Artificial Grammar Learning in Dyslexic and Nondyslexic Adults: Implications for Orthographic Learning

Anna Samara and Markéta Caravolas

School of Psychology, Bangor University, Bangor, LL57 2AS, UK

Corresponding Author:

Anna Samara

Division of Psychology and Language Sciences

University College London

Chandler House, 2 Wakefield Street

London, WC1N 1PF

United Kingdom

Email: anna.samara@ucl.ac.uk

Abstract

Potential implicit orthographic learning deficits were investigated in adults with dyslexia. An artificial grammar learning paradigm served to assess dyslexic and typical readers' ability to exploit information about chunk frequency, letter-position patterns, and specific string similarity, all of which have analogous constructs in real orthographies. We also investigated whether implicit learning deficits in dyslexia held for letter strings (experiment 1) and symbol strings (experiment 2). Experiment 1 results indicated that dyslexic adults were mildly impaired in memorizing letter strings, although this finding proved inconclusive in a more stringent analysis of the data across experiments. There were no signs of difficulty during symbol string memorization in experiment 2. In each experiment, dyslexic and nondyslexic readers were comparably sensitive to chunk frequencies and showed reliable sensitivity to letter and shape position patterns and string similarities. These findings challenge the claim that a general learning deficit contributes to literacy difficulties in dyslexia.

Keywords: artificial grammar learning, implicit learning, dyslexia, orthographic learning

Deficits in implicit, or incidental, skill learning have been put forward as a potential explanation for several, language-based developmental disorders, including dyslexia. Impaired implicit skill learning has been proposed to hinder the development of well-specified phonological representations or to have a detrimental effect on dyslexic children's ability to form grapheme-phoneme associations (Gombert, 2003; Sperling, Lu, & Manis, 2004). Despite the appeal of such parsimonious causal hypotheses, research findings are mixed, and several questions remain to be adequately addressed. Are the putative implicit learning difficulties of learners with dyslexia persistent, lasting into adulthood? Are they domain-general or specific to the domain of symbol-sound processing?

Among the few studies of implicit skill learning in child and adult developmental dyslexic populations, only five, to our knowledge, have employed Reber's (1967) Artificial Grammar Learning (AGL) task—an implicit learning task that assesses pattern knowledge using strings generated from an artificial, finite-state grammar, such as the one shown in Figure 1.

------------Figure 1------------

In a standard version of the AGL experiment, participants are warned of an impending memory task and are presented with seemingly arbitrary stimuli (e.g., letter strings) to be observed, mentally rehearsed, or reproduced. Unbeknown to them, all memorization stimuli have been generated by traversing left-to-right through one of the possible pathways of an artificial grammar (e.g., the pathway F→X→D→H→F→M→L in Figure 1 generates the permissible string FXDHFML). Following memorization, participants are informed that the stimuli conformed to an untaught grammar and are asked to discriminate between novel strings that are either grammatical or ungrammatical (e.g., RTGMCQV is ungrammatical, as there is no permissible pathway between M as a fourth letter and C as a fifth letter in Figure 1). Although participants are unaware of the grammaticality manipulation during training, discrimination between grammar-conforming and nonconforming stimuli reliably exceeds chance. Whatever participants learn while memorizing the grammar-conforming instances, they use it to judge the grammaticality of new strings with above-chance accuracy.

There is less agreement on the nature of the acquired representations (see Pothos, 2007, for a review). It was originally proposed that learners base their grammaticality judgments on an abstract representation of the rules governing the training stimuli (*rule-based theories*; e.g., Reber, 1989), and/or the similarity between test and individual training stimuli (*instance-based theories*; e.g., Brooks & Vokey, 1991). However, accruing evidence suggests that participants' performance reflects sensitivity to frequency-based chunk (e.g., bigram and trigram) information (*fragment-based theories*; e.g., Perruchet & Pacteau, 1990), although the presence of additional unexplained variance indicates that other string aspects may also be encoded (e.g., Johnstone & Shanks, 1999). Moreover, what is learnt (e.g. rules vs. small/large fragments), may depend on different training parameters such as task demands and type of instructions (*episodic processing account*; e.g., Wright & Whittlesea, 1998).

In one of the four AGL studies with dyslexic adults (Rüsseler et al., 2006), skilled and dyslexic readers who had been trained to criterion on 4-letter, grammatical strings outperformed control participants trained with random (i.e., uninformative) strings. Moreover, the reading ability groups did not differ in memorization performance or grammaticality sensitivity at test. Thus, there was no evidence of a learning deficit among dyslexic adults. Pothos and Kirk (2004) assessed learning using (a) geometrical shapes arranged sequentially and (b) nested figures embedding the shapes, pitting grammaticality against *chunk strength* information (i.e., frequency of bigrams and trigrams within training strings) (Knowlton & Squire, 1996). While skilled and dyslexic readers performed statistically similarly to each other in the embedded condition, only dyslexic individuals performed reliably better than guessing in the sequential variant. However, the finding that dyslexics outperformed skilled readers is at best tentative, given that learning was not reliably demonstrated among control participants. Indeed, more recently Laasonen et al. (2013) demonstrated a significant impairment in dyslexic adults' ability to learn the underlying structure of the grammar used by Pothos and Kirk (2004). An implicit learning deficit among dyslexic compensated adults was also shown by Kahta and Schiff (2016). Dyslexic participants were significantly worse than skilled readers in discriminating grammatical from ungrammatical letter strings under implicit task instructions, but not under explicit task instructions. It is clear that these discrepancies in results across studies may be due to differences in training phase manipulations (e.g., learning via simple exposure vs. memorization to criterion), or stimuli (e.g., Kahta and Schiff's ungrammatical stimuli were easy to reject in that most began with an illegal initial letter). It is also noteworthy that the cognitive profiles of adults with dyslexia, including those with well-compensated reading accuracy skills, are highly heterogeneous; thus, inconsistent results may also reflect aggregated cognitive profile differences among samples of dyslexic adults.

To investigate possibly more prominent implicit learning susceptibilities among different groups during development, Ise, Arnoldi, Bartling, and Schulte-

Körne (2012) measured chunk strength sensitivity in 9-year-old typical and poor spellers following exposure to pronounceable (e.g., XABOZ) versus unpronounceable (e.g., FTGCZ) strings. Good spellers were more accurate than poor spellers, although both groups performed above chance in both experimental conditions. Using a nonlinguistic task variant and a chunk-strength-balanced design whereby grammatical and ungrammatical items are made of frequent training chunks to the *same* extent (for details, see Knowlton and Squire, 1996), Pavlidou, Kelly, and Williams (2010) investigated dyslexic children's sensitivity to chunk frequency versus abstract, rule-based, information. Pavlidou et al. (2010) reported that only typically developing children were able to reliably classify the sequences on the basis of grammaticality and chunk strength. However, chunk strength sensitivity was analyzed by separate one-sample *t* tests, against chance, on the proportion of "yes" responses for high chunk strength items and the proportion of "yes" responses for low chunk strength items; this approach may confound participants' sensitivity and response criterion (signal detection theory; Macmillan & Kaplan, 1985). Nigro, Jiménez-Fernández, Simpson, and Defior (2015) used an AGL variant to assess Spanish dyslexic children's ability to implicitly learn and generalize over simple letter-position patterns (e.g., A can begin letter strings, B cannot) embedded either within linguistic or nonlinguistic strings. Regardless of stimulus format, there was no statistical evidence of impairment among dyslexic children.

In sum, the evidence of impaired AGL in children and adults with developmental dyslexia is scant and partially conflicting across age groups and reading ability groups. While it seems plausible that implicit learning performance might improve with experience over the course of typical development, whether this type of learning comprises a core impairment in dyslexia is not yet clear. Whether conflicting findings are due to differences in dyslexic individuals' age and experience with print, developmental differences in implicit skill learning (e.g., Thomas et al., 2004), or confounds of different training and item manipulations is subject to future research. In this study, we sought to establish among adult learners whether implicit learning differences could be observed under different stimulus manipulations, and as a function of the participants' status as dyslexic or non-dyslexic readers.

The present study

As discussed, various types of acquired information potentially contribute to successful AGL performance. Thus, it is pertinent to ask whether dyslexic individuals are less sensitive to specific *types* of knowledge that can be learned in this context. The two adult studies are silent on this question, and in studies with dyslexic children, Nigro et al. (2015) did not control for chunk strength, while Pavlidou et al.'s (2010) study did not provide the most stringent analysis of chunk strength sensitivity. Therefore, a well-controlled study of typical and dyslexic readers is required, which investigates sensitivity to chunk strength information, as well as the ability to exploit letter-position patterns or learn whole exemplars, controlling for chunk frequency.

