Journal of Experimental Child Psychology 145 (2016) 139–156

Contents lists available at ScienceDirect

Journal of Experimental Child Psychology

journal homepage: www.elsevier.com/locate/jecp

The involvement of long-term serial-order memory in reading development: A longitudinal study

Louisa Bogaerts [a,*], Arnaud Szmalec [b], Marjolijn De Maeyer [a], Mike P.A. Page [c], Wouter Duyck [a]

[a] Ghent University, B-9000 Ghent, Belgium
[b] Université Catholique de Louvain, B-1348 Louvain-La-Neuve, Belgium
[c] University of Hertfordshire, Hatfield AL10 9AB, UK

ARTICLE INFO

Article history:
Received 29 January 2015
Revised 22 December 2015
Available online 5 February 2016

Keywords:
Reading
Reading disability
Memory
Language acquisition
Hebb repetition learning
Sequence learning
Serial order

ABSTRACT

Recent findings suggest that Hebb repetition learning—a paradigmatic example of long-term serial-order learning—is impaired in adults with dyslexia. The current study further investigated the link between serial-order learning and reading using a longitudinal developmental design. With this aim, verbal and visual Hebb repetition learning performance and reading skills were assessed in 96 Dutch-speaking children who we followed from first through second grade of primary school. We observed a positive association between order learning capacities and reading ability as well as weaker Hebb learning performance in early readers with poor reading skills even at the onset of reading instruction. Hebb learning further predicted individual differences in later (nonword) reading skills. Finally, Hebb learning was shown to explain a significant part of the variance in reading performance above and beyond phonological awareness. These findings highlight the role of serial-order memory in reading ability.

* Corresponding author.
E-mail address: bog.louisa@gmail.com (L. Bogaerts).

http://dx.doi.org/10.1016/j.jecp.2015.12.008

Introduction

Whereas most children achieve fluent reading skills with relative ease, for others learning to read involves significant difficulties. Approximately 5 to 10% of the population develops dyslexia, characterized by unexpected and persistent difficulties with reading in the context of normal intelligence, adequate sensory functions, and typical educational opportunities (e.g., Lyon, Shaywitz, & Shaywitz, 2003). What determines this important variability in the acquisition of reading skills? What underlies the difficulties of poor readers? A long tradition of research on literacy acquisition has suggested factors such as phonological skills (e.g., see Melby-Lervag, Lyster, & Hulme, 2012, for a review), letter knowledge (e.g., Bond & Dykstra, 1997; Muter & Diethelm, 2001), and short-term memory capacity (e.g., Gathercole & Baddeley, 1993; Rohl & Pratt, 1995) as important predictors for individual differences in reading ability. In parallel, research on reading disability has focused mostly on phonological problems (impaired phonological representations: e.g., Snowling, 2000; problematic phonological access and retrieval: e.g., Boets et al., 2013; Ramus & Szenkovits, 2008) as well as other impairments in (cognitive) functions such as vision (e.g., Bosse, Tainturier, & Valdois, 2007; Chase & Stein, 2003), attention (e.g., Hari & Renvall, 2001), perceptual anchoring (Ahissar, 2007), and memory (e.g., Hachmann et al., 2014; Martinez Perez, Majerus, Mahot, & Poncelet, 2012; Smith-Spark & Fisk, 2007; Szmalec, Loncke, Page, & Duyck, 2011; see Ramus & Ahissar, 2012, for a nuanced overview).

The current study was specifically concerned with the contribution of *serial-order* memory (i.e., memory for the order in which items are presented within a sequence) to early reading. Building on the assumption that learning to read words may be conceived as the acquisition of ordered sequences of graphemes and their corresponding phonemes (Page & Norris, 2009), and following the study by Szmalec et al. (2011) demonstrating deficient serial-order learning capacities in adults with dyslexia, we hypothesized that serial-order learning may play an important role in normal and pathological reading development.

Linking serial-order memory and language

How does memory for serial order relate to language learning and literacy? From an evolutionary perspective, it has been assumed that short-term memory for verbal information developed primarily to support language learning. Baddeley, Gathercole, and Papagno (1998) argued that verbal working memory represents "the processes and mechanisms by which the sound patterns of the words of the (native) language are learned by the child" (p. 159). Indeed, natural language can be regarded as a well-structured environmental input with an inherently sequential nature. A limited number of phonemes and letters form different words, depending on the order of their arrangement, and these words in turn are sequentially arranged to form sentences. It is becoming increasingly clear that both (a) the ability to *temporarily* represent the order of discrete elements occurring in a sequence (i.e., *short-term order memory*) and (b) the ability to consolidate this sequential information in *long-term memory* (referred to as *serial-order learning* or sequential learning) are implicated in several aspects of human language such as lexical acquisition and reading ability.

Lexical acquisition

Evidence for a tight link between short-term memory for order and lexical development comes, inter alia, from the reports of robust correlations (mostly in the range of .40–.50) between performance on verbal immediate serial recall tasks and both nonword repetition (e.g., Gathercole et al., 1999; Gupta, 2003) and vocabulary scores (in either a first or second language; e.g., Gathercole & Baddeley, 1989; Service, 1992). Furthermore, several recent studies by Majerus and colleagues have highlighted the importance of the serial-order component of short-term memory (STM), as opposed to memory for item identity, in relation to early oral language learning (e.g., Leclercq & Majerus, 2010; Majerus & Boukebza, 2013; Majerus, Poncelet, Greffe, & Van der Linden, 2006) and literacy acquisition (Martinez Perez, Majerus, & Poncelet, 2012). Finally, recent research has demonstrated that the order component of STM seems to be affected in both children and adults with dyslexia

(Hachmann et al., 2014; Martinez Perez, Majerus, & Poncelet, 2013; Martinez Perez, Majerus, Mahot, et al., 2012; but see Staels & Van den Broeck, 2014, for a nonreplication).

Multiple authors have proposed that these order–STM mechanisms contribute to long-term learning of new phoneme (and by extension orthographic) sequences via Hebbian learning (e.g., Burgess & Hitch, 2006; Gupta, 2003; Page & Norris, 2009). Hebb (1961) showed that when a particular ordered sequence of stimuli was repeated several times in an immediate or short-term serial recall task, recall of that sequence (known as the Hebb sequence) improved compared with recall of nonrepeated sequences (known as filler sequences). This phenomenon is known as the *Hebb repetition effect* and reflects incidental (repetition-driven) long-term sequence learning. A number of researchers have argued that long-term serial-order learning, operationalized by the Hebb repetition paradigm, provides an analogue for the processes involved in naturalistic vocabulary learning to the extent that the acquisition of novel word forms also requires the retention of letter or phoneme sequences in a specified serial order (e.g., Cumming, Page, & Norris, 2003; Page & Norris, 2009). Consistent with the assumption that Hebb repetition learning mimics naturalistic word-form learning is the observation of a positive correlation between Hebb repetition learning performance and nonword learning in a sample of typically developing 5- and 6-year-olds (Mosse & Jarrold, 2008). Recent experimental evidence was provided by Szmalec, Duyck, Vandierendonck, Barberá-Mata, and Page (2009), who showed that repeating syllabic sequences in the Hebb repetition learning paradigm (e.g., the grouped sequence of nine CV [consonant–vowel] items "ri-zo-bu_ni-li-na_sa-ba-du") resulted in representations in the mental lexicon that are functionally similar to those of newly learned words (e.g., in this case, the novel "words" rizobu, nilina, and sabadu). More recent work also showed that these newly acquired Hebb sequences engage in lexical competition with existing words, just like novel word forms do (Szmalec, Page, & Duyck, 2012).

