scientific reports

Check for updates

OPEN

Cost and benefit of parafoveal information during reading acquisition as revealed by finger movement patterns

Viet Chau Linh Nguyen[1,3], Thomas Perret[1,3], Valentine Fabre[1], Alice Gomez[1,3,4] & Angela Sirigu[1,2,4]✉

Contrary to expert readers, children learning to read have limited ability to preprocess letters in parafoveal vision. Parafoveal letters induce crowding cost: the features of neighboring letters interfere with target letter identification. We longitudinally studied the weight of parafoveal cost and benefit in two group of children (*N* = 42), during their first school year (Group 1) and at the end of second school year (Groupe 2). Using a novel digit-tracking method, a blurred text was presented and rendered unblurred by touching the screen, allowing the user to discover a window of visible text as the finger moved along it. We compared two conditions: (1) a large window, where crowding was enhanced by the presence of parafoveal information; (2) a small window, where crowding was suppressed by blurred parafoveal information. Finger kinematics were simultaneously recorded. We found that at the beginning of first-grade, digital fixations - brief slowing or stopping of the finger on a specific point - are significantly longer in the large compared to the small window condition, as parafoveal crowding increases text processing difficulty. This effect diminishes and disappears at the end of second-grade as reading performance improves. In the large window condition, longer digital saccades - rapid movements of the finger changing position - appear by the end of first grade suggesting that parafoveal exposure become more beneficial than harmful when children acquire basic reading skills. Our results show that in beginning readers, crowding has a cognitive cost that interfere with the speed of the learning reading process. Our findings are relevant to the field of education by showing that visual crowding in first grade should not be underestimated.

Keywords Reading acquisition, Digital tracking, Parafoveal cost, Parafoveal benefit

During reading, expert readers can pre-process a certain amount of information in parafoveal vision, a phenomenon known as 'parafoveal preview benefit', which contributes to reading speed[1,2]. This process is affected by the difficulty of the reading content, and the benefit is greater for high frequency words[3,4]. The cost and benefit of parafoveal previewing vary according to reading expertise. Readers with advanced reading skills benefit most from parafoveal preview[5–8] and this start usually at Grade 2[9,10]. In young beginners, on the other hand, exposure to parafoveal information is costly due to the crowding effect[11,12]. Crowding is a basic low-level to mid-level general phenomenon of visual perception in which identifying a target item is more difficult when surrounded by flankers than when presented in isolation[13–16]. Visual crowding not only sets a bottleneck for object recognition but also impacts reading[17,18]: Resistance to crowding in pre-school children has been shown to predict future reading skills[19], and elevated sensitivity to crowding is found in dyslexia, a reading disorder associated with problems in identifying speech sounds and in learning how these correspond to written letters and words[20,21]. Crowding is more detrimental as eccentricity increases[13,17,22,23], and dyslexics exhibit a stronger crowding effect in the parafoveal field than neurotypical readers[24]. Despite several studies on reading disorders, we still do not know exactly to what extent parafoveal crowding impacts reading when children learn to read, especially in their first grade. Yet, it can be hypothesized that young beginners, like adults, are affected

[1]Institute of Cognitive Science Marc Jeannerod, Centre National de la Recherche Scientifique, Bron 69675, France. [2]iMIND, Center of Excellence for Autism, Bron, France. [3]Present address: Trajectoires team (VCLN), EDUWELL team (AG), Lyon Neuroscience Research Center, Inserm U1028, CNRS, Lyon 1 University, Lyon, France. [4]These authors contributed equally: Alice Gomez and Angela Sirigu. ✉email: sirigu@isc.cnrs.fr

by crowding effects in parafoveal vision[13,17,22,23]. Therefore, removing parafoveal letters should reduce parafoveal crowding and facilitate reading processes.

A way of studying the mechanisms of parafoveal processing during reading and the impact of crowding in children is through the analysis of eye movements. Studies have shown that the pattern of eye movements during reading are considered a good reflective of reading skills[25,26], and that their signature can reveal the amount and the type of written materiel being processed[27].

A reading strategy based on a serial mode of word processing (due to the limited number of written units that can be processed in parallel) could raise the number of fixations and reduce saccades length[28–30].

Longer fixation times in reading can be ascribed to decoding difficulties caused by insufficient language skills[3,31,32] but can also be attributed to the crowding effect[33]. Saccades, are required to bring the word into the foveal field where it is analyzed with the highest acuity[34]. In expert readers, parafoveal processing enables the retrieval of different types of information in a single fixation[1,35–38], facilitating longer saccades to reach a more distant position[34,39]. There is a position near the center of the word that maximizes its recognition rate, known as the optimal viewing position[40]. Word characteristics and parafoveal processing may also impact the position where the eye lands within each written word[41,42].

Despite their great potential in providing more in-depth understanding of children's reading skills, eye-movements data in beginner readers remains rare given the technical constraints[25]. For example, the study of large cohorts of children in schools is tricky because of the time spent calibrating the eyes, fixing the head, or the presence of external factors such as glasses, the intensity of the light in the test room, etc.

In this context, an alternative research strategy is to collect finger movement measures to gain insights about children reading strategies. As shown by Lio et al.[43], when subjects are asked to explore blurred images with their index finger, measurements of digital movements are strongly correlated with eye movements[43] and therefore can be considered a reliable proxy of eye movements exploration. This method, known as digit-tracking, can be applied in the field of reading, where blurred text can be unblur by touching the screen with index finger, revealing a window of visible text. In this way, finger position and movement dynamics can be recorded in real time to reveal reading strategies. With this method, the same measures used in the field of eye movements can be applied this time to the fingers, i.e., digital saccades and digital fixations.