Both of the latter properties are ecologically relevant to reading and spelling skill. Children and adults are sensitive to statistical information regarding the allowable position of double letters within their written language (e.g., they prefer nonwords like *baff* to nonwords like *bbaf*; Cassar & Treiman, 1997) and some letter-position patterns can be learnt, at least to some extent, under minimal incidental exposure conditions from a young age (Samara & Caravolas, 2014). Impaired sensitivity to this attribute among dyslexic individuals would suggest that they are less sensitive to legal/illegal letter-position patterns that support spelling in skilled readers. Sensitivity to overall similarity between learned exemplars and test strings may also have relevance for real orthographic learning because reading and spelling performance may benefit in part from learning by analogy (e.g., Ehri, 1997; Goswami, 1988). If impaired among dyslexic individuals, it may suggest that they are less efficient in reading/spelling by an analogical inference strategy.

The AGL paradigm used in this study was based on that developed by Kinder and Lotz (2009; see also Kinder, 2000). Specifically, we investigated participants' ability to categorize strings as grammatical/ungrammatical at test depending on: (a) the extent to which they contained allowable/familiar training chunks (frequency-based chunk strength), (b) their adherence to patterns on letter position, set by the grammar, and (c) their degree of similarity to specific letter strings presented during training. Kinder and Lotz's (2009) design is well-suited for a detailed investigation of orthographic frequency-based learning abilities in several respects: It allows for systematic manipulation of letter positions, a type of information that is statistically constrained in written language and relevant to literacy skills (e.g., Cassar & Treiman, 1997), yet is relatively overlooked in AGL studies. Sensitivity to chunk frequency information and specific similarity are also important for orthographic learning (Ehri, 1995; Goswami, 1988; Nation, 1997). Finally, the grammar generates strings of equal length and prohibits salient patterns of repetitions (e.g., MT*RRR*), obviating the need to control for these variables.

Experiment 1

In experiment 1, we asked whether skilled and dyslexic readers demonstrate implicit chunk strength sensitivity and sensitivity to letter-position patterns, and specific string similarity. Training was provided through a memorization to criterion task, during which participants reproduced stimuli by dragging and dropping their constituent elements into 7 response boxes, as opposed to typing them from memory as in Kinder and Lotz's (2009) study. This is an easier memorization task, known to reduce the cognitive load on individual item memory by relying less heavily on recall (Tremblay, Macken, & Jones, 2000). Furthermore, the drag and drop task bypassed keyboard skills—another potential source of variation affecting learning performance.

Method

Participants

Thirty-one skilled readers (13 male; mean age = 20.49 years) and 22 students diagnosed with dyslexia (7 male; mean age = 20.80 years) participated. Inclusion criteria for skilled readers were: no documented history of dyslexia or other learning difficulties, WRAT IV Reading/Spelling (Wilkinson & Robertson, 2006) standard score performance above 95 (checked a posteriori), and verbal and performance IQ above 80 (checked a posteriori). Dyslexic readers were recruited through the University's dyslexia Center (see Supplementary Materials for details on the criteria used for diagnosis by the University dyslexia Center). Study inclusion criteria were: formal diagnosis of dyslexia and no known co-occurrence of another developmental disorder (confirmed a posteriori through access to all students' dyslexia assessment reports, following their written consent) and verbal and performance IQ above 80 (checked a posteriori). Three dyslexic participants failed to reproduce all training strings correctly within a reasonable time frame (~75 minutes) and could not proceed to the test-phase judgment task. All participants were monolingual English speakers and reported having normal or corrected-to-normal visual acuity.

Background measures

Participants were assessed on a cognitive, literacy and literacy-related test battery. Single-word reading and spelling were assessed by the Word Reading and Spelling subtests of the WRAT-IV (Wilkinson & Robertson, 2006). General cognitive ability was assessed using the Vocabulary and Matrices subtests of the WRIT (Glutting, Adams, & Sheslow, 2000). Verbal short-term memory and speed of processing ability were assessed with the Digit Span and Symbol Search subtests of the WAIS-III (Wechsler, 1998). Finally, a nonword phoneme deletion task and two Rapid Automatized Naming (RAN) tasks were used to index participants' phonemic processing skill and naming speed.

RAN. Two RAN tasks (digits and objects; Caravolas et al., 2012) were administered. Participants were presented with a quasi-random array of five high frequency items, repeated eight times over five lines of an A4 card, in each condition, and over two trials. They were asked to name all 40 items, sequentially from left to right, as fast as they could. Each trial was timed from the onset of the first to the offset of participants' last response. Response latencies from the two blocks were combined for each participant. The inter-trial correlations were r = .96 for RAN Digits and r = .88 for RAN Objects. Participants' expectedly negligible error rates were less than 1% in all cases and therefore were not analyzed.

Nonword phoneme deletion. Phonemic awareness was measured by a nonword phoneme deletion task from Judge, Caravolas, and Knox (2006). Participants had to identify and delete the second phoneme of 12 CC̲VC (e.g., /stɛk/→ /sɛk/ in block 1) or the penultimate phoneme of 12 CVC̲C (e.g., /fɛsp/→/fɛp/ in block 2) nonwords, and correctly reproduce each resulting nonword as quickly as possible. The maximum accuracy score averaged across blocks was 12. The Cronbach's Alpha reliability was .92. In anticipation of low error rates, latencies were measured and recorded separately for each block from the onset of participants' repetition of the first stimulus to the offset of their last response. The reliability was r = .95.

AGL material

All material (listed in Appendix A) was generated using the grammar shown in Figure 1 and was identical to that used in Kinder and Lotz (2009, Experiment 2). There were two lists of stimuli, the presentation of which was counterbalanced within participant groups[1].

Twenty-four 7-letter strings (12 per counterbalanced list) served as memorization items and ninety-six letter strings (48 per counterbalanced list) served as test-phase items. Half (i.e., 24) of the test-phase items in each list consisted solely of *permissible* (legal) *chunks*, the majority of which appeared *frequently* during memorization (for details on how AGL performance pertains to chunk formation, see Perruchet & Pacton, 2006). The remaining 24 items contained some permissible chunks shown during memorization, but in addition, they all contained *illegal chunks*. Differences between the two sets in each list were quantified in terms of global Associative Chunk Strength (ACS), following Knowlton and Squire (1994; see Appendix B). An independent *t* test confirmed that the mean global ACS of items that contained some illegal chunks ($M = 0.64$, $SD = 0.22$) was significantly lower than the mean global ACS of items consisting completely of legal chunks ($M = 1.27$, $SD = 0.12$), $t(71.68) = 17.14$, $p < .001$, $d = 3.50$.

Letter-position patterns. The 24 (per counterbalanced list) items that contained some illegal chunks were divided into two subsets of items which served to assess participants' sensitivity to letter-position patterns. Half of them violated the

---

[1]There were two list effect in the memorization phase analyses, which suggest that list 2 items were somewhat easier to memorize than list 1 items, however, this finding did not hold in the omnibus analyses reported in Appendix D (and neither of these findings hold in an unpublished dataset using the same materials). Further to the effect not being robust, it is unclear what aspect of stimulus structure may have contributed to this finding (for example, there are no immediate repetitions of elements within any items- this sometimes explains list effects in AGL studies). For these reasons, we have not elaborated on this issue. Turning to the test-phase analyses, the only effect that reached significance was a marginal (p = .049) three-way interaction in the analyses of sensitivity to shape-position patterns in experiment 2. Simple effect analyses correcting for family-wise error again did not reach significance.

grammar in terms of the distribution of adjacent elements, *as well as* the absolute allowable position of letters. For example, the illegal bigram FM within FXDHFML also violates the grammar's constraint that F and M cannot occur in position 5 and 6, respectively. The other subset of items always respected the absolute position of letters set by the grammar, yet contained one or more illegal chunks (e.g., the bigram RB of the string FSKRBWJ is illegal, even though R and B are allowed in position 4 and 5, respectively). The two subsets were not significantly different in the amount of chunk violation (global ACS: $M_{\text{illegal chunks}} = 0.66$, $SD_{\text{illegal chunks}} = 0.24$ vs. $M_{\text{illegal chunks \& positions}} = 0.62$, $SD_{\text{illegal chunks \& positions}} = 0.21$; $t(46) = 0.53$, $p > .05$, $d = 0.15$); thus, differential sensitivity to these items could only be accounted for by the presence/absence of letter-position violations.

Specific similarity. The 24 (per counterbalanced list) items that consisted of legal chunks were divided into two subsets of items which served to assess participants' sensitivity to *specific similarity* (for a thorough review on the concept of similarity, see Pothos, 2005). Half of the items (e.g., FXDHCXJ) deviated from their closest training item (JXDHCWH) by three letters or more, thus, they were dissimilar to training items. The remaining 12 items (e.g., MP*D*RTXL) differed by one letter from their closest training item (MP*V*RTXL), thus, they were specifically similar to one of the memorization items. Differential sensitivity to these subsets could only be attributed to differences in specific similarity, and not differences in global ACS ($M_{\text{dissimilar}} = 1.27$, $SD_{\text{dissimilar}} = 0.11$ vs. $M_{\text{similar}} = 1.27$, $SD_{\text{similar}} = 0.13$).

Procedure

Participants were tested individually on the AGL task followed by the background measures in a 1.5-hour long session.

