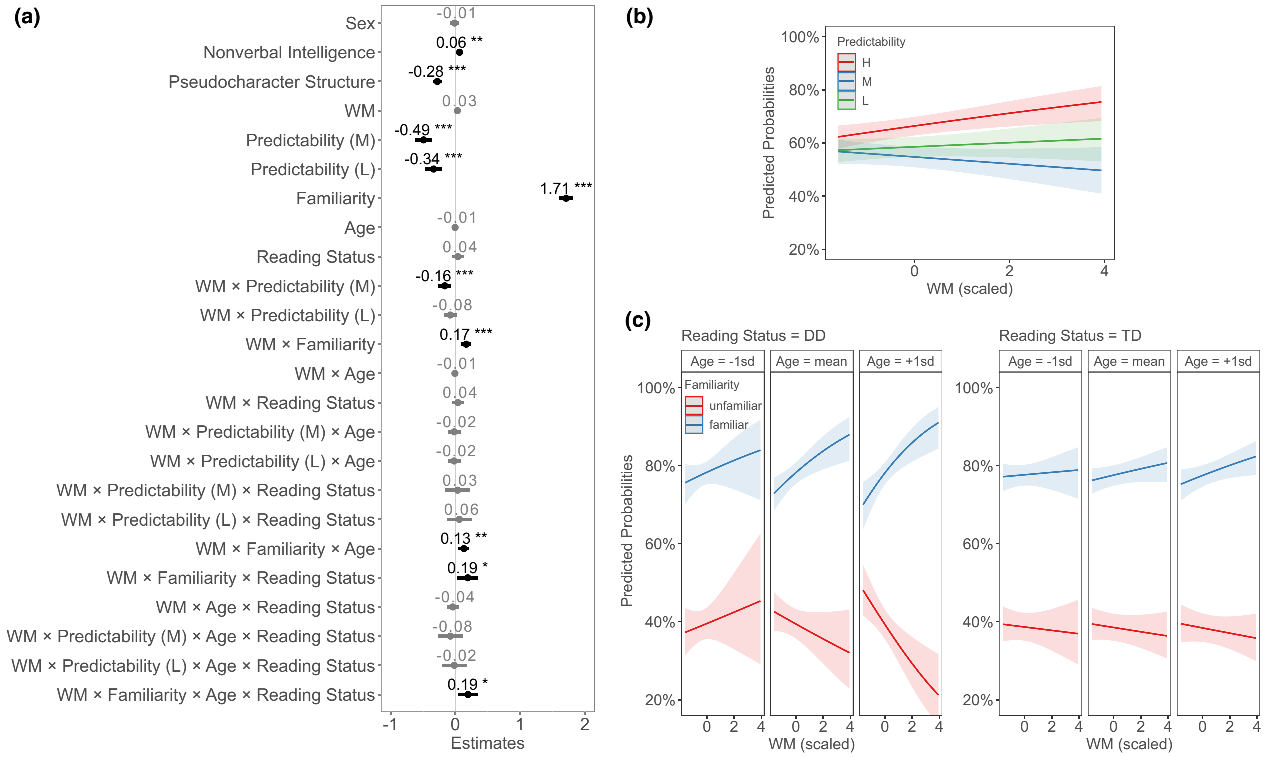

FIGURE 2 Influences of item characteristics, age, and reading status (DD: children with developmental dyslexia; TD: typically developing children) on the relation between distributional statistical learning and working memory (WM). Panel (a) displays the coefficient estimates of the generalized linear mixed model. Panels (b, c) show that the associations between working memory and statistical learning varied under high (H)-, moderate (M)- and low (L)-predictable items (b), and familiar and unfamiliar items in the DD and TD groups across different ages (c). The error bars in panel (a) and the shade areas in panels (b, c) indicate 95% confidence intervals. *$p<.05$; **$p<.01$; ***$p<.001$.

found for working memory, age, and reading status ($ps>.05$).

Item predictability modulated the effect of working memory on statistical learning (Figure 2b), with high-predictable items having a stronger effect than moderate-predictable items ($\beta=-0.16$, 95% CI [−0.26, −0.07], $z=-3.30$, $p<.001$) but a comparable effect with low-predictable items ($p=.106$). However, no significant evidence showed reading status or age modulated this effect ($ps>.05$).