Reading (dis)ability

The above theoretical framework clarifies the link between memory for serial order and lexical acquisition, and it can be extended to the domain of (early) word reading. Models of reading such as the SERIOL model (Whitney, 2001; Whitney & Cornelissen, 2005) stress the importance of the (temporal) alignment of the serial orthographic representations (i.e., letters position and identity) and phonological representations in early reading. Imagine an early reader processing the word CAT. The child will typically use a decoding strategy through which he or she converts each individual letter (or grapheme) into its corresponding sound (or phoneme) while integrating a representation of the entire sequence of sounds (/k/–/æ/–/t/) into a single word form. Repeatedly processing this visual sequence of letters will then, through Hebbian learning, gradually develop into an orthographic representation in the mental lexicon, which allows more automatic and proficient processing of the known letter string. This framework, and the observation that many of the experimental tasks (including tasks from outside the linguistic domain) that yield difficulties for people with dyslexia involve sequentiality, has inspired a new account of reading impairment that we labeled the SOLID (serial order learning in dyslexia) hypothesis (Szmalec et al., 2011). This memory-based account of dyslexia offers an alternative view to the prominent etiological stance that dyslexia results from a phonological deficit, that is, problems with the representation and processing of speech sounds (Snowling, 2000; Stanovich, 1988). It proposes that "dyslexia, and its associated cognitive dysfunctions, may be traced back specifically to the learning of serial order" (Szmalec et al., 2011, p. 1271). Szmalec et al. (2011) indeed demonstrated that dyslexic adults showed reduced Hebb repetition learning across both verbal and visuospatial modalities. The demonstration of a deficit in a visuospatial task implies that Hebb learning deficits in dyslexia extend beyond the verbal domain and that a domain-general serial-order component may be the source of impairment. In support of this view, we recently showed that the learning deficit is persistent in the sense that drastically increasing the number of Hebb repetitions, thereby maximizing learning opportunity, does not mitigate the adverse effect of dyslexia on Hebb learning (Bogaerts, Szmalec, Hachmann, Page, & Duyck, 2015). The same study also suggested poorer lexicalization of verbal Hebb sequences in adults with dyslexia, suggesting that problems with serial-order learning may lead to impaired lexical representations, which are in turn assumed to affect reading performance (Perfetti, 2007). The earliest evidence for an association between serial-order learning difficulties and reading problems comes from Gould and Glencross (1990), who reported a reliable verbal Hebb learning impairment in reading-disabled children aged 11 years but no group difference in a visuospatial Hebb task. Recently, Staels and Van den Broeck (2015) failed to find evidence for weaker Hebb learning in children (sixth graders) and adolescents with dyslexia, leading them to question a Hebb learning impairment as (one of) the underlying problem(s) in dyslexia.

In summary, the research described so far suggests a theoretical link between serial-order learning (of which Hebb learning is a paradigmatic example) and language skills. However, whereas the role of serial-order learning in lexical acquisition (or vocabulary development) has been demonstrated rather convincingly, the exact role it plays in reading acquisition and reading (dis)abilities remains less clear. The evidence linking Hebb learning to dyslexia is not unequivocal, and the generalizability of these findings to early readers requires additional investigation. This was the focus of the current study, which is the first to test Hebb repetition learning in children using a longitudinal approach.

The current study

The current study investigated the relationship between long-term serial-order learning (here operationalized as Hebb repetition learning) and (poor) reading ability using a longitudinal design. It is generally accepted that reading difficulty should be considered as a continuum rather than an all-or-none condition (Fletcher, 2009; Shaywitz, Escobar, Shaywitz, Fletcher, & Makuch, 1992). Yet, the empirical evidence supporting the link between serial-order learning—measured as Hebb repetition learning—and literacy hinges, to our knowledge, exclusively on group studies comparing dyslexic participants and matched controls. Here we looked, for the first time, both at poor readers versus matched controls and at the relationship between Hebb learning and reading, considering the full reading continuum. Thus, the study had three major objectives. First (Objective 1a), we aimed to investigate whether the observation of Hebb learning deficits in dyslexic adults extends to poor-reading children. Testing children at the very start of reading instruction deals, at least partly, with the alternative explanation that difficulties with serial-order learning are not producing the dyslexic symptoms but are instead the result of impaired reading abilities. In the same vein (Objective 1b), we aimed to examine the relationship between long-term Hebb repetition learning and reading skills using a large sample of readers along the reading continuum. Second (Objective 2), the use of a longitudinal design provides a unique opportunity to estimate the potential of the Hebb repetition paradigm as a predictive tool for (pathological) reading development. Third (Objective 3), a final research question was whether Hebb repetition learning ability contributes to word and nonword reading skills *independently* of phonological awareness, a well-established and commonly accepted predictor of individual differences in reading ability (Melby-Lervag et al., 2012).

Two types of children were included in the study: children at risk for dyslexia and children without risk. During a first test period (first grade, 6–7 years of age), both auditory–verbal and visuospatial Hebb learning were assessed in addition to word reading abilities. Importantly, at this time children had received only initial reading instruction and had little reading experience. One year later (second grade, 7–8 years of age), we tested the same children with the same Hebb learning tasks, and in addition a nonverbal intelligence measure and word/nonword reading tasks were administered. We further included a spelling task to obtain an estimate of orthographic skills as well as measures of phonological awareness and naming speed to obtain an estimate of the quality and accessibility of phonological, sublexical, and lexical representations. The predictions regarding the major objectives are outlined below:

1a. We predicted a Hebb repetition effect in the children's group notwithstanding the fact that Hebb learning has been found to be somewhat weaker in children compared with adults (e.g., Hitch, McNeil, Page, Cumming, & Norris, 2015; Mosse & Jarrold, 2008). In line with the adult dyslexia data (Bogaerts et al., 2015; Szmalec et al., 2011), we also predicted that poor readers would display weaker Hebb learning compared with good readers in both the verbal and visuospatial stimulus modalities.

1b. Considering reading skill as a continuous variable, we predicted a positive relationship between both word and nonword reading scores and Hebb learning performance.

2. We expected a predictive correlation between the Hebb learning effect and reading performance 1 year later.
3. We anticipated that the Hebb measure would explain a unique portion of the variance in reading above and beyond phonological awareness.

Method

Participants

Primary school children were tested in the first grade (Time Point 1 = T1) and again 1 year later, when they attended the second grade (Time Point 2 = T2). At both time points, testing took place between March and May. A total of 96 children (47 boys; mean age at T1 = 6.7 years, $SD = 0.41$, range = 6.0–7.9) participated at both time points. Of these children, 47 were selected for the study on the basis of their increased risk for reading difficulties. This risk was assessed through parental report of a delay in language development or through family risk, meaning that the child had a family member (within the third degree of consanguinity) who reported reading difficulties. The children were recruited from 15 primary schools in Flanders, Belgium, and were all monolingual Dutch speakers. They had no history of sensorimotor or neurological disorders according to the parents' reports.

Materials and procedure

All children were tested individually in a quiet room at their school. At T1 (first grade), all participants underwent the verbal and spatial Hebb task sessions, with the order of presentation of the two sessions counterbalanced across participants. After completing the Hebb tasks, a measure of word reading was administered. At T2 (second grade), the same verbal and spatial Hebb tasks were administered to the participants. To avoid the influence of material-specific savings, we made sure that participants were presented with different Hebb sequences from those they had learned 1 year earlier. The order of the Hebb tasks was again counterbalanced. The remaining tests were administered in a fixed order; after completing the Hebb tasks, nonverbal intelligence was assessed, followed by word reading, nonword reading, and the four subscales of the Dyslexia Screening Test (Kort et al., 2005). Children received a cartoon sticker as a reward.

Hebb learning tasks

Verbal domain

The verbal Hebb learning task was an adaptation of the procedure used by Mosse and Jarrold (2010). The task was presented on a 15-inch laptop computer and was introduced to the children as an animal race task. On each trial, a sequence of six Dutch animal words was presented auditorily for immediate serial recall. All animal words—duif, hert, hond, stier, leeuw, and paard (pigeon, deer, dog, bull, lion, and horse, respectively)—were monosyllabic and had a high frequency (log freqency/million: $M = 1.5039$, range = 0.8451–2.2253; calculated using WordGen by Duyck, Desmet, Verbeke, and Brysbaert (2004). The names were recorded by a female voice, and all audio files were edited to have a length of 1000 ms. Immediately after the auditory presentation of the six animal names, a visual recall screen appeared showing six simple black and white animal drawings arranged randomly in a "noisy" circle around a central question mark. Participants were instructed to recall the order of the animal words by clicking on the corresponding pictures; the question mark could be used for a missing animal word. Note that this procedure allows children to click the same animal more than once. However, it was not possible to recall an animal that was not in the stimulus list. After the participants had clicked six animal pictures, a black screen was presented and the following trial was initiated after a self-paced press on the spacebar. The task consisted of 16 trials in total: 8 repetitions of the repeating Hebb sequence interspersed with 8 random filler sequences. Three different Hebb lists were used (across participants) to avoid list-specific effects. On the filler trials, which alternated with the Hebb trials, the order of the six animal names was random. The dependent variable was the accuracy of each individual item's recall (i.e., 1 if the animal word was recalled in the correct serial position, 0 otherwise).