AGL task. The memorization phase of the experiment (phase 1) was followed by a surprise test-phase judgment task (phase 2). In phase 1, participants were told that they were taking part in a short-term memory task. During each trial, a training string appeared on white background for unlimited time, participants were instructed to memorize it, and press the spacebar when ready. String presentation was followed by a 3000 ms interval (centered black cross) after which participants were presented with 10 individual letters (40 ppt; Arial Font), only 7 of which matched the letters of the "memorized" letter string (*target* letters). Three *distractor* letters were selected randomly without replacement from the pool of consonants which did not comprise the letter string. Participants were asked to recreate the string just memorized by dragging and dropping only the relevant letters into the boxes. The left-right order of the candidate letters was randomly determined for each participant during each trial. There was no time limit and letters could be dragged and dropped (in any order) or rearranged into the boxes until the response was submitted. There was no feedback, however, incorrect trials were repeated.[2] The cycle repeated until all 96 memorization trials (8 repetitions/string) were correctly reproduced. Training trials were presented randomly in a single block; breaks were allowed.

In phase 2, the test-phase judgment task was administered. Participants were informed that all previous letter strings followed hard to unravel rules and they were to decide whether new strings were grammatical or ungrammatical. To avoid extreme "yes"/"no" response biases, participants were informed that only half of the strings were rule-conforming. Each test string was displayed on white background, remained

[2] Due to programming constraints, incorrect responses were paired with identical letter/distractor sets and left-right order arrangements.

on screen until a response was collected, and was followed by a 1000 ms fixation (centered black cross). Stimuli were presented in a single block without feedback.

Data Analyses

Memorization performance was measured by the proportion of strings reproduced correctly within a single attempt, the mean number of trials to criterion (i.e., correct reproduction of all 96 strings), and mean correct memorization RTs (Intercorrelations between these measures are reported in the Supplementary Materials). To assess classification accuracy on the basis of chunk strength, a measure of sensitivity (*d*') was computed by calculating the difference between the *z*-transformed proportion of "yes" responses to sequences that did not contain illegal chunks (hits) and the *z*-transformed proportion of "yes" responses to sequences that contained illegal chunks (false alarms, FAs). Rates of 0 were replaced with 1/2*n* where n corresponds to the number of signal or noise trials, respectively. Rates of 1 were replaced with 1 – 1/2*n* (Macmillan & Kaplan, 1985). Separate measures of *d*' sensitivity were calculated for strings containing chunk and letter-position violations ($z(\text{hits}) - z(\text{FAs}_{\text{illegal chunks \& positions}})$) relative to strings that contained chunk violations only ($z(\text{hits}) - z(\text{FAs}_{\text{illegal chunks only}})$) and were compared to investigate participants' sensitivity to letter-position patterns. Separate measures of *d*' sensitivity were also calculated for specifically similar strings ($z(\text{hits}_{\text{similar}}) - z(\text{FAs})$) relative to dissimilar strings ($z(\text{hits}_{\text{dissimilar}}) - z(\text{FAs})$) and were compared to investigate participants' sensitivity to specific similarity. Intercorrelations between these measures are also reported in the Supplementary Materials.

Bayes factor (BF) analyses were carried out on all critical reader group comparisons associated with nonsignificant *p* values to quantify evidence for/against the theory that skilled and dyslexic readers will perform comparably on the AGL task.

We used the BayesFactor package, version 0.9.11–1 (Morey & Rouder, 2015) for the R software package (R Development Core Team, 2015). We estimated BFs using the Jeffrey–Zellner–Siow priors with a default scaling factor for the predictions of the alternative hypothesis and Monte-Carlo sampling for each model of interest (500,000 iterations). BFs were computed by comparing a full/more complex model (e.g., a model that includes a main effect) to a simpler one (e.g., null model) are denoted as B_{10}. BFs that were computed by comparing a simpler model (e.g., main effect(s), no interaction) to a more complex one (e.g., main effect(s) and interaction) are denoted as B_{01}. Following Jeffreys' (1961) convention, values of 0.33 and below were taken to suggest *evidence for* the model in the numerator relative to the model in the denominator; values of 3 and above were considered as *evidence against* the model in the numerator relative to the model in the denominator; values between 0.33 and 3 were interpreted as *no evidence* for either model (i.e. inconclusive evidence).

Results

Background measures

Skilled and dyslexic readers' performance on the cognitive and literacy measures is shown in Table 1. The groups were matched for age, Vocabulary and Matrices task performance. Dyslexic participants performed within the average range on both WRAT Reading and Spelling subtests (i.e., they were relatively high functioning) but, as a group, were significantly impaired relative to the skilled readers, and unlike their peers, they showed significant discrepancies between their general abilities and their literacy skills. Furthermore, the dyslexic group was impaired on Symbol Search, Digit Span, phonemic awareness and rapid naming skill, the latter three measures being behavioral markers of a phonological processing impairment typical in dyslexia (Snowling & Hulme, 2012).

------------Table 1------------

AGL

Memorization phase. Readers with dyslexia were less accurate (85%, *SD* = 0.10) in correctly reproducing the training strings within a single attempt than skilled readers (91%, *SD* = 0.07), *t*(29.76) = 2.23, *p* = .033, *d* = 0.67, but did not require significantly more trials to reach criterion (*M* = 1.21, *SD* = 0.16) when compared to skilled readers (*M* = 1.13, *SD* = 0.11), *t*(27.91) = 1.98, *p* = .058, *d* = 0.60. Similarly, there was no significant difference in correct memorization RTs (trimmed to the values of 2 *SD*s from each group mean), *t*(48) = 0.77, *p* = .448, *d* = 0.23, between skilled (*M* =5707.99, *SD* = 2382.93) and dyslexic readers (*M* = 6204.37, *SD* = 1933.96).

Bayes factor analyses were carried out to establish whether the data (number of trials to criterion; memorization response latencies) supported the model that included the effect of group over the null model. They indicated data insensitivity (B_{10} = 1.78 and 0.37, respectively). Thus, it is not possible to draw strong conclusions on the basis of these comparisons alone about whether string memorization posed a greater challenge to dyslexic students.

Test phase

Chunk strength sensitivity. Discrimination ability between skilled and dyslexic readers (Table 2) was not statistically different, *t*(48) = 0.23, *p* = .817, *d* = 0.07. Both groups performed significantly better than chance, skilled readers: *t*(30) =

10.31, $p < .001$, $d = 1.85$; dyslexic readers: $t(18) = 7.46$, $p < .001$, $d = 1.71$ and comparably to each other ($B_{10} = 0.296$).

------------Table 2------------

***Sensitivity to letter-position patterns*.** To investigate whether participants were sensitive to letter-position patterns, *d*'s for the two types of strings that contained some illegal chunks (Table 2) were entered into a mixed factorial ANOVA with type of violation as a within-subject variable and group as a between-subjects variable[3]. This showed a significant main effect of type of violation, $F(1, 48) = 11.84$, $p = .001$, $\eta^2 = .20$, such that participants were better at detecting ungrammaticality for strings that contained chunk violations and letter-position violations ($M = 1.09$, $SE = 0.09$) relative to strings that contained chunk violations only ($M = 0.84$, $SE = 0.09$). There was no interaction and no effect of group, $Fs(1, 48) < 1$. The model without the interaction term was preferred to the model that included the group by type of violation interaction ($B_{01} = 3.290$), which suggests that sensitivity to letter-position patterns was comparable in skilled and dyslexic readers.

***Specific similarity sensitivity*.** To assess participants' sensitivity to specific similarity, *d*'s for the two types of strings that consisted of legal chunks (Table 2) were subjected to a 2-way similarity by group ANOVA[4]. This revealed a significant main effect of similarity, $F(1, 48) = 8.35$, $p = .006$, $\eta^2 = .14$, with higher *d*'s for

[3]Including participants' memorization accuracy (number of trials to criterion) as a covariate to statistically control for differences in skilled and dyslexic readers' memorization performance did not alter the pattern of results in these analyses (or, in fact, any of the test-phase analyses reported in the manuscript).

[4]One covariate that influenced participants' sensitivity to specific similarity was their nonverbal IQ (i.e., WRIT Matrices subtest performance). The ANCOVA on *d*'s for strings that consisted of legal chunks showed a significant interaction between type of item and the covariate ($p = .005$), which suggests that sensitivity to specific similarity depended on participants' nonverbal IQ. This effect held across groups and was not replicated in Experiment 2, thus, we do not further elaborate on it.

similar (M = 1.08, SE = 0.09) than dissimilar items (M = 0.86, SE = 0.08). Neither the interaction, $F(1, 47) = 1.61$, $p = .211$, $\eta^2 = .03$, nor the main effect of group, $F(1, 48) < 1$, was significant. However, the Bayes factor between the model with and without the group by similarity interaction was between 1/3 and 3 (B_{01} = 2.643) indicating that the nonsignificant interaction was, in fact, insensitive.

Discussion

Experiment 1 assessed AGL performance in dyslexic and nondyslexic adults by means of a test-phase judgment task preceded by memorization and correct reproduction of training letter strings. There was some evidence that dyslexic participants were less accurate in string memorization relative to skilled readers, yet it was not possible to draw strong conclusions regarding memorization response latencies.