Familiarity influenced the relation between statistical learning and working memory in a more complex way. A two-way interaction between working memory and familiarity ($\beta=0.17$, 95% CI [0.09, 0.25], $z=4.10$, $p<.001$) was significant, indicating a stronger effect of working memory on recognizing familiar items compared to unfamiliar items. This interaction was further modulated by age ($\beta=0.13$, 95% CI [0.05, 0.21], $z=3.14$, $p=.002$) and reading status ($\beta=0.19$, 95% CI [0.03, 0.35], $z=2.34$, $p=.019$) in two three-way interactions, such that the difference between the effect of working memory on familiar and unfamiliar items was larger at an older age relative to a younger age and in children with DD relative to TD children. Accordingly, a significant four-way interaction ($\beta=0.19$, 95% CI [0.04, 0.35], $z=2.44$, $p=.015$) indicated that the relation between working memory and statistical learning was modulated by item familiarity, age, and reading status (Figure 2c).

The post hoc analysis on simple effects of this four-way interaction further revealed that children with DD's working memory showed an age-increased relation with statistical learning for familiar items ($\beta_{\text{age: mean-1sd}}=0.09$, 95% CI [−0.08, 0.27]; $\beta_{\text{age: mean}}=0.18$, 95% CI [0.06, 0.30]; $\beta_{\text{age: mean+1sd}}=0.27$, 95% CI [0.11, 0.42]), but an age-decreased relation with statistical learning for unfamiliar items ($\beta_{\text{age: mean-1sd}}=0.06$, 95% CI [−0.10, 0.22]; $\beta_{\text{age: mean}}=-0.08$, 95% CI [−0.19, 0.02]; $\beta_{\text{age: mean+1sd}}=-0.22$, 95% CI [−0.36, −0.09]). However, TD children showed a small age-increased pattern for familiar items ($\beta_{\text{age: mean-1sd}}=0.02$, 95% CI [−0.07, 0.11]; $\beta_{\text{age: mean}}=0.05$, 95% CI [−0.01, 0.11]; $\beta_{\text{age: mean+1sd}}=0.08$, 95% CI [0.01, 0.15]) and an age-invariant pattern for unfamiliar items ($\beta_{\text{age: mean-1sd}}=-0.02$, 95% CI [−0.10, 0.06]; $\beta_{\text{age: mean}}=-0.02$, 95% CI [−0.08, 0.03]; $\beta_{\text{age: mean+1sd}}=-0.03$, 95% CI [−0.09, 0.03]).

Conditional statistical learning: Working memory is positively associated with the recognition of familiar–unfamiliar items in children with DD across different ages

Another generalized linear mixed model was conducted to examine the relation between working memory and conditional statistical learning and the factors affecting it, with outcome responses in the visual triplet learning experiment as the dependent variable (Figure 3a; see Table S2 for details). The main effects of the two covariates, sex (p = .606) and nonverbal intelligence (p = .151), were not significant. The model revealed a significant main effect of age (β = 0.06, 95% CI [0.02, 0.10], z = 2.95, p = .003), indicating an increase in conditional statistical learning across development. Compared to TD children, children with DD showed significantly lower performance on the 2AFC test (β = −0.09, 95% CI [−0.18, −0.01], z = −2.13, p = .033). Working memory (p = .908) and familiarity (p = .914) did not yield significant main effects.

The relation between statistical learning and working memory was modulated by familiarity and reading status (Figure 3b), indicated by a significant three-way interaction among working memory, familiarity, and reading status (β = 0.11, 95% CI [0.02, 0.21], z = 2.26, p = .024). Post hoc tests showed that children with DD had a larger working memory effect on familiar–unfamiliar than on unfamiliar–familiar items ($\beta_{\text{familiar–unfamiliar}}$ = 0.04, 95% CI [−0.06, 0.13]; $\beta_{\text{unfamiliar–familiar}}$ = −0.05, 95% CI [−0.14, 0.04]), while TD children exhibited no difference in the effects involving the two kinds of items ($\beta_{\text{familiar–unfamiliar}}$ = −0.01, 95% CI [−0.06, 0.04]; $\beta_{\text{unfamiliar–familiar}}$ = 0.02, 95% CI [−0.03, 0.07]). Age did not further influence this pattern, as the four-way interaction among working memory, familiarity, age, and reading status was not significant (p = .488).