Visuospatial domain

The visuospatial Hebb learning task was similar to the one used by Mosse and Jarrold (2008, 2010). Seven images of green lily pads and a frog were presented on the screen. Each trial consisted of an animated frog appearing on one lily pad and jumping in sequence onto the remaining six lily pads with an interstimulus interval of 1000 ms. After jumping onto the final pad, the frog disappeared from the screen; however, the pads remained. Participants responded by clicking the sequence of pads in the correct serial order. Pads changed color (from light green to darker green) when they were clicked. Three different versions of the task were made using different spatial background compositions for the seven lily pads and a different Hebb sequence. All three Hebb lists contained one single path crossing (i.e., the frog crosses the virtual path between two previously visited lily pads). The dependent variable was again the accuracy of each individual item's recall.

Intelligence

The Raven Standard Progressive Matrices test (Raven, Court, & Raven, 1992) was used to obtain a measure of nonverbal intelligence so that intelligence scores would not be confounded with linguistic performance, especially in the poor-reading group.

Reading

Three Minutes Test

The Drie Minuten Toets (DMT; Three Minutes Test; Verhoeven, 1995) is a standardized word reading test. The test consists of three reading cards with increasing difficulty. Participants are instructed to correctly read aloud as many words as possible within the time limit of 1 min per card. There are three different versions of all reading cards. To avoid effects of retesting, we used Version B at T1 (first reading card only) and Version C at T2 (all three cards). The total score is calculated for each card as the total number of words read minus the number of reading errors.

Klepel

The Klepel (Van den Bos, Spelberg, Scheepstra, & de Vries, 1994) is a nonword reading test in which participants are instructed to read aloud correctly as many nonwords as possible within 2 min. Because the test is suitable and normed only for children from the second grade onward, we administered the test only at T2 and used Version B.

Other language tests

Spelling

Spelling was administered with a subtest of the Dyslexia Screening Test (DST; Kort et al., 2005) named "two minutes spelling." Children have 2 min to write down as many words as possible. The words are read out loud by the experimenter (a new word is read only after the previous word was spelled) and have an increasing level of difficulty. The number of correctly spelled words is taken as the total score on the subtest.

Phonological awareness

The phonological awareness subtest of the DST consists of two tasks: phoneme deletion and spoonerisms. In the phoneme deletion task, children are asked to repeat orally presented words omitting one or multiple phonemes (e.g., "spi*n*" [spider] without "*n*"). In the spoonerisms task, the first letters of two orally presented words must be switched (e.g., *H*arry *P*otter becomes *P*arry *H*otter). For each task, the number of correct answers determines the raw score. A total phonological awareness score was calculated by summing up the raw scores of both tasks.

Rapid automatized naming

The rapid automatized naming subtest of the DST was used to assess the speed of lexical access. In this task, participants are asked to rapidly name a set of objects—boom, eend, fiets, stoel, and schaar (tree, duck, bicycle, chair, and scissors, respectively)—and a set of letters that are each presented in a 5×10 matrix. The dependent variable is the time in seconds needed to name all items. For incorrectly named items, 5 s is added.

Results

In the Hebb repetition task, an item was scored as correct if it was recalled in its correct serial position in the sequence. To analyze these binary data, we made use of mixed logit models (see Jaeger, 2008). We first focus on the results of the extreme groups (poor/clinical readers vs. controls). A second section considers the full sample and tests the relation between serial-order learning and reading scores across the entire reading continuum. The models reported are those including the fullest random effects structure justified by the design that still allows the model to converge. As a modeling procedure, when the full random model did not converge, we removed, in a stepwise procedure, nonessential terms such as the random intercept (Barr, Levy, Scheepers, & Tily, 2013), if necessary followed by random (interaction) terms that were not of theoretical interest. Once the model converged, we always tried to re-include the random intercept in the final model.

Extreme groups analysis

For this initial analysis, we selected those children with a clinical (i.e., below the 10th percentile) word and/or nonword reading score at T2 ($n = 23$; 11 boys) and a matched group of 23 children (14 boys) with good reading performance. Of the 23 children with a clinical score, 8 had a clinical score on both reading tests. An additional 15 children had a clinical score on word reading only. Fully 70% of the children (i.e., 16/23) with a clinical score at T2 were initially identified (i.e., at T1) as children at risk for dyslexia based on our screening. Good readers were chosen so that the groups differed maximally in reading performance yet matched (on the group level) as closely as possible on intelligence and age.

Table 1 summarizes the participant characteristics. The poor-reading children and the control children did not differ significantly in their age and nonverbal intelligence. The groups did, however, differ significantly on all language tests except for the rapid automatized naming of objects. Note that we opt for the term "poor-reading children" rather than "children with dyslexia" because, according to the definition of Dyslexie Nederland (Foundation Dyslexia Netherlands), not only does the level of reading and/or writing need to be significantly lower than what can be expected based on the educational level and age of the individual, but also the resistance to instruction (i.e., defective response to therapeutic remediation) needs to be confirmed before one can legitimately speak of dyslexia (Stichting Dyslexie Nederland, 2008).

Fig. 1 shows the learning curves for the verbal and visual tasks as a function of group at T1 and T2. In addition, Table 2 displays the mean percentages of correctly recalled Hebb and filler items.

We used the lme4 package in R (CRAN project; R Foundation for Statistical Computing, 2009) to run a mixed logit model with accuracy as the dependent variable. The fixed effect variables included in the model were type (filler or Hebb), presentation (list position in the task block: 1–8), domain (visual or verbal), and group (controls or poor readers) as well as their two-way interactions. We included the higher order interactions of interest, Type $\times$ Presentation $\times$ Domain and Type $\times$ Presentation $\times$ Group, as well as Type $\times$ Presentation $\times$ Group $\times$ Lag (Session 1 or 2) and Type $\times$ Presentation $\times$ Group $\times$ Domain. We further included IQ as a control variable. The model included the fullest random effects structure for participants and for items that still allowed the model to converge (Barr et al., 2013). Concretely, the random effects structure of the model included a random intercept for participant and a random by-participant slope for type, presentation, domain, Type $\times$ Presentation, and Type $\times$ Presentation $\times$ Domain. The random effect structure for item, defined as the serial position of an item within a sequence, included a random intercept. All continuous predictors were

Table 1
Participant characteristics.

	Extreme groups			Full sample
	Poor readers (n = 23)	Good readers (n = 23)	Group difference	All children (n = 94)
Control variables				
Age (T2, months)	95.1 (4.1)	95.6 (5.6)	p = .77	94.8 (4.6)
Raven PM (%)	54.9 (18.4)	64.4 (23.0)	p = .10	61.6 (22.4)
Reading tests				
T1 TMT (words/1 min, Card 1)	18.0 (7.0)	44.2 (18.2)	p < .001	32.3 (16.2)
T2 TMT (words/1 min, Card 1-3)	28.5 (13.4)	68.0 (10.7)	p < .001	49.3 (19.2)
T2 Klepel (nonwords/2 min)	18.2 (6.4)	44.1 (12.8)	p < .001	9.2 (3.1)
Other language tests				
Spelling (words/2 min)	7.0 (3.0)	12.3 (2.2)	p < .001	10.0 (3.7)
Phoneme deletion (max = 12)	6.6 (2.8)	8.4 (2.4)	p < .05	8.1 (2.5)
Spoonerism (max = 11)	0.8 (1.9)	3.7 (3.5)	p < .01	2.0 (3.0)
RAN objects (s)	61.2 (16.0)	55.3 (12.2)	p = .16	58.3 (14.4)
RAN letters (s)	54.6 (15.0)	36.7 (8.5)	p < .001	46.3 (15.0)

Note. Values are means per group with standard deviations shown in parentheses for T1 and T2 variables. Group differences were tested with a one-way ANOVA with $df(1, 44)$. Raven PM, Raven's Progressive Matrices; TMT, Three Minute Test.