Test-phase analyses of *d*' scores confirmed that skilled and dyslexic readers reliably discriminated between items that did and did not contain some illegal chunks. Importantly, chunk sensitivity led to *comparable* levels of discrimination ability in skilled and dyslexic adults. Regarding skilled and dyslexic readers' sensitivity to letter-position patterns and specific similarity, we replicated Kinder and Lotz (2009) by showing that both factors guided classification performance in the task (replicating Kinder & Lotz, 2009), and conclusively demonstrated that sensitivity to letter-position patterns influenced skilled and dyslexic adult readers *to the same extent*. Some uncertainty remains, however, regarding potential group differences in sensitivity to specific similarity.

Experiment 2

An important question in dyslexia research is whether dyslexics' well-known impairment in processing alphanumeric stimuli (e.g., Wolf, Bally, & Morris, 1986;

Ziegler, Pech-Georgel, Dufau, & Grainger, 2010), is domain-specific or signals a domain-general learning deficit. This was examined in experiment 2 which employed a nonlinguistic AGL variant.

Rationale

Experiment 2 was directly analogous to experiment 1 in all respects but stimulus format: letter strings were replaced by sequences of novel nonlinguistic symbols, which were unfamiliar and therefore did not map onto any specific verbal information (stimuli are shown in Appendix A). We investigated whether dyslexic participants experience difficulties during the training phase when learning nonalphanumeric patterns, suggesting a generalised implicit learning weakness; and whether, under this stimulus format, dyslexic and control groups demonstrate differential sensitivity to chunk strength, symbol position patterns and specific similarity.

Method

Participants

Thirty-one skilled readers (12 male; mean age = 20.60 years) and 21 dyslexic students (9 male; mean age = 20.54 years) naïve to the study, took part in experiment 2. Six additional participants (1 skilled and 5 dyslexic readers) who failed to complete the memorization phase of the task within 75 minutes were not considered in any analyses. The same inclusion criteria as in experiment 1 applied.

Background Measures

The same background measures were used as in experiment 1.

Material and Procedure

The letter strings used in experiment 1 were converted to shape strings by mapping each of the 20 letters to an abstract easily distinguishable shape from Taylor, Plunkett, and Nation (2011). We did not opt for geometrical shapes to prevent, as much as possible, participants from adopting a verbal encoding strategy. All other aspects of the study were identical to experiment 1.

Results

Background Measures

As in experiment 1, skilled and dyslexic readers (Table 3) did not differ significantly in terms of age, verbal and nonverbal IQ. There was, again, some overlap in literacy performance between the two reader groups, however, only the individuals with dyslexia experienced an IQ-literacy skills discrepancy. Dyslexic participants performed within the average range of the WRAT Reading and Spelling subtests but significantly worse than skilled readers. As in experiment 1, their performance indicated significant phonological processing difficulties with significantly lower scores than skilled readers, and large effect sizes in Digit Span, phonemic awareness and RAN digits, but not RAN objects, p = .068, or Symbol Search.

------------Table 3------------

AGL Task

Memorization phase. No significant difference emerged between skilled (M = 83%, SD = 0.11) and dyslexic readers' (M = 83%, SD = 0.12) ability to correctly reproduce the shape sequences within a single attempt, $t(50)$ = 0.12, p = .903, d = 0.03, B_{10} = 0.284. The mean number of trials to criterion did not differ for skilled (M = 1.22, SD = 0.15) and dyslexic readers (M = 1.22, SD = 0.17), $t(50)$ = 0.14, p = .889, d = 0.04, B_{10} = 0.285, nor did memorization reaction times (skilled: M = 10544.46, SD

= 3100.80, dyslexic: M = 10573.41, SD = 2966.43), $t(47) = 0.03$, $p = .973$, $d = 0.01$, $B_{10} = 0.283$). In sum, in experiment 2, dyslexic readers' memorization performance matched skilled readers' performance.

Test phase

Chunk strength sensitivity. Dyslexic readers' discrimination ability (Table 4) was not significantly different from that of skilled readers, $t(50) = 0.63$, $p = .531$, d = 0.18. Mean d' values reliably exceeded chance for skilled readers, $t(30) = 8.09$, $p < .001$, $d = 1.45$, and dyslexic readers, $t(20) = 7.08$, $p < .001$, $d = 1.54$, and the Bayes factor provided some (albeit weak) support for the null hypothesis, $B_{10} = 0.333$.

------------Table 4------------

Sensitivity to shape position patterns. A type of violation by group ANOVA on d's for strings that contained some illegal chunks showed a main effect of type of violation, $F(1, 50) = 29.50$, $p < .001$, $\eta^2 = .36$, reflecting higher sensitivity to the ungrammaticality of strings that contained chunk and symbol-position violations (M = 1.02, SE = 0.10) relative to strings that contained chunk violations only (M = 0.53, SE = 0.07). There was no interaction, $F(1, 50) = 2.14$, $p = .150$, $\eta^2 = .03$, or main effect of group, $F(1, 50) < 1$. The Bayes factor between the model with and without the interaction term (B_{01}) was 2.104, indicating lack of sensitivity.

Specific similarity sensitivity. A similarity by group ANOVA on strings that consisted completely of legal chunks (Table 4) revealed a significant main effect of similarity, $F(1, 50) = 9.48$, $p = .003$, $\eta^2 = .16$, with higher sensitivity to the grammaticality of specifically similar strings (M = 0.89, SE = 0.08) than dissimilar strings (*M* = 0.64, *SE* = 0.08). There was no effect of group, $F(1, 50) < 1$, or group by similarity interaction, $F(1, 50) < 1$. The model without the interaction term was preferred to the model that included the group by similarity interaction ($B_{01} = 3.364$). Thus, sensitivity to specific similarity was comparable in skilled and dyslexic readers.

Discussion

Experiment 2 investigated AGL of symbol sequences that do not map onto specific verbal information. This manipulation allowed us to evaluate whether (a) memorization performance and (b) sensitivity to different patterns which may map onto orthographic learning constructs would be modulated by stimulus format, and whether any such differences might be moderated by reading ability. Importantly, in a direct comparison of the cognitive and literacy scores of participants in experiments 1 and 2 (Appendix C), the groups were not statistically different, precluding the likelihood that differences in results were attributable to differences in background profiles.

In sum, in contrast to the previous experiment, it appeared that skilled and dyslexic readers were comparable in all aspects of memorization performance. To statistically support this claim, we examined the stimulus type (letter vs. symbol strings) by group (skilled vs. dyslexic readers) interactions in the omnibus ANOVAs on all three measures of memorization performance. These were all nonsignificant and there was no significant effect of group in any of the analyses (Appendix D), yet, only the speed of correct responding during training was conclusively similar between readers groups regardless of stimulus format. The aggregated analysis on the measures of memorization accuracy proved inconclusive. We return to the issues that these mixed findings raise in the General Discussion.

In terms of test-phase performance, Bayes factor analyses confirmed that skilled and dyslexic readers were comparably influenced by the manipulation of chunk strength sensitivity and specific similarity. Both groups could differentiate between stimuli that did and did not adhere to letter-position patterns, yet it was not possible to strongly conclude whether groups differed in this ability.

General Discussion

Previous research investigating implicit skill learning in dyslexia has been inconclusive, reporting unimpaired (e.g., Rüsseler et al., 2006) as well as impaired (e.g., Pavlidou et al., 2010) performance on different implicit learning tasks. These findings are inconsistent across age groups as well as reading ability groups. Here, we examined implicit knowledge acquisition in skilled versus dyslexic adults by means of an AGL task. We adapted Kinder and Lotz's (2009) task variant that required the reconstruction of letter strings using a "drag and drop" procedure instead of a typing-from-memory procedure; we also adapted the stimulus format to shape strings in experiment 2. In both experiments, participants memorized the stimuli and subsequently, undertook a test-phase judgment task assessing sensitivity to chunk strength, positional constraints on letters/symbols, and specific similarity of letter/shape strings.

Going beyond previous studies, our comparison of dyslexic and typical readers was considered not only at test but also during stimulus memorization. The series of cross-experiment analyses and relevant Bayes factor analyses clearly indicated that students with and without dyslexia responded similarly quickly to correct learning trials. Some of the analyses, however, lacked robustness. For example, the indication that students with dyslexia were, on average, less accurate in memorizing letter strings (experiment 1), was not confirmed in the cross-experiment Bayes factor analyses.

These inconclusive results highlight the need for further exploration of the interesting possibility that dyslexic readers may have subtle implicit learning deficits specifically for letters strings (and possibly all symbols mapping onto phonological codes; Wolf et al., 1986).

The analyses of primary interest were those assessing test-phase performance. Across experiments, these revealed little evidence of impairment among dyslexic adults. Whether the stimuli were letters or shapes, both groups demonstrated reliable levels of sensitivity to the three types of attributes embedded within the grammar's strings. Moreover, dyslexic and typical readers' level of sensitivity to chunk strength, was *comparable*, suggesting that, what dyslexic individuals learn about orthographic patterns within an artificial system is generally no different to what their nondyslexic peers learn. Consistent with this, Bourassa and Treiman (2003) have reported similar patterns of misspellings among dyslexic and typical child spellers.