DISCUSSION

This study investigated the relation between working memory and learning of distributional and conditional statistics in a large group of children with DD and their TD peers. In general, the results showed no significant difference between the two groups in the association of working memory with statistical learning. However, compared to their TD peers, children with DD displayed distinct patterns in this relation, particularly in terms of item characteristics and age. In the context of distributional statistical learning of an artificial orthography, both groups exhibited a stronger working memory effect when recognizing high-predictable items compared to moderate-predictable items, and no difference between high-predictable and low-predictable items. However, the influence of item familiarity on the working memory effect differed between the two groups. As age increased, children with DD and their TD peers exhibited an increased positive effect of working memory on recognizing familiar items. However, an opposite pattern was observed in the recognition of unfamiliar items, specifically among children with DD. Additionally, in the context of conditional statistical learning of visual triplets,

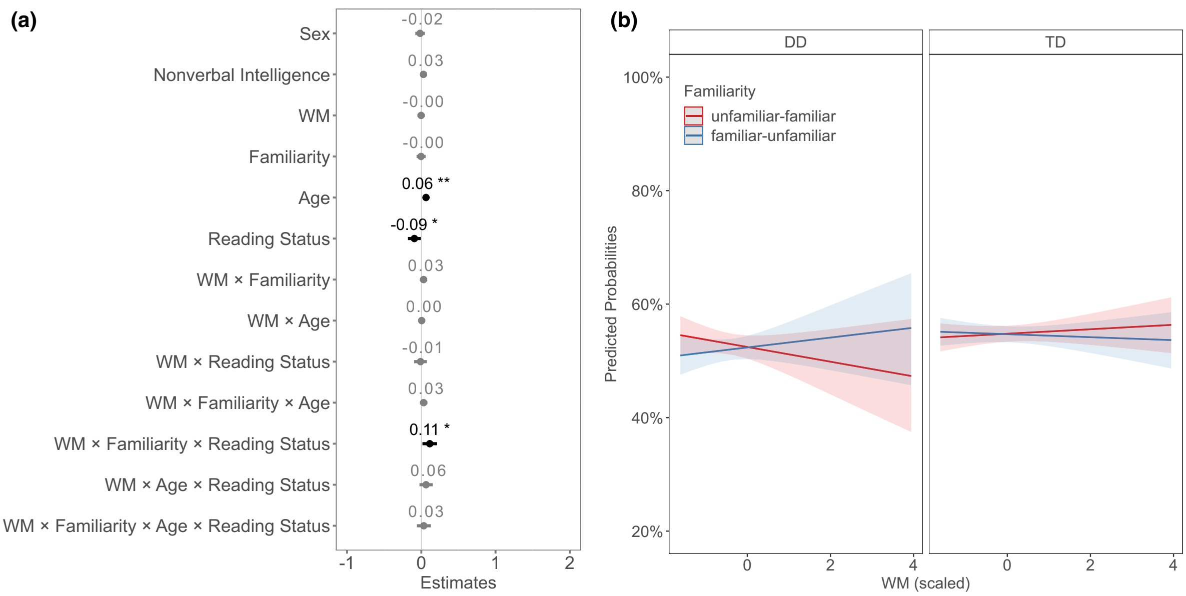

FIGURE 3 Influences of item characteristics, age, and reading status (DD: children with developmental dyslexia; TD: typically developing children) on the relation between conditional statistical learning and working memory (WM). Panel (a) displays the coefficient estimates of the generalized linear mixed model. Panel (b) shows that the associations between working memory and statistical learning varied under the familiar–unfamiliar and unfamiliar–familiar items in the DD and TD groups. The error bars in panel (a) and the shade areas in panel (b) indicate 95% confidence intervals. *p < .05; **p < .01.

children with DD showed a stronger effect of working memory on familiar–unfamiliar items compared to unfamiliar–familiar items, while no such association was found in TD.

Consistent with a previous EEG study (Tong et al., 2023) that suggested the involvement of attentional and working memory resources during learning high-predictable and low-predictable items, our study found that working memory played a greater role in processing highly regular (i.e., high-predictable) and highly irregular (i.e., low-predictable) inputs during distributional statistical learning of an artificial orthography. Specifically, working memory exhibited a stronger positive association with the recognition of high-predictable compared to moderate-predictable but not low-predictable items in children with and without DD. In Tong et al.'s (2023) study, high-predictable items elicited a larger initial-stage P1 component while low-predictable items elicited a larger later-stage P300 component. Thus, it is possible that the comparable positive effect of working memory on high-predictable and low-predictable items may operate in different stages during statistical learning. Another explanation is based on the distinction between deterministic and probabilistic rules. According to Lee et al. (2022), deterministic rules may be stored in the explicit declarative system with more working memory resources involved, while probabilistic rules are more likely to operate in the implicit procedural system. In our study, the radical positions were deterministic for high-predictable (100%) pseudocharacters but probabilistic for moderate-predictable (80%) items. Therefore, children may utilize conscious resources or explicit memory to identify high-predictable items during the retrieval stage, resulting in a positive relation between working memory and the recognition of high-predictable items but not moderate-predictable items.