standardized. Factors were sum coded (a.k.a. contrast or analysis of variance [ANOVA]-style coding): type (filler = −1 or Hebb = 1), and domain (visual = −1 or verbal = 1), group (controls = −1 or poor readers = 1).[1] The results of this mixed logit model are summarized in Table 3. We found a significant main effect of type, with the positive coefficient for this effect reflecting higher performance for Hebb sequences than for fillers. A significant interaction of Type × Presentation confirms the presence of a Hebb repetition effect in the developmental sample. A simple slopes analysis revealed that this interaction was driven by a positive correlation between accuracy and presentation for Hebb trials (β = .12, z = 5.51, p < .001) and a negative correlation for filler trials (β = −.10, z = −4.54, p < .001). Marginally significant effects for the interaction terms Type × Domain and Type × Presentation × Domain, both with a negative coefficient, indicate more learning for Hebb lists relative to fillers in the visuospatial modality compared with the verbal modality. More important, we predicted that Hebb repetition learning would be weaker in the poor reading group. Because the two-way interaction capturing the learning over multiple presentations has a coefficient with a positive sign, a group difference in the disadvantage of the poor readers would surface as a three-way interaction, Type × Presentation × Group, with a negative coefficient. We observed a marginally significant interaction of Type × Presentation × Group (as well as Type × Group). This effect did not interact with domain or lag. A simple slopes analysis revealed that the Type × Presentation interaction was significant for the control group (β = .14, z = 6.74, p < .001) and for the poor reading group (β = .07, z = 3.26, p < .01), suggesting that Hebb learning is present in both groups but to a lesser extent for the poor readers, $\chi^2(2)$ = 56.04, p < .001 (see Fig. 1).

When running the maximal mixed logit model on T1 data only, a significant three way-interaction, Type × Presentation × Group, was observed, β = −.07, $\chi^2(1)$ = 3.90, p < .05. This shows that worse Hebb learning for the anticipated poor readers was already present at T1.

Finally, we considered whether the group difference in Hebb learning might be associated with differences in baseline serial-recall performance. Filler performance (i.e., performance on nonrepeating lists) did not differ significantly between the groups, $\chi^2(1)$ = 0.68, p = .41, and is controlled for in the mixed logit model by the fact that Hebb learning is evaluated as the improvement over presentations on Hebb trials relative to improvement over matched filler trials and by the inclusion of a random by-participant intercept. This makes it an unlikely cause, on its own, of the group differences in

[1] Note, however, that the quantity really modeled in a mixed logit model is the log odds. The coefficient β, therefore, implies that a one-unit change in type (i.e., from filler [−1] to average performance [0] or from average performance to Hebb [1]) results in difference of β in the logit probability of getting a correct response.

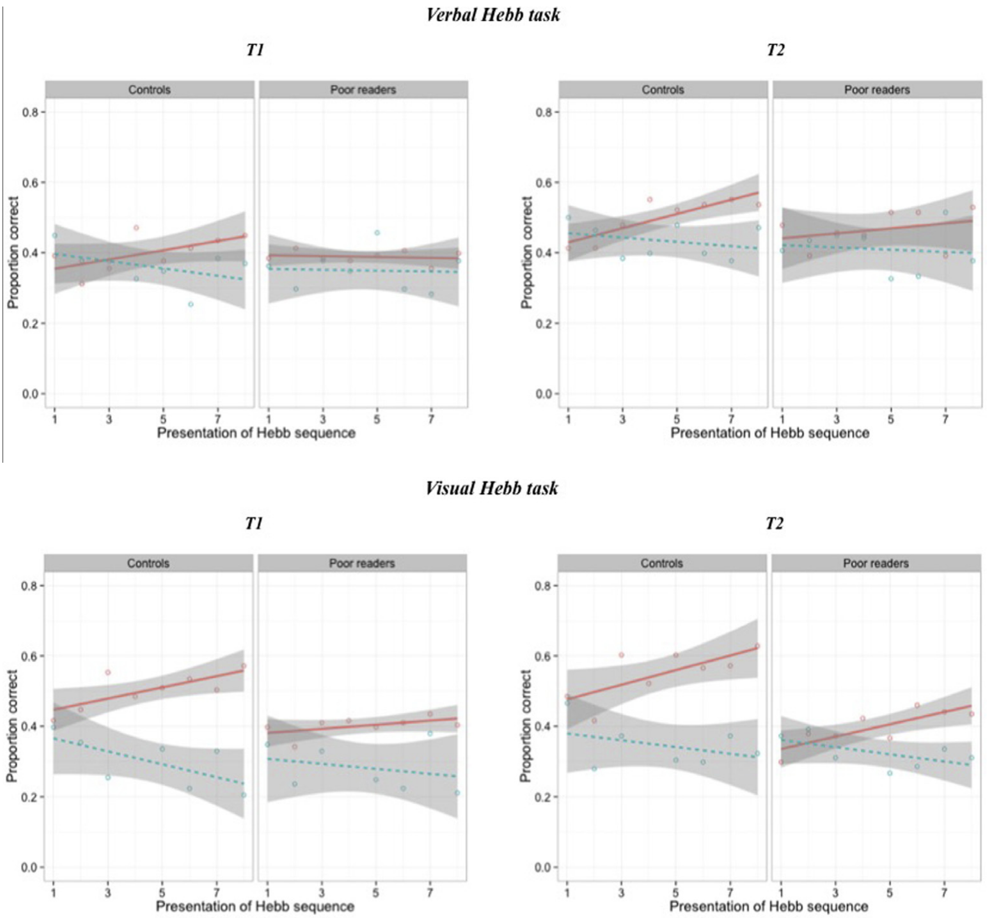

Fig. 1. Plots of the average proportion of correctly recalled items for Hebb (red) and filler (blue) in function of the presentation position of the Hebb sequence. Regression lines have been added to show the change in performance. The gray shading denotes the 95% confidence interval. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

Hebb learning. Indeed, as a check, including average filler performance as an additional control variable in the model did not change the result regarding the Type × Presentation × Group interaction.

In summary, by showing the presence of a clear Hebb repetition effect in the children's group and weaker Hebb learning for poor readers compared with good readers (significantly at T1 and marginally across sessions), the analysis presented above largely confirms the first of our experimental predictions.

Full sample analysis

In total, 96 children (47 boys) participated in the study. At T1, 26 children were tested with preliminary versions of the Hebb learning task that contained fewer items (i.e., a shorter Hebb sequence) than the final versions and that, therefore, could not be used in the analysis. More precisely, 21 children completed versions of the verbal and visual Hebb learning task with fewer items at T1 than in the final versions at T2. In addition, 5 children completed just the visual Hebb learning task with fewer

Table 2
Performance (%) on the two types of trials averaged across repetitions.

		Extreme groups		Full sample
		Poor readers ($n = 23$)	Good readers ($n = 23$)	All children ($n = 94$)
Verbal Hebb task				
T1	Filler	34.96 (12.36)	36.05 (17.33)	35.03 (13.85)
	Hebb	38.86 (16.92)	40.04 (15.87)	42.08 (16.30)
T2	Filler	41.03 (17.47)	43.39 (14.73)	41.45 (16.19)
	Hebb	46.56 (19.86)	50.00 (18.74)	45.05 (18.14)
Visuospatial Hebb task				
T1	Filler	28.22 (10.91)	30.05 (8.52)	29.66 (9.56)
	Hebb	40.14 (20.19)	50.05 (19.55)	45.63 (18.41)
T2	Filler	32.48 (14.52)	34.52 (11.90)	36.50 (11.95)
	Hebb	39.67 (14.32)	54.89 (13.37)	50.52 (17.14)

Note. Standard deviations are in parentheses.

Table 3
Summary of the fixed effects in the mixed logit model.

Predictor	β coefficient	χ^2	p
Intercept	−.329	1.69	.19
Raven PM	.185	7.67	<.01**
Type	.141	26.90	<.001***
Presentation	.012	0.23	.63
Domain	.013	0.07	.80
Group	−.102	2.43	.12
Type × Presentation	.116	26.11	<.001***
Type × Domain	−.029	3.22	.07°
Presentation × Domain	.012	0.28	.60
Presentation × Group	−.018	0.55	.46
Type × Group	−.048	3.00	.08°
Type × Presentation × Domain	−.037	3.36	.07°
Type × Presentation × Group	−.042	3.29	.07°
Type × Pres × Group × Lag		5.29	.26
Type × Pres × Group × Domain		1.32	.52

Note. The quantity modeled in a mixed logit model is the log odds. The coefficient β, such as for type (.141), implies that a one-unit change in type (i.e., from filler [−1] to average performance [0] or from average performance to Hebb [1]) results in difference of .141 in the logit probability of getting a correct response. Raven PM, Raven's Progressive Matrices; Pres, Presentation.

items. This means that at T1 we had complete data for 70 participants and a number of missing values for 26 participants. At T2, the data for all 96 participants were complete.