Our findings are not consistent with the notion that a general implicit learning deficit is causally related to reading and spelling disability. The demonstration of unimpaired implicit chunk strength sensitivity—a frequency-based sensitivity—in dyslexia is important for ruling out a statistical learning deficit interpretation of dyslexic individuals' literacy difficulties. The ability to process words in terms of common letter chunks or spelling units (e.g., _ell as in b*ell*, s*ell*, t*ell*) is widely acknowledged as an important aspect of skilled word reading (e.g., Ehri, 1995, 1997) and spelling, and may be attributed at least in part to implicit skill learning (e.g., Nation, 1997). The present study demonstrates that, at least in compensated dyslexic adults learning under laboratory conditions, the ability to detect chunk frequency information—be it for letters or shapes—is completely normal in dyslexia. As pointed out earlier, however, well-compensated, high functioning dyslexic adults comprise a highly heterogeneous group. A challenge for future research will be to determine whether a subtle frequency-based learning impairment exists among less compensated (e.g., nonuniversity dyslexic populations) or younger learners with dyslexia. It is also an empirical question whether skilled and compensated dyslexic readers' implicit learning abilities are qualitatively different (e.g., associated with different domain-general learning abilities).

Turning to sensitivity to letter-position patterns and specific similarity learning, our results suggest reliable sensitivity to both of these factors across groups and stimulus types. However, the present study could not provide strong evidence for or against relative differences in the levels of sensitivity to these attributes among typical and dyslexic readers. In both cases, the results were inconclusive in one of the two experiments leaving open the possibility that adults with dyslexia may find these aspects of orthographic learning relatively more difficult than typical readers. As both of these effects were notably weaker than those assessing chunk strength sensitivity, paradigms other than the AGL may be more suitable to reliably compare dyslexic and control groups' ability to benefit from these frequency-based orthographic cues. In addition to seeking more conclusive results, further manipulations of AGL and other paradigms could effectively address other interesting questions about these important frequency-based cues. Would constraints on the allowable position of individual letters have been reliably extracted had all positional violations occurred in medial (i.e. nonanchor) positions? Is learning by analogy important for literacy development in reading unimpaired/impaired individuals or was the effect of specific similarity in the AGL task a by-product of the lengthy training phase duration adopted in our study (Kinder, 2000)? These important questions get to the core of the complexities of real orthographic learning.

Several factors, over and above random noise effects, may account for differences between the current and previous studies. Heterogeneity of dyslexic samples is one likely explanation of the inconsistent findings in the literature. For example, in contrast to some previous research, we were able to ascertain that individuals in our dyslexia groups did not have comorbid disorders, e.g. impaired attention in the context of Attention Deficit Disorder, which could additively impair AGL performance (e.g., Tanaka, Kiyokawa, Yamada, Dienes, & Shigemasu, 2008). In line with typical university samples, dyslexic individuals in our study were relatively high functioning or well-compensated in terms of reading and spelling accuracy, therefore, not representative of all individuals with dyslexia in terms of literacy levels. Nevertheless, despite their relative cognitive strengths, their reading, spelling and phonological processing weaknesses were well below those of their university peers, and, reflected typical dyslexia profiles. While it is possible that our well-compensated participants had better implicit learning abilities relative to participants in previous studies, this is, in our view, unlikely.

It is also possible, and remains a question for further empirical research, that stimulus type, a factor we manipulated directly, may also account for some discrepant findings in the literature. In our study, separate dyslexic and control groups showed somewhat differential implicit learning patterns as a function of stimulus-based factors rather than participant-based factors (on which the two samples were very similar, see Appendix C). Differences as a function of literacy status, however, were subtle and generally tentative. A possible exception was the memorization accuracy of grammatical letter strings, which appeared to be relatively more compromised among dyslexics than was symbol memorization. Thus, stimulus type may play a significant role in AGL performance, and dyslexic individuals' implicit learning difficulties, to the extent that they exist, may be amplified when learning alphanumeric materials. It should be noted, however, that memorization of shape strings to criterion was prohibitively lengthy for five dyslexic participants who failed to complete the task. Post hoc inspection of their clinical profiles revealed that most had severe additional disorders, at least some of which were likely to explain their increased learning difficulty (e.g., clinical depression, motor coordination difficulties, developmental language delay, obsessive compulsive disorder). It is nevertheless conceivable that we may have excluded a few very impaired dyslexic (implicit) learners whose performance might have influenced the outcomes of experiment 2. Given the challenging nature of our lengthy training phase, this issue would have to be resolved in a future study that used fewer or a fixed numbers of training trials. Another avenue for future research may be to investigate whether varying the demands of the training phase procedure has an effect on dyslexic individuals' AGL performance. For example, would our findings replicate if participants were asked to type rather than "drag and drop" letter strings during memorization?

Conclusion

An orthographic frequency-based learning deficit has been sometimes considered to provide a parsimonious account of dyslexic individuals' difficulties in literacy and other literacy-related skills. For example, according to Sperling et al. (2004), an implicit learning deficit may result in less automatized decoding skill or reduced implicit orthographic pattern acquisition in dyslexia. The only tentative evidence suggestive of impairment among dyslexic participants in the present studies involved memorizing letter strings. There was no evidence that dyslexic adults were insensitive to *either* chunk strength *or* the more subtle aspects of the grammar (letter positions and whole exemplars) and, most notably, frequency-based, chunk strength sensitivity was comparable across groups in both experiments. Future studies with larger sample sizes and perhaps with other test paradigms may be more decisive regarding the relative sensitivities of typical versus dyslexic readers to the remaining two features of the grammar. Overall, our findings weaken the claim that deficits in implicit skill learning are a direct cause of literacy impairment. Future studies should investigate whether these findings generalize to dyslexic children, nonuniversity dyslexic populations, and adults in the lower ability end of the dyslexia spectrum.

References

Bourassa, D., & Treiman, R. (2003). Spelling in children with dyslexia: Analyses from the Treiman-Bourassa early spelling test. *Scientific Studies of Reading, 7*, 309-333. doi: 10.1207/S1532799XSSR0704_1

Brooks, L. R., & Vokey, J. R. (1991). Abstract analogies and abstracted grammars: Comments on Reber (1989) and Mathews et al. (1989). *Journal of Experimental Psychology: General, 120,* 316-323. doi: 10.1037/0096-3445.120.3.316

Caravolas, M., Lervåg, A., Mousikou, P., Efrim, C., Litaviský, M., Onochie-Quintanilla, E., … Hulme, C. (2012). Common patterns of prediction of literacy development in different alphabetic orthographies. *Psychological Science, 23*, 678-686. doi: 10.1177/0956797611434536

Cassar, M., & Treiman, R. (1997). The beginnings of orthographic knowledge: Children's knowledge of double letters in words. *Journal of Educational Psychology, 89*, 631-644. doi: 10.1037/0022-0663.89.4.631

Ehri, L. C. (1995). Phases of development in learning to read words by sight. *Journal of Research in Reading, 18,* 116-125. doi: 10.1111/j.1467-9817.1995.tb00077.x

Ehri, L. C. (1997). Learning to read and learning to spell are one and the same, almost. In C. Perfetti, L. Rieben, & M. Fayol (Eds.), *Learning to spell: Research, theory and practice* (pp. 237-269). Mahwah, NJ: Erlbaum.

Glutting, J. J., Adams, W., & Sheslow, D. (2000). *Wide Range Intelligence Test Manual*. Wilmington, DE: Wide Range.

Gombert, J. E. (2003). Implicit and explicit learning to read: implication as for subtypes of dyslexia. *Current Psychology Letters, 10*.

Goswami, U. (1988). Orthographic analogies and reading development. *The Quarterly Journal of Experimental Psychology, 40A*, 239-268. doi: 10.1080/02724988843000113

Ise, E., Arnoldi, C. J., Bartling, J., & Schulte-Körne, G. (2012). Implicit learning in children with spelling disability: Evidence from artificial grammar learning. *Journal of Neural Transmission, 119*, 999-1010. doi: 10.1007/s00702-012-0830-y

Jeffreys, H. (1961). *Theory of probability* (3rd ed.). Oxford: Oxford University Press, Clarendon Press.

Johnstone, T., & Shanks, D. R. (1999). Two mechanisms in implicit grammar learning? Comment on Meulemans and Van der Linden (1997). *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*, 524-531. doi: 10.1037/0278-7393.25.2.524

Judge, J., Caravolas, M., & Knox, P. C. (2006). Smooth pursuit eye movements and phonological processing in adults with dyslexia. *Cognitive Neuropsychology, 23*, 1174-1189. doi: 10.1080/02643290600785931

Kahta, S., & Schiff, R. (2016). Implicit learning deficits among adults with developmental dyslexia. *Annals of Dyslexia, 66*, 235-250. doi: 10.1007/s11881-016-0121-7.