Contrary to our hypothesis that assumed more involvement of working memory in processing unfamiliar items, we found a stronger positive effect of working memory on familiar, but not unfamiliar items in distributional statistical learning of an artificial orthography regardless of children's reading status. However, we observed different age-related patterns between children with DD and their TD peers. Specifically, TD children showed an age-increased working memory relation with the recognition of familiar but not unfamiliar items, indicating an increase in the use of working memory resources to retrieve acquired knowledge (i.e., familiar items). This pattern reflects an exploitation strategy that develops with age. Although children with DD exhibited a similar age-related increase in the effect of working memory on recognizing familiar items, we observed a negative association between working memory and statistical learning for unfamiliar items, particularly among older children with DD, but not TD children. This pattern in children with DD does not reflect an exploitation strategy because if they were able to utilize working memory resources to exploit the correct positional regularity, their working memory capacity would also have a positive, or at least not a negative, effect on predicting the recognition of unfamiliar items.

One plausible explanation for the opposite effects of working memory on recognizing familiar and unfamiliar items is that children with DD overgeneralize the similarities among the inputs. They may utilize their working memory resources to seek connections and relevance between the inputs and their long-term memory traces, influencing their decision criterion to choose more items as "familiar," thus resulting in an increased accuracy on familiar items but a decreased accuracy on unfamiliar items as working memory capacity increases. This explanation is supported by an additional analysis of the decision criterion index (C) based on the signal detection theory (Stanislaw & Todorov, 1999). Specifically, we found that in older children with DD, a higher working memory capacity was related to a lower negative value of C, indicating a more lenient criterion to consider the stimulus as "familiar" (see the details in Table S3; Figure S1). This explanation is also consistent with the fuzzy memory trace theory, which suggests that individuals with DD have an enhanced gist memory but less accuracy on verbatim items (Obidziński & Nieznański, 2017), reflecting their reliance on the exploration strategy (Taylor & Vestergaard, 2022).

In the context of conditional learning of visual triplets, we found that the presentation order of the familiar and unfamiliar triplets influenced the association between working memory and statistical learning in children with DD, but not their TD peers. This finding partially aligns with those of previous studies that found no performance differences for different target sequences during visual statistical learning in TD learners (Raviv & Arnon, 2018; Shufaniya & Arnon, 2018). Interestingly, children with DD exhibited a larger positive effect of working memory on recognizing triplets from familiar–unfamiliar trials compared to unfamiliar–familiar trials. However, their accuracy was lower in the familiar–unfamiliar condition compared to their TD peers. This is because the six items comprising the two triplets exceed the working memory capacity of children with DD, making it difficult for them to perceive and discriminate between the two triplets. Therefore, their decision-making process relied mostly on the second triplet. Additionally, as these children tended to seek similarity between the presented items and their internal memory of items, they preferentially chose the second triplet as "familiar." This resulted in more incorrect responses for familiar–unfamiliar trials. In contrast, children with higher working memory capacity found it more feasible to recall the first triplet, leading to a positive effect of working memory on the familiar–unfamiliar trials.

The item-regulated relation between statistical learning and working memory observed at varying stages of development elucidates different strategies employed by children with DD and TD peers in optimizing working memory resources. Given the limited capacity of working memory, humans need to efficiently process a vast amount of information and make predictions and decisions based on predictable and unpredictable, familiar and unfamiliar inputs. A common view is that inputs with more complex information, such as less predictable and unfamiliar information, may require more explicit efforts through working memory functions (Conway, 2020). However, our findings suggest an alternative possibility: working memory may preferentially retrieve more relevant and salient information to optimize statistical learning outcomes. This is evident in both children with DD and their TD peers, as they optimize their use of working memory by allocating more resources when processing relatively determined items (i.e., high- or low-predictable) and familiar items, rather than uncertain items that cause confusion (i.e., moderate-predictable) or unfamiliar items.

Furthermore, the optimization strategy observed in children with DD involves accepting or exploring the similarity between different kinds of inputs, both familiar and unfamiliar. This is likely due to their limited working memory capacity, which makes it challenging for them to accurately discriminate an unfamiliar input from internal representations. Understanding how children with DD develop this optimization function for learning is an important area for further investigation.