A model similar to the one described above was run, once with T2 word reading performance (DMT score) as a continuous predictor replacing the factor group and once with nonword reading performance (Klepel score). Thus, the structure of the model included the fixed predictors type, presentation, domain, and reading score; their two-way interactions; the interaction terms Type × Presentation × Domain, Type × Presentation × Reading Score, Type × Presentation × Reading Score × Lag (Session 1 or 2), and Type × Presentation × Reading Score × Domain; and the control variables IQ and age. The random effects structure of the model includes a random intercept for participant and a random by-participant slope for type, presentation, domain, Type × Presentation, Type × Domain, and Type × Presentation × Domain as well as a random intercept for item. All continuous predictors were standardized, and the variables type, condition, and domain were again sum coded (see above).

The results of the mixed logit model with word reading score and nonword reading score as predictors are summarized in Sections A and B, respectively, of Table 4. Crucially, we observed a significant Type × Presentation × DMT interaction as well as a significant Type × Presentation × Klepel

Table 4
Summary of the fixed effects in the mixed logit model.

Predictor	β coefficient	χ^2	p
A. Word reading			
Intercept	−.306	1.52	.22
Raven PM	.172	14.91	<.001***
Type	.151	56.07	<.001***
Presentation	−.005	0.06	.80
Domain	−.014	0.14	.71
DMT	.081	3.03	.08°
Type × Presentation	.097	28.44	<.001***
Type × Domain	−.025	1.53	.22
Presentation × Domain	.001	0.002	.97
Presentation × DMT	.028	2.00	.16
Type × DMT	.051	6.02	.01*
Type × Presentation × Domain	−.039	5.29	.02*
Type × Presentation × DMT	.041	4.86	.03*
Type × Pres × DMT × Lag		0.30	.86
Type × Pres × DMT × Domain		0.34	.85
B. Nonword reading			
Intercept	−.306	1.52	.22
Raven PM	.178	15.36	<.001***
Type	.151	54.37	<.001***
Presentation	−.005	0.92	.34
Domain	−.014	0.20	.65
Klepel	.121	2.67	.10°
Type × Presentation	.096	30.27	<.001***
Type × Domain	−.025	0.96	.33
Presentation × Domain	.001	0.02	.88
Presentation × Klepel	.008	0.01	.94
Type × Klepel	.073	12.55	<.001***
Type × Presentation × Domain	−.039	5.24	.02*
Type × Presentation × Klepel	.038	4.44	.04*
Type × Pres × Klepel × Lag		0.37	.83
Type × Pres × Klepel × Domain		0.05	.97

Note. Raven PM, Raven's Progressive Matrices; DMT, Three Minutes Test; Pres, Presentation.

interaction, both with positive coefficients, indicating stronger Hebb repetition learning in children with higher word and nonword reading scores (see Fig. 2). This confirms the subsidiary part of our first experimental prediction.

Predicting reading performance

Longitudinal regression

In the following analysis, we tested the predictive value of T1 filler performance and T1 Hebb learning for later (i.e., T2) word and nonword reading scores, respectively. The degree of Hebb learning (i.e., the size of the Hebb repetition effect) for a given participant is measured by the individual's coefficient for the Type × Presentation interaction, as extracted from a mixed logit model with accuracy as the dependent variable, run on the T1 data only (see Karuza, Farmer, Fine, Smith, & Jaeger, 2014, and Bogaerts, Siegelman, & Frost, in press, for a similar approach). The model included the fixed predictors type, presentation, and their interaction. Its random effect structure included a random intercept for participant, a random by-participant slope for all fixed terms, and a random intercept for item.

Subsequently, two linear regression models were run: one with word reading as the dependent (i.e., to-be-predicted) variable and one with nonword reading as the dependent variable. Average filler performance and Hebb learning were included as predictors as well as the control variable IQ. All predictors were standardized.

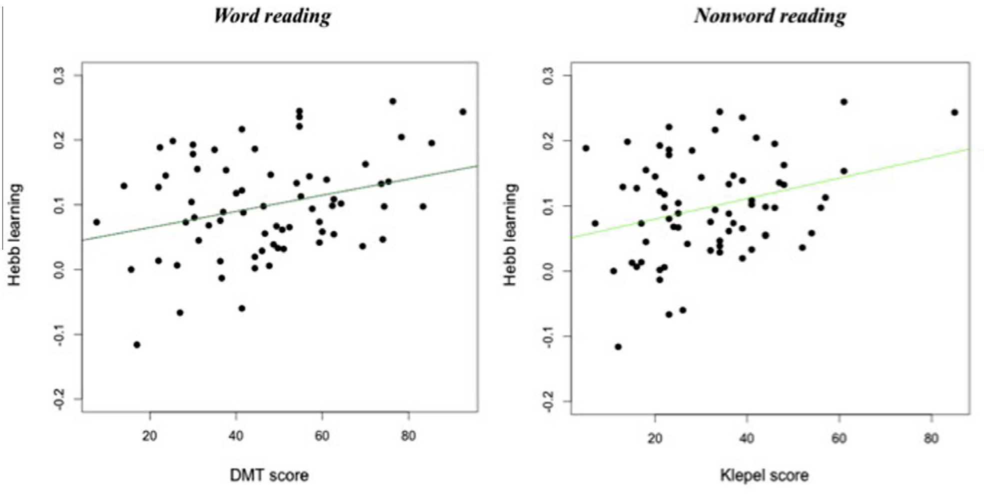

Fig. 2. Scatterplots clarifying the relationship between reading scores and Hebb learning measured by the coefficient for the Type × Presentation interaction (see below). Green lines represent the regression lines. Pearson's r values are .30 for word reading (p = .01) and .29 for nonword reading (p = .01). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 5
Summary of the longitudinal linear regression results.

Predictor	β	t	p
A. Word reading			
Intercept	47.05	21.51	<.001***
Raven PM	3.37	1.40	.17
T1 filler	−0.37	−0.15	.88
T1 Hebb learning	3.53	1.41	.16
B. Nonword reading			
Intercept	31.94	18.66	<.001***
Raven PM	1.13	0.60	.54
T1 filler	−0.07	−0.04	.97
T1 Hebb learning	3.99	2.04	.046*

Note. Raven PM, Raven's Progressive Matrices.

Table 5 shows that neither T1 filler performance nor T1 Hebb learning accounts for significant variance in word reading. However, Hebb learning does reliably predict T2 nonword reading. A model comparison confirms the significant unique contribution of Hebb learning, $F(1) = 4.14$, $p < .05$, $\Delta R^2 = .057$, above and beyond all other predictors.

In a second longitudinal regression, we included reading at T1 (DMT score) as an additional predictor. This linear model predicts variance due to the *growth* in reading over time rather than variance in T2 reading scores per se. Unsurprisingly, word reading at T1 significantly predicts T2 word reading ($\beta = 14.36$, $t = 8.98$, $p < .001$) and nonword reading ($\beta = 10.16$, $t = 7.46$, $p < .001$). Hebb learning performance, which did not significantly predict later word reading, is also not a significant predictor of growth in reading ($\beta = 1.70$, $t = 0.99$, $p = .33$). More interesting, Hebb learning still does reliably predict T2 nonword reading in this more conservative model ($\beta = 2.87$, $t = 1.97$, $p = .05$) and has a unique contribution ($\Delta R^2 = .029$).

In summary, Hebb repetition learning qualifies as a reliable predictor for later nonword reading performance but not word reading performance; our second experimental prediction, therefore, could be partially confirmed.

Explaining variance in reading skills

Finally, we tested our third experimental prediction by evaluating the contributions of filler performance, Hebb learning (across T1 and T2), and phonological awareness in explaining the variance in reading skills. The coefficient for the Type $\times$ Presentation interaction, extracted from the mixed logit model, was again taken as a measure of the size of Hebb learning. A linear regression model with reading score as the dependent variable and average filler performance, Hebb learning, and phonological awareness (all standardized) as predictors was run. As can be seen in Table 6, filler performance did not account for a significant proportion of any of the reading variables variance at T2. However, both Hebb learning and phonological awareness explained a significant amount of variance in word reading and nonword reading. Model comparisons confirm the significant unique contributions of phonological awareness and Hebb learning in explaining the variance in word reading (ΔR^2 phon = .202, ΔR^2 Hebb = .042) and in nonword reading (ΔR^2 phon = .155, ΔR^2 Hebb = .048) above and beyond all other predictors.