Kinder, A. (2000). The knowledge acquired during artificial grammar learning: Testing the predictions of two connectionist models. *Psychological Research, 63*, 95-105. doi: 10.1007/s004260000038

Kinder, A, & Lotz, A. (2009). Connectionist models of artificial grammar learning: what type of knowledge is acquired? *Psychological Research, 73*, 659-673. doi: 10.1007/s00426-008-0177-z

Knowlton, B. J., & Squire, L. R. (1994). The information acquired during artificial grammar learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*, 79-91. doi: 10.1037/0278-7393.20.1.79

Knowlton, B. J., & Squire, L. R. (1996). Artificial grammar learning depends on implicit acquisition of both abstract and exemplar-specific information. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*, 169-181. doi: 10.1037/0278-7393.22.1.169

Laasonen, M., Väre, J., Oksanen-Hennah, H., Leppämäki, S., Tani, P., Harno, H., ... & Cleeremans, A. (2014). Project DyAdd: Implicit learning in adult dyslexia and ADHD. *Annals of dyslexia, 64*, 1-33. doi: 10.1007/s11881-013-0083-y

Macmillan, N. A., & Kaplan, H. L. (1985). Detection theory analysis of group data: Estimating sensitivity from average hit and false-alarm rates. *Psychological Bulletin, 98*, 185-199. doi: 10.1037/0033-2909.98.1.185

Morey, R. D., & Rouder, J. N. (2015). BayesFactor: Computation of Bayes factors for common designs (Version 0.9.11-1) [Software]. Retrieved from http://bayesfactorpcl.r-forge.r-project.org/

Nation, K. (1997). Children's sensitivity to rime unit frequency when spelling words and nonwords. *Reading and Writing: An Interdisciplinary Journal, 9*, 321-338. doi: 10.1007/978-94-017-3054-9_2

Nigro, L., Jiménez-Fernández, G., Simpson, I. C., & Defior, S. (2015). Implicit learning of non-linguistic and linguistic regularities in children with dyslexia. *Annals of Dyslexia, 66*, 202-218. doi: 10.1007/s11881-015-0116-9

Pavlidou, E. V., Kelly, L. M, & Williams, J. M. (2010). Do children with developmental dyslexia have impairments in implicit learning? *Dyslexia, 16*, 143-161. doi: 10.1002/dys.400

Perruchet, P., & Pacteau, C. (1990). Synthetic grammar learning: Implicit rule abstraction or explicit fragmentary knowledge. *Journal of Experimental Psychology: General, 119*, 264-275. doi: 10.1037/0096-3445.119.3.264

Perruchet, P., & Pacton, S. (2006). Implicit learning and statistical learning: One phenomenon, two approaches. *Trends in Cognitive Sciences, 10(5)*, 233-238. doi:10.1016/j.tics.2006.03.006.

Pothos, E. M. (2005). The rules versus similarity distinction. *Behavioral and Brain Sciences, 28(01)*, 1-14. http://dx.doi.org/10.1017/S0140525X05000014

Pothos, E. M. (2007). Theories of artificial grammar learning. Psychological Bulletin, 133, 227-244. doi: 10.1037/0033-2909.133.2.227

Pothos, E. M., & Kirk, J. (2004). Investigating learning deficits associated with dyslexia. *Dyslexia, 10*, 61-76. doi: 10.1002/dys.266

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

Reber, A. S. (1989). Implicit learning and tacit knowledge. *Journal of Experimental Psychology: General, 118*, 219-235. doi: 10.1037/0096-3445.118.3.219

Rüsseler, J., Gerth, I., & Münte, T. (2006). Implicit learning is intact in adult developmental dyslexic readers: Evidence from the serial reaction time task and artificial grammar learning. *Journal of Clinical and Experimental Neuropsychology, 28*, 808-827. doi: 10.1080/13803390591001007

Samara, A., & Caravolas, M. (2014). Statistical learning of novel graphotactic constraints in children and adults. *Journal of Experimental Child Psychology, 121*, 137-155. doi: 10.1016/j.jecp.2013.11.009

Snowling, M. J., & Hulme, C. (2012). Annual Research Review: The nature and classification of reading disorders–a commentary on proposals for DSM-5. *Journal of Child Psychology and Psychiatry, 53,* 593-607. doi: 10.1111/j.1469-7610.2011.02495.x

Sperling, A. J., Lu, Z. L., & Manis, F. R. (2004). Slower implicit categorical learning in adult poor readers. *Annals of Dyslexia, 54,* 281-303. doi: 10.1007/s11881-004-0014-z

Tanaka, D., Kiyokawa, S., Yamada, A., Dienes, Z., Shigemasu, K. (2008). Role of selective attention in artificial grammar learning. *Psychonomic Bulletin & Review, 15*, 1154-1159. doi: 10.3758/PBR.15.6.1154

Taylor, J. S. H., Plunkett, K., & Nation, K. (2011). The influence of consistency, frequency, and semantics on learning to read: An artificial orthography paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37*, 60-76. doi: 10.1037/a0020126

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

Tremblay, S., Macken, W. J., & Jones, D. M. (2000). Elimination of the word length effect by irrelevant sound revisited. *Memory & Cognition, 28*, 841-846. doi: 10.3758/BF03198419

Wechsler, D. (1998). *Wechsler Adult Intelligence Scale – Third Edition*. London: The Psychological Corporation.

Wilkinson, G. S., & Robertson, G. J. (2006). *The Wide Range Achievement Test–4*. Lutz, FL: Psychological Assessment Resources.

Wolf, M., Bally, H., & Morris, R. (1986). Automaticity, retrieval processes, and reading: A longitudinal study in average and impaired readers. *Child Development, 57,* 988-1000. doi: 10.1111/j.1467-8624.1986.tb00260.x

Wright, R. L., & Whittlesea, B. W. A. (1998). Implicit learning of complex structures: Active adaptation and selective processing in acquisition and application. *Memory & Cognition, 26*, 402-420. doi: 10.3758/BF03201149

Ziegler, J. C., Pech-Georgel, C., Dufau, S., & Grainger, J. (2010). Rapid processing of letters, digits and symbols: What purely visual-attentional deficit in developmental dyslexia? *Developmental Science*, *13*, F8–F14. doi: 10.1111/j.1467-7687.2010.00983.x

Table 1

Mean scores (SDs) and Ranges on the Cognitive and Literacy Related Measures as a Function of Group and Results from Independent t tests (Experiment 1).

Variable	Skilled readers (n = 31)		Dyslexic readers (n = 19)		t	sig.	Cohen's d
	Mean (*SD*s)	Range	Mean (*SD*s)	Range			
Handedness (r; l)	25; 6		14; 5				
Age (years)	20.49 (3.47)	18.25 – 37.83	20.80 (3.41)	18.33 – 33.17	0.31	.756	0.09
WRIT Vocabulary[a]	102.90 (9.42)	82.00 – 119.00	103.11 (10.08)	84.00 – 119.00	0.07	.943	0.02
WRIT Matrices[a]	105.55 (11.36)	82.00 – 130.00	111.11 (10.84)	83.00 – 125.00	1.71	.094	0.50
WRAT Reading[a]	105.97 (7.60)	96.00 – 130.00	92.05 (8.68)	77.00 – 105.00	5.95	***	1.71
WRAT Spelling[a]	110.52 (9.20)	95.00 – 129.00	93.63 (9.77)	78.00 – 113.00	6.15	***	1.78
WAIS Digit Span[b]	11.00 (2.86)	4.00 – 16.00	8.26 (2.58)	6.00 – 17.00	3.38[e]	.001	1.00
WAIS Symbol Search[b]	13.67(2.60)	9.00 – 19.00	11.84 (2.87)	6.00 – 19.00	2.30[e]	.026	0.67
RAN digits mean time[c]	14.37 (2.46)	9.57 – 20.60	18.72 (4.68)	12.05 – 27.72	3.74[f]	.001	1.16

RAN objects mean time[c]	21.05 (3.03)	15.72 – 29.38	25.30 (5.27)	17.47 – 37.77	3.20[f]	.004	0.99
NWPD accuracy[d]	10.94 (0.88)	9.00 – 12.00	9.45 (2.08)	5.50 – 12.00	2.96[f]	.007	0.93
NWPD latencies[c]	38.81 (8.40)	25.49 – 60.35	69.16 (19.55)	30.00 – 109.91	6.41[f]	***	2.02

Note. r = right-handed. l = left-handed. WRIT = Wide Range Intelligence Test. WRAT = Wide Range Achievement Test. WAIS = Wechsler Adult Intelligence Scale. RAN= Rapid Automatized Naming. NWPD = NonWord Phoneme Deletion.

[a]Standard Scores. [b]Scaled scores. [c]In seconds. [d]Out of 12. [e]Skilled readers: $n = 30$ due to missing data from one participant..

[f]Correction for unequal variances applied.

*** $p < .001$

Table 2

Skilled and Dyslexic Readers' Sensitivity (d's ± 95% CIs) to the Grammaticality/Ungrammaticality of (i) Strings Containing Chunk and Letter-position Violations, (ii) Strings Containing Chunk Violations Only, (iii) Specifically Similar Strings, (iv) Dissimilar Strings, and (v) Overall Chunk Strength Sensitivity in Experiment 1.