Finally, our findings on the crucial role of working memory in the development of statistical learning have implications for two competing claims regarding the mechanisms underlying statistical learning. One line of research suggests that statistical learning is a multifaceted construct, as evidenced by the variant developmental patterns observed across different modalities (e.g., visual vs. auditory; Raviv & Arnon, 2018) and stimulus types (e.g., linguistic vs. non-linguistic; Shufaniya & Arnon, 2018), as well as the contributions of various cognitive functions (e.g., working memory and attention) to statistical learning (Hall et al., 2015; Richter & de Lange, 2019; Roembke & McMurray, 2021; Toro et al., 2005). However, another line of research argues that statistical learning is supported by a unifying, domain-general mechanism (Kirkham et al., 2002), such that an individual's statistical learning ability develops as a shared pattern under different contexts and is not related to other cognitive functions (Gao & Theeuwes, 2020; Horváth et al., 2020; Vickery et al., 2010). Our study suggests that the non-significant relations that have been found may be due to distinct manifestations of working memory under different item characteristics in children with and without DD. Thus, we support the multiple-component view by showing that the multifaceted mechanisms of statistical learning operate at the item level, rather than at the overall learning performance level.

Furthermore, the item- and age-regulated relations that we discovered between statistical learning and working memory in children with DD highlight the need for a more fine-grained approach to studying statistical learning in dyslexia. Previous studies examining group differences between children with DD and their TD peers have generated inconsistent results, with some indicating impaired learning (e.g., Sigurdardottir et al., 2017) while others show intact learning (e.g., van Witteloostuijn et al., 2019). Similarly, in our study, we observed lower overall performance in conditional statistical learning but not in distributional statistical learning. These conflicting results in simple comparisons have made it challenging to determine whether DD is associated with deficits in statistical learning.

However, our findings suggest that regardless of impaired or intact statistical learning, children with DD allocated their working memory resources differently compared to their TD peers. This aligns with a previous study showing that individuals with DD tend to have a broader focus on high-predictable and low-predictable items compared to TD children (Singh et al., 2018). Thus, instead of interpreting the distinct patterns between children with DD and their TD peers as atypical statistical learning processes, our study illuminates that these patterns reflect an exploration strategy adopted by children with DD. As Taylor and Vestergaard (2022) suggest, such strategy is not inherently detrimental and may lead to strengths such as creative thinking in individuals with DD. Therefore, future interventions tailored to their reading difficulties should be designed to preserve their strengths.

Despite the important implications of our findings, it should be noted that we measured a general working memory capacity using a single-digit span task. However, previous research has suggested that the relation between statistical learning and working memory may be sensitive to the specific modality of the statistical learning and working memory measures (Janacsek & Nemeth, 2013). Future studies could consider measuring different modalities (i.e., verbal or visual) or subcomponents (e.g., central executive) of working memory, and include other age-variant cognitive skills (e.g., visual attention) as covariates. Additionally, aligning with previous studies on statistical learning, our assessment of statistical learning was limited to the testing phase, which represents the ability to retrieve long-term knowledge. Moreover, compared to the memory recognition task that involved responding to a single item in the artificial orthography learning experiment, the 2AFC task required memorizing six items at a time in the visual triplet learning experiment and thus may have demanded more working memory resources, influencing the relation between statistical learning and working memory. It would be beneficial for future studies to include consistent procedures or online indexes of statistical learning to investigate the role of working memory in the encoding and maintenance of different items during learning. Future studies should also be cautious in using 2AFC tasks to assess statistical learning abilities in children with DD who have low working memory capacities. Finally, it is worth noting that our study focused on Chinese children learning an artificial orthography and visual transitions. This raises the possibility that the observed item- and age-regulated relations between statistical learning and working memory may be specific to a particular context. Therefore, to establish the generalizability of our results, future studies are needed to replicate our findings in other contexts involving distributional and conditional statistics in different languages.

In summary, our study utilized two statistical learning paradigms and one working memory task to investigate the role of working memory in statistical learning among children with and without DD. Our findings demonstrate that the relation between working memory and statistical learning is contingent upon the item characteristics embedded within a particular type of statistical learning. Both children with DD and their TD peers allocated more working memory resources to high-predictable over moderate-predictable items. However, while TD children utilized working memory resources to retrieve familiar items based on acquired regularities, children with DD employed limited working memory resources to seek similarity among items as age increased. These results highlight the distinct optimization functions between children with DD and their TD peers in the memory-based mechanisms of statistical learning across childhood and suggest that children with DD could benefit from interventions that optimize the utilization of working memory resources when processing unfamiliar or uncertain inputs.