General discussion

The question of how memory supports language development has been a topic of wide scientific interest during the past decades (see Baddeley et al., 1998). An increasing number of studies suggest that both short- and long-term memory processes underlie various aspects of language development such as vocabulary acquisition (e.g., Gathercole et al., 1999; Leclercq & Majerus, 2010; Page & Norris, 2009) and reading (e.g., Bogaerts et al., 2015; Martinez Perez, Majerus, & Poncelet, 2012; Szmalec et al., 2011). In previous work, Szmalec and colleagues (2009, 2012) clarified the role of long-term serial-order learning in novel word-form acquisition. In contrast, the role of this type of learning in reading remains less well understood. The goal of the current study was to clarify the involvement of serial-order memory in the development of reading skills.

First, we investigated whether the association between serial-order learning problems and dyslexia that has been demonstrated in adults (Bogaerts et al., 2015; Szmalec et al., 2011) may be generalized to early reading development. This is crucial because dyslexia is, of course, primarily a developmental disorder. The results of our extreme groups analysis suggest weaker Hebb repetition learning in poor readers even at the beginning of reading instruction. As such, the results provide evidence of an association between long-term serial-order learning and reading difficulties. Importantly, weaker Hebb repetition learning for poor readers could not be attributed to worse baseline (short-term) memory capacity. Poor-reading children did not differ significantly from controls in baseline serial recall (or filler) performance, and filler performance was controlled for in the statistical analysis. We further explored the relationship between short-term serial recall (i.e., filler performance) and long-term serial-order learning (i.e., Hebb repetition performance), on the one hand, and reading skill as a continuous variable, on the other. As predicted, we observed a significant positive relationship between the degree of Hebb learning and reading performance measured at T2. In addition, note that the association between reading and Hebb learning did not interact with domain, suggesting that the serial-order deficiency is independent of the modality of memory content and, thus, probably reflects a core deficit in serial-order learning.

Second, concerning the predictive value of the Hebb task, we observed that the magnitude of Hebb learning measured at T1 predicts individual differences in nonword reading abilities 1 year later (with a similar, although nonsignificant, result for word reading), thereby hinting at a possible underlying role of serial-order learning in reading acquisition.

Third, as we administered both measures of phonological awareness and Hebb learning, we were able to show that both variables explained a significant and unique part of the variance in T2 reading performance. This suggests that, in addition to the well-established phoneme awareness deficit, other factors such as impairments in serial-order memory contribute to reading difficulties.

Our joint findings of (a) weaker Hebb learning in poor-reading children, (b) the positive association between Hebb learning and reading performance, and (c) a predictive correlation between the magnitude of Hebb learning and future (nonword) reading abilities support the view that difficulties with the long-term learning of serial-order information may, at least to some extent, underlie reading disability (cf. the SOLID hypothesis; Szmalec et al., 2011). From the SOLID hypothesis perspective, serial-order learning is crucial for registering sequence regularities in the phonological and orthographic

Table 6
Summary of the linear regression results.

Predictor	β	t	p
A. Word reading			
Intercept	47.05	24.64	<.001***
Raven PM	0.97	0.46	.65
Filler	−3.14	−1.44	.15
Phon awareness	9.32	4.38	<.001***
Hebb learning	4.30	2.00	.049*
B. Nonword reading			
Intercept	31.94	20.35	<.001***
Raven PM	−0.23	−0.13	.90
Filler	−1.69	−0.94	.35
Phon awareness	6.41	3.67	<.001***
Hebb learning	3.59	2.03	.046*

Note. Raven PM, Raven's Progressive Matrices; Phon, Phonological.

input. Especially for early readers, who assemble phonology for reading, the correct alignment of letters in written words (i.e., serial orthographic representations) and their conversion to spoken forms (i.e., serial phonological representations) is seen as a key aspect of the reading mechanism (Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001; see also Whitney, 2001, and Whitney & Cornelissen, 2005). Proficient reading is further dependent on the development of long-term, stable phonological and orthographic lexical representations. These representations presumably develop through repeated exposure to phoneme or letter sequences, a process that is an instantiation of Hebb learning. If, due to poor long-term serial-order learning, the order of the individual sublexical items in a sequence is not optimally consolidated as a single lexical entry in lexical memory, the quality of the phonological and orthographical word-form representation will be poor (see Bogaerts et al., 2015). Impoverished representations complicate lexical access during reading in the sense of disrupting the (usually highly automatized) procedures for mapping grapheme and phoneme sequences in word identification (e.g., Whitney & Cornelissen, 2005). This way, poor serial-order learning skills may affect novel word-form acquisition (see also Di Betta & Romani, 2006) and reading performance (e.g., Kuperman & Van Dyke, 2011; Perfetti, 2007).

Considering the interrelationship among serial-order learning, novel word-form acquisition, and reading, we acknowledge the possibility that not only the quality of orthographic lexical representations but also vocabulary size could (partially) mediate the link between Hebb repetition learning and reading skill. In this context, it is noteworthy that poor serial-order learning abilities have recently also been observed in children with specific language impairment (SLI), which is diagnosed when oral language lags behind normal expectations (Hsu & Bishop, 2014). SLI and reading disability are closely related language disorders (see Bishop & Snowling, 2004, for a discussion). On the one hand, the oral language deficits that are typically observed in SLI have also been reported in children with dyslexia (e.g., McArthur et al., 2000; Starck & Tallal, 1988). On the other hand, high rates of literacy problems that are characteristic of dyslexia have also been demonstrated in children with SLI (e.g., Conti-Ramsden, Botting, Simkin, & Knox, 2001; Haynes & Naidoo, 1991; Tallal, Allard, Miller, & Curtiss, 1997). In this sense, our Hebb learning account of language development may be potentially useful to investigate the still poorly understood sources of comorbidity between language disorders.

Interestingly, the positive association between reading and Hebb learning appears to be domain general in nature. These results are consistent with the results of Mosse and Jarrold (2008), who reported an association between Hebb learning across modalities and nonword learning, and with the finding of a general Hebb learning impairment in dyslexic adults (Szmalec et al., 2011). Taken together, this suggests that both vocabulary acquisition and reading do not depend on a uniquely verbal (e.g., phonological, orthographical) sequential learning mechanism but rather seem to rely on the core ability to represent serial-order information (see Depoorter & Vandierendonck, 2009, and Parmentier, 2011, for discussions on the domain specificity of order representation).

Caveats, limitations, and future research

We should emphasize that although our results (especially the weaker Hebb learning performance in children who just began reading instruction and who turned out to experience reading difficulties 1 year later) are consistent with the SOLID account, they do not preclude other existing etiological hypotheses of reading disability. Our findings should not be taken to demonstrate that deficient serial-order learning ability is the single core deficit underlying reading difficulty. Rather, we suggest that serial-order learning provides a novel perspective for understanding both normal and pathological language development, one that merits further investigation. More precisely, problems with serial-order information can explain why people with dyslexia also show impairments outside the linguistic domain such as impaired procedural learning (e.g., Lum, Ullman, & Conti-Ramsden, 2013; Pavlidou, Kelly, & Williams, 2010; see also Howard, Howard, Japikse, & Eden, 2006), motor sequencing (e.g., De Kleine & Verwey, 2009), and working memory functions (e.g., Smith-Spark & Fisk, 2007; see Szmalec et al., 2011, for a discussion). Our demonstration that both phonological awareness and Hebb learning explain unique variance in reading ability is compatible with a multi-deficit view of reading disability (Menghini et al., 2010; Pennington, 2006); it suggests that a serial-order learning deficit can be seen as one of the sources of reading difficulty next to phonological awareness and possibly other factors. Future research requires further examination of the relative importance of the different factors (in both normal and abnormal reading development) and of the extent to which they are interrelated.