Group	Strings that contained some illegal chunks		Strings that consisted of legal chunks		Chunk strength sensitivity
	strings containing chunk & letter-position violations	strings containing chunk violations only	specifically similar strings	dissimilar strings	
Skilled readers	1.13 ± 0.20	0.83 ± 0.21	0.84 ± 0.21	1.15 ± 0.21	0.96 ± 0.18
Dyslexic readers	1.05 ± 0.29	0.84 ± 0.26	0.88 ± 0.25	1.00 ± 0.29	0.93 ± 0.24

Note. Skilled readers: *n* = 31. Dyslexic readers: *n* =19.

CIs = Confidence Intervals.

Table 3

Mean scores (SDs) and Ranges on the Cognitive and Literacy Related Measures as a Function of Group and Results from Independent t tests (Experiment 2).

Variable	Skilled readers (*n* = 31)		Dyslexic readers (*n* = 20)		*t*	sig.	Cohen's *d*
	Mean (*SD*s)	Range	Mean (*SD*s)	Range			
Handedness (r; l; a)	28; 3; 0		17; 3; 1				
Age (years)	20.60 (2.49)	18.25 – 31.25	20.54 (2.23)	18.42 – 27.92	0.10	.923	0.03
WRIT Vocabulary[a]	105.19 (11.40)	80.00 – 128.00	102.15 (8.75)	85.00 – 122.00	1.02	.315	0.30
WRIT Matrices[a]	101.87 (11.38)	81.00 – 127.00	103.65 (13.58)	83.00 – 130.00	0.51	.616	0.14
WRAT Reading[a]	104.35 (5.35)	95.00 – 116.00	97.90 (7.17)	86.00 – 114.00	3.68	.001	1.02
WRAT Spelling[a]	109.13 (8.81)	95.00 – 129.00	97.65 (7.23)	85.00 – 114.00	4.86	***	1.42
WAIS Digit Span[b]	10.43 (2.97)	6.00 – 17.00	8.35 (1.66)	6.00 – 11.00	3.17[e,f]	.003	0.86
WAIS Symbol Search[b]	13.81 (2.41)	10.00 – 19.00	12.95 (3.28)	7.00 – 19.00	1.07	.289	0.30
RAN digits mean time[c]	13.97 (2.63)	7.80 – 19.95	17.70 (4.43)	11.80 – 27.95	3.40[f]	.002	1.02

RAN objects mean time[c]	22.28 (3.33)	17.27 – 31.50	24.91 (5.59)	17.45 – 34.58	1.90[f]	.068	0.57
NWPD accuracy[d]	10.89 (1.11)	7.50 – 12.00	10.33 (1.20)	8.00 – 12.00	1.71	.093	0.49
NWPD latencies[c]	38.91 (9.56)	17.52 – 72.77	54.36 (15.67)	34.83 – 94.13	3.96[f]	***	1.19

Note. Dyslexic readers: n =20 due to missing background data from one participant. r = right-handed. l = left-handed. a = ambidextrous. WRIT = Wide Range Intelligence Test. WRAT = Wide Range Achievement Test. WAIS = Wechsler Adult Intelligence Scale. RAN= Rapid Automatized Naming. NWPD = Nonword Phoneme Deletion.

[a]Standard Scores. [b]Scaled scores. [c]In seconds. [d]Out of 12. [e]Skilled readers: n = 27 due to missing data from one participant.

[f]Correction for unequal variances applied.

*** $p < .001$.

Table 4

Skilled and Dyslexic Readers' Sensitivity (d's ± 95% CIs) to the Grammaticality/Ungrammaticality of (i) Strings Containing Chunk and Shape-position Violations, (ii) Strings Containing Chunk Violations Only, (iii) Specifically Similar Strings, (iv) Dissimilar Strings, and (v) Overall Chunk Strength Sensitivity in Experiment 2.

Group	Strings that contained some illegal chunks		Strings that consisted of legal chunks		Chunk strength sensitivity
	strings containing chunk & shape-position violations	strings containing chunk violations only	specifically similar strings	dissimilar strings	
Skilled readers	1.03 ± 0.22	0.40 ± 0.19	0.60 ± 0.19	0.83 ± 0.19	0.71 ± 0.17
Dyslexic readers	1.01 ± 0.35	0.65 ± 0.19	0.68 ± 0.26	0.95 ± 0.28	0.80 ± 0.22

Note. Skilled readers: n = 31. Dyslexic readers: n = 21.

CIs = Confidence Intervals.

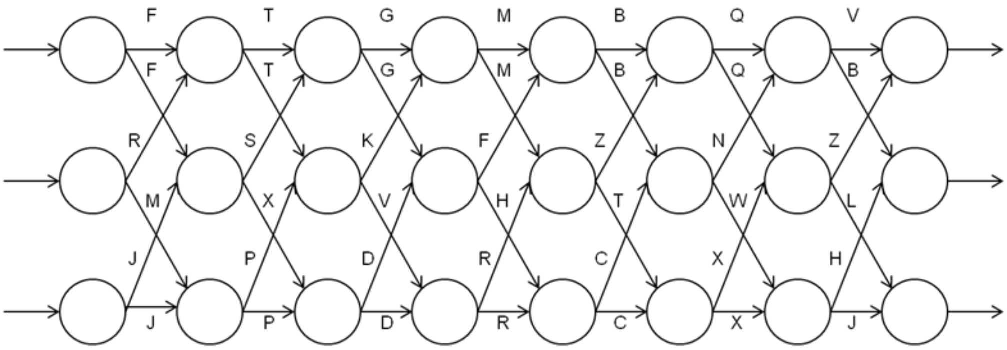


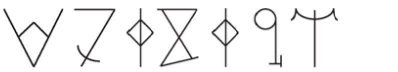
Figure 1. Finite state grammar used in Experiments 1-2. Adapted with permission from Kinder and Lotz (2009).

Appendix A

Table A1. Stimuli used in the memorization phase of experiment 1 and 2.

List1		**List 2**	
Experiment1	**Experiment2**	**Experiment1**	**Experiment2**
FSGFBNB		FSGMBWH	
FTGMBQV		FTKMBQL	
FTVRCNV		FTVRTXJ	
FXDRZQB		FXDHCXH	
JPDRCXJ		JPKMZQV	
JPKMBWJ		JPVRCXZ	
JSGMTXH		JSGFBQB	
JXDHCWH		JXDRCNB	
MPDFBQL		MPDFBNV	
MPVRTXL		MPDRZQZ	
RTGHCXZ		RTGHCWJ	
RTKMZQZ		RTGMTXL	

Table A2. Stimuli used in the test phase of experiment 1 and 2.

List 1		List 2	
Experiment1	**Experiment2**	**Experiment1**	**Experiment2**
Strings containing chunk and letter/shape-position violations			
FXDHFML		FXMSZQL	
FXLKCXJ		FXDRXCV	
FSXLBWJ		JSPGTXJ	
FTVRTKX		FRFRCNB	
RTVNLXJ		RTKZJQB	
MBXMZQV		MPVRTDM	
JKFMZQB		LBKMBWH	
JSGFGDZ		JSGSQXZ	
MPVLPQZ		DCKMBWJ	
BQDRZQV		MPDRTFX	
HLKMTXL		RWQMZQZ	
RTGMTVS		RTVRFKV	
Strings containing chunk violations only			
FXDHZNL		FXKFZQL	

FXVFCXJ		FXDRBXV	
FSKRBWJ		JSKHTXJ	
FTVRTWZ		FPGRCNB	
RTVFZXJ		RTKFCQB	
MTDMZQV		MPVRTQH	
JTDMZQB		MXKMBWH	
JSGFZNZ		JSGRBXZ	
MPVFCQZ		MXKMBWJ	
JTDRZQV		MPDRTWL	
MXKMTXL		RSVMZQZ	
RTGMTQJ		RTVRBXV	

Strings that differ from any training item by 3 letters or more, i.e., they are dissimilar to the items used in the memorization phase

FXDHCXL		FXDRZQL	
FXDHCXJ		FXDRZQV	
FSGMBWJ		JSGMTXJ	
FTVRTXH		FTVRCNB	
RTVRTXJ		RTKMZQB	
MPKMZQV		MPVRTXH	

JPKMZQB		JPKMBWH	
JSGFBQZ		JSGMTXZ	
MPVRZQZ		MPKMBWJ	
MPDRZQV		MPDRTXL	
RTKMTXL		RTVRCNV	
RTGMTXJ		RTGMZQZ	

Strings that differ from the closest training item by one letter, i.e., they are similar to a specific training item

FXDRZQL		FXDHCXL	
FXDRZQV		FXDHCXJ	
JSGMTXJ		FSGMBWJ	
FTVRCNB		FTVRTXH	
RTKMZQB		RTVRTXJ	
MPVRTXH		MPKMZQV	
JPKMBWH		JPKMZQB	
JSGMTXZ		JSGFBQZ	
MPKMBWJ		MPVRZQZ	
MPDRTXL		MPDRZQV	
RTVRCNV		RTKMTXL	

RTGMZQZ  RTGMTXJ 

Appendix B

Computation of global Associative Chunk Strength (ACS) of the stimuli (based on Knowlton & Squire, 1994).