ACKNOWLEDGMENTS

The work described in this manuscript was supported by the Research Fellow Scheme (RFS 2021-7H05) and General Research Fund (17612222 and 17616323) from the Research Grants Council of the Hong Kong Special Administrative Region, China, to Shelley Xiuli Tong. We would like to thank all participating families and children. Additionally, we thank Dr. John Martino for proofreading the final draft of our manuscript.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

The data, analytic codes and materials for the replication of the analyses presented here are not publicly accessible due to concerns of protecting confidentiality of participants' data. The analyses presented in this study were not preregistered.

ORCID

Shelley Xiuli Tong https://orcid.org/0000-0003-3319-4609

REFERENCES

Alloway, T. P., & Alloway, R. G. (2013). Working memory across the lifespan: A cross-sectional approach. *Journal of Cognitive Psychology*, *25*(1), 84–93. https://doi.org/10.1080/20445911.2012.748027

Arciuli, J., & Simpson, I. C. (2011). Statistical learning in typically developing children: The role of age and speed of stimulus presentation. *Developmental Science*, *14*(3), 464–473. https://doi.org/10.1111/j.1467-7687.2009.00937.x

Arciuli, J., & Simpson, I. C. (2012). Statistical learning is lasting and consistent over time. *Neuroscience Letters*, *517*(2), 133–135. https://doi.org/10.1016/j.neulet.2012.04.045

Ballan, R., Durrant, S. J., Manoach, D. S., & Gabay, Y. (2023). Failure to consolidate statistical learning in developmental dyslexia. *Psychonomic Bulletin & Review*, *30*(1), 160–173. https://doi.org/10.3758/s13423-022-02169-y

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. https://doi.org/10.1016/j.jml.2015.04.004

Bo, J., Borza, V., & Seidler, R. D. (2009). Age-related declines in visuospatial working memory correlate with deficits in explicit motor sequence learning. *Journal of Neurophysiology*, *102*(5), 2744–2754. https://doi.org/10.1152/JN.00393.2009/ASSET/IMAGES/LARGE/Z9K0110997710004.JPEG

Bo, J., Jennett, S., & Seidler, R. D. (2012). Differential working memory correlates for implicit sequence performance in young and older adults. *Experimental Brain Research*, *221*(4), 467–477. https://doi.org/10.1007/s00221-012-3189-2

Conway, C. M. (2020). How does the brain learn environmental structure? Ten core principles for understanding the neurocognitive mechanisms of statistical learning. *Neuroscience & Biobehavioral Reviews*, *112*, 279–299. https://doi.org/10.1016/j.neubiorev.2020.01.032

Cowan, N. (1998). Visual and auditory working memory capacity. *Trends in Cognitive Sciences*, *2*(3), 77. https://doi.org/10.1016/S1364-6613(98)01144-9

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

Emberson, L. L., Misyak, J. B., Schwade, J. A., Christiansen, M. H., & Goldstein, M. H. (2019). Comparing statistical learning across perceptual modalities in infancy: An investigation of underlying learning mechanism(s). *Developmental Science*, *22*(6), e12847. https://doi.org/10.1111/desc.12847

Erickson, L. C., & Thiessen, E. D. (2015). Statistical learning of language: Theory, validity, and predictions of a statistical learning account of language acquisition. *Developmental Review*, *37*, 66–108. https://doi.org/10.1016/j.dr.2015.05.002

Frankenhuis, W. E., & Gopnik, A. (2023). Early adversity and the development of explore–exploit tradeoffs. *Trends in Cognitive Sciences*, *27*, 616–630. https://doi.org/10.1016/j.tics.2023.04.001

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

Gabay, Y., Thiessen, E. D., & Holt, L. L. (2015). Impaired statistical learning in developmental dyslexia. *Journal of Speech, Language, and Hearing Research*, *58*(3), 934–945. https://doi.org/10.1044/2015_JSLHR-L-14-0324

Gao, Y., & Theeuwes, J. (2020). Learning to suppress a distractor is not affected by working memory load. *Psychonomic Bulletin &*

Review, *27*(1), 96–104. https://doi.org/10.3758/s13423-019-01679-6

Gopnik, A. (2020). Childhood as a solution to explore–exploit tensions. *Philosophical Transactions of the Royal Society, B: Biological Sciences*, *375*(1803), 20190502. https://doi.org/10.1098/rstb.2019.0502