Turning to the limitations of the current study, it should be noted that we tested children within a relatively small age range and that, despite the relatively large sample, our group of clinical readers (at T2) was limited to 23 participants. A second limitation arises from our choice of tasks. We opted to administer nonverbal IQ as a control task and a maximal amount of language tests. Given constraints of testing length, this prevented us from also assessing attentional functioning. Given some reported comorbidity between reading disability and attention problems (Boada, Willcutt, & Pennington, 2012, see Staels & Van den Broeck, 2015, for a discussion of attentional factors in Hebb learning), future research should explicitly examine to what extent Hebb repetition performance is modulated by attentional factors. Note, however, that the Hebb learning impairment observed in our poor-reading group is unlikely to have risen from attentional functioning problems. Attentional problems in the poor-reading group would have resulted in a larger drop in filler performance for poor readers compared with the controls. This, however, is not what we observed. A final concern is the unknown test–retest reliability of performance in the Hebb repetition paradigm. Our current investigation indeed provides preliminary evidence for differences between individuals in their (repetition-driven) long-term sequence learning abilities and relates these to reading skills. However, whether sequence learning ability—as measured through the Hebb task—is a stable individual capacity, and whether it correlates with other measures of (incidental) sequence learning, is still unknown. This limitation suggests an important avenue for future research.

Conclusion

The current study aimed to investigate the link between serial-order learning and the development of reading skills in young children. Our results suggest that children who had just begun reading instruction and turned out to experience reading difficulties 1 year later demonstrated weaker Hebb learning performance when compared with normal-reading controls. In the same vein, we observed a positive association between serial-order learning capacities and both reading and nonword reading skills across the full reading continuum. Moreover, Hebb learning was reliably shown to predict later nonword decoding abilities, providing the very first evidence for a possible causal role of serial-order learning in reading acquisition. These results highlight the contribution of serial-order learning to reading and suggest that Hebb repetition learning performance explains a significant and unique amount of variance in reading performance. Overall, these findings contribute to a growing body of evidence for the involvement of serial-order memory in normal and pathological language development.

Acknowledgments

This research was made possible by the Research Foundation–Flanders, of which Louisa Bogaerts is a research fellow, and by the Flemish Agency for Disabled Persons. We are grateful to Alex Fine and Michael Stevens for helping us with the data analysis and to Ram Frost and Noam Siegelman for their helpful comments.

References

Ahissar, M. (2007). Dyslexia and the anchoring-deficit hypothesis. *Trends in Cognitive Sciences, 11*, 458–465.
Baddeley, A., Gathercole, S., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review, 105*, 158–173.
Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*, 255–278.
Bishop, D. V. M., & Snowling, M. J. (2004). Developmental dyslexia and specific language impairment: Same or different? *Psychological Bulletin, 130*, 858–886.
Boada, R., Willcutt, E. G., & Pennington, B. F. (2012). Understanding the comorbidity between dyslexia and attention-deficit/hyperactivity disorder. *Topics in Language Disorders, 32*, 264–284.
Boets, B., Op de Beeck, H. P., Vandermosten, M., Scott, S. K., Gillebert, C. R., Mantini, D., ... Ghesquiere, P. (2013). Intact but less accessible phonetic representations in adults with dyslexia. *Science, 342*, 1251–1254.
Bogaerts, L., Siegelman, N., & Frost, R. (in press). Splitting the variance of statistical learning performance: A parametric investigation of exposure duration and transitional probabilities. *Psychonomic Bulletin and Review*. http://dx.doi.org/10.3758/s13423-015-0996-z.
Bogaerts, L., Szmalec, A., Hachmann, M. W., Page, M. P. A., & Duyck, W. (2015). Linking memory and language: Evidence for a serial-order learning impairment in dyslexia. *Research in Developmental Disabilities, 43–44*, 106–122.
Bond, G. L., & Dykstra, R. (1997). The cooperative research program in first-grade reading instruction. *Reading Research Quarterly, 32*, 348–427.
Bosse, M.-L., Tainturier, M. J., & Valdois, S. (2007). Developmental dyslexia: The visual attention span deficit hypothesis. *Cognition, 104*, 198–230.
Burgess, N., & Hitch, G. J. (2006). A revised model of short-term memory and long-term learning of verbal sequences. *Journal of Memory and Language, 55*, 627–652.
Chase, C., & Stein, J. (2003). Visual magnocellular deficits in dyslexia. *Brain, 126*, E2.
Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review, 108*, 204–256.
Conti-Ramsden, G. M., Botting, N. F., Simkin, Z., & Knox, E. (2001). Follow-up of children attending infant language units: Outcomes at 11 years of age. *International Journal of Language Communication Disorders, 36*, 207–219.
Cumming, N., Page, M., & Norris, D. (2003). Testing a positional model of the Hebb effect. *Memory, 11*, 43–63.
De Kleine, E., & Verwey, W. B. (2009). Motor learning and chunking in dyslexia. *Journal of Motor Behavior, 41*, 331–337.
Depoorter, A., & Vandierendonck, A. (2009). Evidence for modality-independent order coding in working memory. *Quarterly Journal of Experimental Psychology, 62*, 531–549.
Di Betta, A.-M., & Romani, C. (2006). Lexical learning and dysgraphia in a group of adults with developmental dyslexia. *Cognitive Neuropsychology, 23*, 376–400.
Duyck, W., Desmet, T., Verbeke, L. P. C., & Brysbaert, M. (2004). WordGen: A tool for word selection and nonword generation in Dutch, English, German, and French. *Behavior Research Methods, Instruments, & Computers, 36*, 488–499.
Fletcher, J. M. (2009). Dyslexia: The evolution of a scientific concept. *Journal of the International Neuropsychological Society, 15*, 501–508.
Gathercole, S. E., & Baddeley, A. (1989). Evaluation of the role of phonological STM in the development of vocabulary in children: A longitudinal study. *Journal of Memory and Language, 28*, 200–213.
Gathercole, S., & Baddeley, A. (1993). Phonological working memory: A critical building-block for reading development and vocabulary acquisition. *European Journal of Psychology of Education, 8*, 259–272.
Gathercole, S. E., Service, E., Hitch, G. J., Adams, A. M., & Martin, A. J. (1999). Phonological short-term memory and vocabulary development: Further evidence on the nature of the relationship. *Applied Cognitive Psychology, 13*, 65–77.
Gould, J. H., & Glencross, D. J. (1990). Do children with a specific reading disability have a general serial-ordering deficit? *Neuropsychologia, 28*, 271–278.
Gupta, P. (2003). Examining the relationship between word learning, nonword repetition, and immediate serial recall in adults. *Quarterly Journal of Experimental Psychology A: Human Experimental Psychology, 56*, 1213–1236.
Hachmann, W. M., Bogaerts, L., Szmalec, A., Woumans, E., Duyck, W., & Job, R. (2014). Short-term memory for order but not for item information is impaired in developmental dyslexia. *Annals of Dyslexia, 64*, 121–136.
Hari, R., & Renvall, H. (2001). Impaired processing of rapid stimulus sequences in dyslexia. *Trends in Cognitive Sciences, 5*, 525–532.
Haynes, C., & Naidoo, S. (1991). *Children with specific speech and language impairment*. London: Mac Keith Press.
Hebb, D. O. (1961). Distinctive features of learning in the higher animal. In J. F. Delafresnaye (Ed.), *Brain mechanisms and learning* (pp. 37–46). Oxford, UK: Blackwell.
Hitch, G. J., McNeil, A. M., Page, M. P. A., Cumming, N., & Norris, D. (2015). Interactions between short-term and long-term memory in the verbal domain: A series of experiments investigating the Hebb effect in 5-, 7-, and 10-year-old children. Unpublished manuscript.
Howard, J. H., Howard, D. V., Japikse, K. C., & Eden, G. F. (2006). Dyslexics are impaired on implicit higher-order sequence learning, but not on implicit spatial context learning. *Neuropsychologia, 44*, 1131–1144.