ACS was computed by (a) partitioning each test-phase item (e.g., FSGMBWJ) into its constituent bigrams (FS, SG, GM, MB, BW, WJ) and trigrams (FSG, SGM, GMB, MBW, BWJ), (b) summing their frequency of occurrence across the 12 training items and (c) averaging the sum across the 11 chunks which comprised *each* test-phase string.

Appendix C

To rule out the possibility of unexpected group differences regarding performance on the background, literacy, and literacy-related measures in experiment 1 and 2, we subjected the data to a series of ANOVAs with group (skilled vs. dyslexic readers) and stimulus format (letter vs. shape strings) as between-subject factors. These analyses confirmed that age was not significantly different between groups, $F(1, 98) = 0.04$, $p = .836$, $\eta^2 = .00$, or stimulus formats, $F(1, 98) = 0.02$, $p = .902$, $\eta^2 = .00$, and that these factors did not interact, $F(1, 98) = 0.10$, $p = .753$, $\eta^2 = .00$. Similarly, performance on the Vocabulary test did not differ between groups, $F(1, 97) = 0.48$, $p = .492$, $\eta^2 = .00$, or stimulus formats, $F(1, 97) = 0.11$, $p = .746$, $\eta^2 = .00$, and these factors did not interact, $F(1, 97) = 0.62$, $p = .433$, $\eta^2 = .01$. There was an unexpected significant effect of stimulus format on participants' Matrices performance, $F(1, 97) = 5.38$, $p = .023$, $\eta^2 = .05$, suggesting that participants in the letter variant had higher nonverbal IQ ($M = 108.33$, $SE = 1.71$) relative to participants in the shapes variant ($M= 102.76$, $SE = 1.68$). This difference held across the two reading ability groups, $F(1, 97) = 2.33$, $p = .130$, $\eta^2 = .02$, confirming that skilled and dyslexic readers did not differ in terms of nonverbal IQ, and there was no group by stimulus format interaction, $F(1, 97) = 0.62$, $p = .433$, $\eta^2 = .01$.

Turning to performance on the literacy and literacy-related measures, the analyses on WRAT reading performance showed a significant effect of group, $F(1, 97) = 48.93$, $p < .001$, $\eta^2 = .32$, a nonsignificant effect of stimulus format, $F(1, 97) = 2.11$, $p = .149$, $\eta^2 = .01$, and a significant group by stimulus format interaction, $F(1, 97) = 6.56$, $p = .012$, $\eta^2 = .04$. Breaking down the interaction showed that there were no statistically significant differences in performance between the skilled readers in the letters and shapes variants, $t(60) = 0.97$, $p = .337$, $d = 0.25$; however, dyslexic participants in the shapes variant tended to score higher on the WRAT Reading measure relative to dyslexic participants in the letter variant, a difference that was not significant after bonferonni correction for multiple comparisons, $t(37) = 2.30$, $p = .027$, $d = 0.73$. The analyses of WRAT spelling performance showed the expected significant differences in favor of skilled ($M = 109.82$, $SE = 1.12$) relative to dyslexic readers ($M = 95.64$, $SE = 1.42$), $F(1, 97) = 61.57$, $p < .001$, $\eta^2 = .38$, no effect of stimulus format, $F(1, 97) = 0.53$, $p = .468$, $\eta^2 = .00$, and no group by format interaction, $F(1, 97) = 2.24$, $p = .138$, $\eta^2 = .01$. Dyslexic participants were significantly slower ($M = 18.21$, $SE = 0.55$) relative to skilled readers ($M = 14.17$, $SE = 0.44$) in terms of RAN digits performance, $F(1, 97) = 32.76$, $p < .001$, $\eta^2 = .25$, there was no significant difference between participants in the letters and shapes version, $F(1, 97) = 1.03$, $p = .314$, $\eta^2 = .01$, and no group by stimulus format interaction, $F(1, 97) = 0.19$, $p = .663$, $\eta^2 = .00$. The analyses on RAN objects reaction times replicated this pattern: dyslexic participants were significantly slower ($M = 25.11$, $SE = 0.67$) relative to skilled readers ($M = 21.67$, $SE = 0.53$), $F(1, 97) = 16.13$, $p < .001$, $\eta^2 = .14$, there was no significant difference between participants in the letters and shapes variants, $F(1, 97) = 0.24$, $p = .625$, $\eta^2 = .00$, and no group by stimulus format interaction, $F(1, 97) = 0.88$, $p = .351$, $\eta^2 = .01$.

Consistent with the above, the ANOVA on Digit Span performance revealed a significant main effect of group, $F(1, 95) = 19.59$, $p < .001$, $\eta^2 = .17$, due to skilled readers' advantage on this measure ($M = 10.72$, $SE = 0.34$) relative to dyslexic readers ($M = 8.31$, $SE = 0.42$), no effect of stimulus format, $F(1, 95) = 0.19$, $p = .661$, $\eta^2 = .00$, and no group by stimulus format interaction, $F(1, 95) = 0.36$, $p = .550$, $\eta^2 = .00$.

Skilled readers' advantage held also true for Symbol Search performance: They (*M* = 13.74, *SE* = 0.35) outperformed dyslexic readers (*M* = 12.40, *SE* = 0.44), $F(1, 96) = 5.66$, $p = .019$, $\eta^2 = .05$, and there was no effect of stimulus format, $F(1, 96) = 1.23$, $p = .271$, $\eta^2 = .01$, or interaction with this factor, $F(1, 96) = 0.74$, $p = .393$, $\eta^2 = .01$.

Finally, the analyses on nonword phoneme deletion latencies showed a significant effect of group, $F(1, 97) = 74.18$, $p < .001$, $\eta^2 = .40$, a significant effect of stimulus format, $F(1, 97) = 7.65$, $p = .007$, $\eta^2 = .04$, and a significant group by stimulus format interaction, $F(1, 97) = 7.85$, $p = .006$, $\eta^2 = .04$. The interaction was caused by a significant difference in terms of dyslexic participants' performance in the shapes variant relative to the letter variant ($t(37) = 2.61$, $p = .013$, $d = 0.84$; not significant after bonferonni correction for multiple comparisons), but no difference in performance between the skilled readers in the letters and shapes variants, $t(60) = 0.04$, $p = .967$, $d = 0.01$.

In sum, (a) skilled readers were well-matched across experiments, except for the aforementioned, unpredicted nonverbal IQ advantage of participants in the letter variant relative to participants in the shapes variant, (b) skilled readers outperformed dyslexic participants in all literacy and literacy-related measures, and (c) dyslexic readers were generally well-matched across experiments, although there were trends in the analyses of WRAT reading and nonword phoneme deletion performance suggesting that dyslexic participants in the drag and drop shapes variant may have been somewhat better compensated than dyslexic participants in the drag and drop letters variant.

Appendix D

Memorization data from experiment 1 and 2 were subjected to a series of omnibus ANOVAs comparing performance across stimulus formats (letter vs. shape strings) and groups (skilled vs. dyslexic readers). The analysis of the proportion of strings reproduced correctly within a single attempt revealed a significant effect of stimulus format, $F(1, 98) = 5.24$, $p = .024$, $\eta^2 = .05$ (letter strings: $M = .88$, $SE = 0.01$; shape strings: $M = .83$, $SE = 0.01$), but no effect of group, $F(1, 98) = 2.38$, $p = .126$, $\eta^2 = .05$, or group by stimulus format interaction, $F(1, 98) = 1.78$, $p = .185$, $\eta^2 = .02$. The analyses on the mean number of trials to criterion revealed no effect of stimulus format, $F(1, 98) = 2.97$, $p = .088$, $\eta^2 = .03$, no effect of group, $F(1, 98) = 2.17$, $p = .144$, $\eta^2 = .02$, and no interaction between group and stimulus format, $F(1, 98) = 1.61$, $p = .208$, $\eta^2 = .02$. The effect of stimulus format on mean correct memorization latencies was strong, $F(1, 98) = 71.79$, $p < .001$, $\eta^2 = .42$, showing that participants' mean correct memorization RTs for the shape strings ($M = 10558.94$, $SE = 378.24$) were almost double when compared to participants' RTs for the letter strings ($M = 5956.18$, $SE = 389.93$). The difference between skilled and dyslexic readers' RTs was not significant, neither was the group by stimulus format interaction, both *Fs* < 1.

With regards to mean correct memorization latencies, the model that included the effect of stimulus format was preferred to the model that further included the main effect of group ($B_{01} = 4.297$) and over the model that further included the group by stimulus format interaction ($B_{01} = 3.207$). This was not the case in the analyses of the proportion of strings reproduced correctly within a single attempt. There was no conclusive evidence that the model that included the effect of stimulus format was preferred to the model that further included the main effect of group ($B_{01} = 1.749$) or the model that further included the group by stimulus format interaction (B_{01} = 1.714). Similarly, in the analyses of the mean number of trials to criterion, there was no conclusive evidence that the intercept-only model was preferred to the model that included the main effect of group (B_{01} = 1.833) or the model that included the group by stimulus format interaction (B_{01} = 1.503).