Gray, S., Fox, A. B., Green, S., Alt, M., Hogan, T. P., Petscher, Y., & Cowan, N. (2019). Working memory profiles of children with dyslexia, developmental language disorder, or both. *Journal of Speech, Language, and Hearing Research*, *62*(6), 1839–1858. https://doi.org/10.1044/2019_JSLHR-L-18-0148

Hall, M. G., Mattingley, J. B., & Dux, P. E. (2015). Distinct contributions of attention and working memory to visual statistical learning and ensemble processing. *Journal of Experimental Psychology: Human Perception and Performance*, *41*(4), 1112–1123. https://doi.org/10.1037/xhp0000069

He, X., & Tong, S. X. (2017). Statistical learning as a key to cracking Chinese orthographic codes. *Scientific Studies of Reading*, *21*(1), 60–75. https://doi.org/10.1080/10888438.2016.1243541

Hendricks, M. A., Conway, C. M., & Kellogg, R. T. (2013). Using dual-task methodology to dissociate automatic from nonautomatic processes involved in artificial grammar learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*(5), 1491–1500. https://doi.org/10.1037/a0032974

Ho, S. H. C., Chan, D., Chung, K. K. H., Tsang, S. M., Lee, S. H., & Cheng, W. Y. R. (2007). *The Hong Kong Test of Specific Learning Difficulties in Reading and Writing for Primary School Students, second edition [HKT-P(II)]*. Hong Kong Specific Learning Difficulties Research Team.

Horváth, K., Török, C., Pesthy, O., Nemeth, D., & Janacsek, K. (2020). Divided attention does not affect the acquisition and consolidation of transitional probabilities. *Scientific Reports*, *10*(1), 22450. https://doi.org/10.1038/s41598-020-79232-y

Janacsek, K., & Nemeth, D. (2013). Implicit sequence learning and working memory: Correlated or complicated? *Cortex*, *49*(8), 2001–2006. https://doi.org/10.1016/J.CORTEX.2013.02.012

Jongbloed-Pereboom, M., Janssen, A. J. W. M., Steiner, K., Steenbergen, B., & Nijhuis-van der Sanden, M. W. G. (2017). Implicit and explicit motor sequence learning in children born very preterm. *Research in Developmental Disabilities*, *60*, 145–152. https://doi.org/10.1016/j.ridd.2016.11.014

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

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

Kraemer, H. C., & Blasey, C. M. (2004). Centring in regression analyses: A strategy to prevent errors in statistical inference. *International Journal of Methods in Psychiatric Research*, *13*(3), 141–151. https://doi.org/10.1002/mpr.170

Lee, S. M. K., Cui, Y., & Tong, S. X. (2022). Toward a model of statistical learning and reading: Evidence from a meta-analysis. *Review of Educational Research*, *92*(4), 651–691. https://doi.org/10.3102/00346543211073188

Lee, S. M. K., & Tong, X. (2020). Spelling in developmental dyslexia in Chinese: Evidence of deficits in statistical learning and over-reliance on phonology. *Cognitive Neuropsychology*, *37*(7–8), 494–510. https://doi.org/10.1080/02643294.2020.1765754

Lyon, G. R., Shaywitz, S. E., & Shaywitz, B. A. (2003). A definition of dyslexia. *Annals of Dyslexia*, *53*, 1–14. https://doi.org/10.1007/s11881-003-0001-9

Obidziński, M., & Nieznański, M. (2017). False memory for orthographically versus semantically similar words in adolescents with dyslexia: A fuzzy-trace theory perspective. *Annals of Dyslexia*, *67*(3), 318–332. https://doi.org/10.1007/s11881-017-0146-6

Raviv, L., & Arnon, I. (2018). The developmental trajectory of children's auditory and visual statistical learning abilities: Modality-based differences in the effect of age. *Developmental Science*, *21*(4), e12593. https://doi.org/10.1111/desc.12593

Richter, D., & de Lange, F. P. (2019). Statistical learning attenuates visual activity only for attended stimuli. *eLife*, *8*, e47869. https://doi.org/10.7554/eLife.47869

Roembke, T. C., & McMurray, B. (2021). Multiple components of statistical word learning are resource dependent: Evidence from a dual-task learning paradigm. *Memory & Cognition*, *49*(5), 984–997. https://doi.org/10.3758/s13421-021-01141-w