Hsu, H. J., & Bishop, D. V. M. (2014). Sequence-specific procedural learning deficits in children with specific language impairment. *Developmental Science, 17*, 352–365.
Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language, 59*, 434–446.
Karuza, E. A., Farmer, T. A., Fine, A. B., Smith, F. X., & Jaeger, T. F. (2014). On-line measures of prediction in a self-paced statistical learning task. In P. Bello, M. Guarini, M. McShane, & B. Scassellati (Eds.), *Proceedings of the 36th annual meeting of the Cognitive Science Society*. Austin, TX: Cognitive Science Society.
Kort, W., Schittekatte, M., Bos, K. V., Vermeir, G., Spelberg, H. L., Verhaege, P., et al (2005). *Dyslexie Screening Test (DST), Handleiding*. London: Harcourt Test Publishers.
Kuperman, V., & Van Dyke, J. A. (2011). Effects of individual differences in verbal skills on eye-movement patterns during sentence reading. *Journal of Memory and Language, 65*, 45–73.
Leclercq, A.-L., & Majerus, S. (2010). Serial-order short-term memory predicts vocabulary development: Evidence from a longitudinal study. *Developmental Psychology, 46*, 417–427.
Lum, J. A. G., Ullman, M. T., & Conti-Ramsden, G. (2013). Procedural learning is impaired in dyslexia: Evidence from a meta-analysis of serial reaction time studies. *Research in Developmental Disabilities, 34*, 3460–3476.
Lyon, G. R., Shaywitz, S. E., & Shaywitz, B. A. (2003). A definition of dyslexia. *Annals of Dyslexia, 53*, 1–14.
Majerus, S., & Boukebza, C. (2013). Short-term memory for serial order supports vocabulary development: New evidence from a novel word learning paradigm. *Journal of Experimental Child Psychology, 116*, 811–828.
Majerus, S., Poncelet, M., Greffe, C., & Van der Linden, M. (2006). Relations between vocabulary development and verbal short-term memory: The relative importance of short-term memory for serial order and item information. *Journal of Experimental Child Psychology, 93*, 95–119.
Martinez Perez, T., Majerus, S., Mahot, A., & Poncelet, M. (2012). Evidence for a specific impairment of serial order short-term memory in dyslexic children. *Dyslexia, 18*, 94–109.
Martinez Perez, T., Majerus, S., & Poncelet, M. (2012). The contribution of short-term memory for serial order to early reading acquisition: Evidence from a longitudinal study. *Journal of Experimental Child Psychology, 111*, 708–723.
Martinez Perez, T., Majerus, S., & Poncelet, M. (2013). Impaired short-term memory for order in adults with dyslexia. *Research in Developmental Disabilities, 34*, 2211–2223.
McArthur, G. M., Hogben, J. H., Edwards, V. T., Heath, S. M., & Mengler, E. D. (2000). On the "specifics" of specific reading disability and specific language impairment. *Journal of Child Psychology and Psychiatry, 41*, 869–874.
Melby-Lervag, M., Lyster, S.-A. H., & Hulme, C. (2012). Phonological skills and their role in learning to read: A meta-analytic review. *Psychological Bulletin, 138*, 322–352.
Menghini, D., Finzi, A., Benassi, M., Bolzani, R., Facoetti, A., Giovagnoli, S., ... Vicari, S. (2010). Different underlying neurocognitive deficits in developmental dyslexia: A comparative study. *Neuropsychologia, 48*, 863–872.
Mosse, E. K., & Jarrold, C. (2008). Hebb learning, verbal short-term memory, and the acquisition of phonological forms in children. *Quarterly Journal of Experimental Psychology, 61*, 505–514.
Mosse, E. K., & Jarrold, C. (2010). Searching for the Hebb effect in Down syndrome: Evidence for a dissociation between verbal short-term memory and domain-general learning of serial order. *Journal of Intellectual Disability Research, 54*, 295–307.
Muter, V., & Diethelm, K. (2001). The contribution of phonological skills and letter knowledge to early reading development in a multilingual population. *Language Learning, 51*, 187–219.
Page, M. P. A., & Norris, D. (2009). A model linking immediate serial recall, the Hebb repetition effect, and the learning of phonological word forms. *Philosophical Transactions of the Royal Society B: Biological Sciences, 364*, 3737–3753.
Parmentier, F. (2011). Exploring the determinants of memory for spatial sequences. In A. Vandierendonck & A. Szmalec (Eds.), *Spatial working memory* (pp. 67–86). Hove, UK: Psychology Press.
Pavlidou, E. V., Kelly, M. L., & Williams, J. M. (2010). Do children with developmental dyslexia have impairments in implicit learning? *Dyslexia, 16*, 143–161.
Pennington, B. F. (2006). From single to multiple deficit models of developmental disorders. *Cognition, 101*, 385–413.
Perfetti, C. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading, 11*, 357–383.
R Development Core Team (2009). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. ISBN 3-900051-07-0. <http://www.R-project.org>.
Ramus, F., & Ahissar, M. (2012). Developmental dyslexia: The difficulties of interpreting poor performance, and the importance of normal performance. *Cognitive Neuropsychology, 29*, 104–122.
Ramus, F., & Szenkovits, G. (2008). What phonological deficit? *Quarterly Journal of Experimental Psychology, 61*, 129–141.
Raven, J. C., Court, J. H., & Raven, J. (1992). *Manual for Raven's Progressive Matrices and vocabulary scales*. Oxford, UK: Oxford Psychologists Press.
Rohl, M., & Pratt, C. (1995). Phonological awareness, verbal working memory, and the acquisition of literacy. *Reading and Writing, 7*, 327–360.
Service, E. (1992). Phonology, working memory, and foreign-language learning. *Quarterly Journal of Experimental Psychology A: Human Experimental Psychology, 45*, 21–50.
Shaywitz, S. E., Escobar, M. D., Shaywitz, B. A., Fletcher, J. M., & Makuch, R. (1992). Evidence that dyslexia may represent the lower tail of a normal distribution of reading ability. *New England Journal of Medicine, 326*, 145–150.
Smith-Spark, J. H., & Fisk, J. E. (2007). Working memory functioning in developmental dyslexia. *Memory, 15*, 34–56.
Snowling, M. J. (2000). *Dyslexia* (2nd ed.). Oxford, UK: Blackwell.
Staels, E., & Van den Broeck, W. (2014). Order short-term memory is not impaired in dyslexia and does not affect orthographic learning. *Frontiers in Human Neuroscience, 8*. http://dx.doi.org/10.3389/fnhum.2014.00732.
Staels, E., & Van den Broeck, W. (2015). No solid empirical evidence for the SOLID (serial order learning impairment) hypothesis of dyslexia. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 41*, 650–669.
Stanovich, K. E. (1988). Explaining the differences between the dyslexic and the garden-variety poor reader: The phonological-core variable-difference model. *Journal of Learning Disabilities, 21*, 590–604.
Starck, R. E., & Tallal, P. (1988). *Language, speech, and reading disorders in children*. Boston: Little, Brown.

Stichting Dyslexie Nederland (2008). *Diagnose dyslexie. Brochure van de Stichting Dyslexie Nederland* Herziene versie. Bilthoven: Stichting Dyslexie Nederland.

Szmalec, A., Duyck, W., Vandierendonck, A., Barberá-Mata, A. B., & Page, M. P. A. (2009). The Hebb repetition effect as a laboratory analogue of novel word learning. *Quarterly Journal of Experimental Psychology, 62*, 435–443.

Szmalec, A., Loncke, M., Page, M. P. A., & Duyck, W. (2011). Order or disorder? Impaired Hebb learning in dyslexia. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37*, 1270–1279.

Szmalec, A., Page, M. P. A., & Duyck, W. (2012). The development of long-term lexical representations through Hebb repetition learning. *Journal of Memory and Language, 67*, 342–354.

Tallal, P., Allard, L., Miller, S., & Curtiss, S. (1997). Academic outcomes of language impaired children. In C. Hulne & M. Snowling (Eds.), *Dyslexia: Biology, cognition, and intervention* (pp. 167–181). London: Whurr Press.

Van den Bos, K., Spelberg, H. L., Scheepstra, A., & de Vries, J. (1994). *De Klepel vorm A en B: Een test voor de leesvaardigheid van pseudowoorden—Verantwoording, handleiding, diagnostiek en behandeling*. Nijmegen, Netherlands: Berkhout.

Verhoeven, L. (1995). *Drie-Minuten-Toets: Handleiding*. Arnhem, Netherlands: Cito.

Whitney, C. (2001). How the brain encodes the order of letters in a printed word: The SERIOL model and selective literature review. *Psychonomic Bulletin and Review, 8*, 221–243.

Whitney, C., & Cornelissen, P. (2005). Letter-position encoding and dyslexia. *Journal of Research in Reading, 28*, 274–301.