Saffran, J. R. (2002). Constraints on statistical language learning. *Journal of Memory and Language*, *47*(1), 172–196. https://doi.org/10.1006/jmla.2001.2839

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

Saito, S., Nakayama, M., & Tanida, Y. (2020). Verbal working memory, long-term knowledge, and statistical learning. *Current Directions in Psychological Science*, *29*(4), 340–345. https://doi.org/10.1177/0963721420920383

Schiff, R., Katan, P., Sasson, A., & Kahta, S. (2017). Effect of chunk strength on the performance of children with developmental dyslexia on artificial grammar learning task may be related to complexity. *Annals of Dyslexia*, *67*(2), 180–199. https://doi.org/10.1007/s11881-017-0141-y

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-prime user's guide*. Psychology Software Tools. Inc.

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

Siegelman, N., & Frost, R. (2015). Statistical learning as an individual ability: Theoretical perspectives and empirical evidence. *Journal of Memory and Language*, *81*, 105–120. https://doi.org/10.1016/j.jml.2015.02.001

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

Singh, S., Walk, A. M., & Conway, C. M. (2018). Atypical predictive processing during visual statistical learning in children with developmental dyslexia: An event-related potential study. *Annals of Dyslexia*, *68*(2), 165–179. https://doi.org/10.1007/s11881-018-0161-2

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, & Computers*, *31*(1), 137–149. https://doi.org/10.3758/BF03207704

Taylor, H., & Vestergaard, M. D. (2022). Developmental dyslexia: Disorder or specialization in exploration? *Frontiers in Psychology*, *13*, 889245. https://doi.org/10.3389/fpsyg.2022.889245

Thiessen, E. D. (2017). What's statistical about learning? Insights from modelling statistical learning as a set of memory processes. *Philosophical Transactions of the Royal Society, B: Biological Sciences*, *372*(1711), 20160056. https://doi.org/10.1098/rstb.2016.0056

Tong, S. X., Duan, R., Shen, W., Yu, Y., & Tong, X. (2023). Multiple mechanisms regulate statistical learning of orthographic regularities in school-age children: Neurophysiological evidence. *Developmental Cognitive Neuroscience*, *59*, 101190. https://doi.org/10.1016/j.dcn.2022.101190

Tong, S. X., Zhang, P., & He, X. (2020). Statistical learning of orthographic regularities in Chinese children with and without dyslexia. *Child Development*, *91*(6), 1953–1969. https://doi.org/10.1111/cdev.13384

Toro, J. M., Sinnett, S., & Soto-Faraco, S. (2005). Speech segmentation by statistical learning depends on attention. *Cognition*, *97*(2), B25–B34. https://doi.org/10.1016/j.cognition.2005.01.006

Tseng, C.-C., Chen, H.-C., Chang, L.-Y., Hu, J.-F., & Chen, S.-Y. (2018). A corpus-based analysis of radical position and the degree of freedom of permissible positions and examination of the influential radical properties. *Bulletin of Educational Psychology*, *49*, 487–511. https://doi.org/10.6251/BEP.201803_49(3).0007

van Witteloostuijn, M., Boersma, P., Wijnen, F., & Rispens, J. (2019). Statistical learning abilities of children with dyslexia across three experimental paradigms. *PLoS One*, *14*(8), e0220041. https://doi.org/10.1371/journal.pone.0220041

Vickery, T. J., Sussman, R. S., & Jiang, Y. V. (2010). Spatial context learning survives interference from working memory load. *Journal of Experimental Psychology: Human Perception and Performance*, *36*(6), 1358–1371. https://doi.org/10.1037/a0020558

Wechsler, D. (2011). *Wechsler Abbreviated Scale of Intelligence—Second edition*. Pearson.

Weinberger, A. B., & Green, A. E. (2022). Dynamic development of intuitions and explicit knowledge during implicit learning. *Cognition*, *222*, 105008. https://doi.org/10.1016/j.cognition.2021.105008

Zhang, G. (2022). *YawMMF* (0.2.1) [R]. https://github.com/usplos/YawMMF

Zhang, P., Chen, H., & Tong, S. X. (2023). Cue predictiveness and uncertainty determine cue representation during visual statistical learning. *Learning & Memory*, *30*(11), 282–295. https://doi.org/10.1101/lm.053777.123

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Zhou, M., Zhang, P., Mimeau, C., & Tong, S. X. (2024). Unraveling the complex interplay between statistical learning and working memory in Chinese children with and without dyslexia across different ages. *Child Development*, *95*, e338–e351. https://doi.org/10.1111/cdev.14121