Another advantage of the digit-tracking technique in the domain of reading acquisition is the display format which makes it possible to manipulate the information available in the parafoveal field. In a default modality, the size of the unblurred window discovered by the finger simulates the size of the foveal field while the blur outside the window represents the low acuity of peripheral vision.

In the present study, we manipulated the size of the parafoveal window to reduce or enhance the crowding effect. We investigated in young children the dynamics and evolution of digital movements (digital fixations and digital saccades) during the process of learning to read by using a longitudinal design to compare in the parafoveal field the effects of high (large unblurring window condition) versus low information load (small unblurring window condition) on reading performance and finger movements pattern. We manipulated the window size, to investigate cost and benefit of parafoveal processing as a function of reading abilities. We assessed children's reading performances at three different points in time: at the beginning and end of Grade 1, and at the end of Grade 2. We expected both temporal and spatial digital measures to change over time as a function of the amount of parafoveal information provided in the window size. As crowding would affect reading performance mainly in Grade 1 (to disappear at Grade 2), the small window size designed to reduce crowding would minimize cost, and this effect was expected to be evident in digital movement patterns as already shown for eye movements[9,10]. Thus, we expected changes over time as children become more expert readers and somewhat less affected by visual crowding[11,12]. We expect no difference between the window conditions at Grade 2, with performances similar to that of an expert reader (longer saccades, fewer fixations and a landing position closer to the centre of the word). Hence, that the crowding effect would disappear as reading skills improved and that the advantage of parafoveal preview would become evident in the second or third year of school. In experts however, we anticipated that, the small-window condition, by preventing parafoveal treatment, may lead to high cost.

Results

Hereafter we report the results on finger kinematics for each time point (time point 1, t1, which occurs at the beginning of Grade 1 and 12 months after time point 0, t0; time point 2, t2, which occurs at the end of Grade 1, and 18 months after time point 0; time point 3, t3, which occurs at the end of Grade 2 and 30 months after time point 0), for window conditions (large vs. small) and decoding ability (good vs. poor decoders) using two sets of sentences.

Following we describe results on the Language and Spelling Assessment Battery *(*BELO, *Batterie d'Evaluation du Langage et de l'Orthographe*) tests while the additional reading tests (meaningless and meaningful text reading, pseudowords reading) are reported in the Supplementary Information.

Reading performance levels on the BELO test were used to classify children as good or poor decoders, and scores on the additional reading tests as good or poor readers.

Digital movements were investigated at both sentences and word level (results for words and sentences processing are reported in Supplementary Information) Changes in finger movements for both small and large windows were examined in relation to cognitive and reading scores using multiple regression analysis (see Supplementary Information). Further, the effect of window size on decoding ability was analyzed at each timepoint for reading fluency. We set a significance level at $p < 0.05$ for all statistical analyses and Bonferroni correction was applied for post-hoc comparisons to control for Type I errors. The significance thresholds were as follows: for Timepoint effect (3 comparisons), $\alpha_{Bonferroni} = 0.05/3 = 0.017$; for Timepoint x decoding ability and Timepoint x unblurring size (15 comparisons), $\alpha_{Bonferroni} = 0.05/15 = 0.0033$; for Timepoint x decoding ability x unblurring size (66 comparisons), $\alpha_{Bonferroni} = 0.05/66 = 0.0008$.

Sentences-level variables

Data on finger kinematics were analyzed with the following Linear Mixed Models (LMM):

variable ~ timepoint x decoding ability x unblurring window size + (1|subject) with *timepoint* as a three-level factor: t1 vs. t2 vs. t3; *decoding ability* as a two-level factor: good vs. poor and *unblurring window size* as a two-level factor: small vs. large). The dependent variable (*variable*) was modeled as a function of the independent variables (time point, decoding ability, unblurring window size) and a random intercept for each subject, using a linear mixed effects model. Contrast effects of the sentence-level variable are reported in Table 1.

The analysis on finger movements at the sentence level showed that digital fixations, i.e., when the finger slows down beyond a defined threshold (see Material and Methods), are significantly longer at the beginning of first grade compared to those observed at the end of first and second grade. The poor decoders showed longer digital fixations than good decoders and these fixations were longer in the large (unblurring) window condition compared to the small one. At the beginning of first grade, good decoders made significantly shorter digital fixations comparing to poor decoders. The difference between the large and small window conditions was significant at the beginning of first grade, but not at later stages.

Below we report the main effects of timepoints and windows size on digital saccade length and the interaction between decoding performance, window size and timepoints. Overall, the results show that poor decoders show shorter digital saccades than good decoders at the beginning of Grade 1, but not at the end of grade 1 and grade 2. The difference between large and small windows size was not significant at the beginning of Grade 1 while it was at the end of grade 1 and 2. Concerning the number of digital fixations, we found a main effect of timepoint, i.e., a significant increase between the beginning and end of first grade while for decoding, poor decoders showed higher digital fixations compared to good decoders and these were modulated by the window size, with a significant higher number of digital fixations in the large window condition compared to the small one.

Duration of digital fixationsc

We predicted a greater cost of crowding at early stages at Grade 1.

As shown in Fig. 1A; Table 1, digitalfixation duration was longer for the large window compared to the small one. We also observed an interaction between the timepoint and the window size: the difference between the large and small window was robust at the beginning of first grade but not at the end of Grade 1 and at Grade 2. The interaction between unblurring window size and decoding ability and the triple interaction were not significant. This suggests that decoding ability was not affected by crowding when measured through the average duration of digital fixations. Hence, the cost of crowding is greater on early than late stage of reading acquisition.

Digital saccade length

We predicted that crowding should have greater cost than benefit on decoding abilities at early stages and thus, that the difference in the digital saccade length between large and small unblurring windows will not be significant at the beginning of Grade 1, but it will emerge at later stages. Thus, the cost of crowding should be reduced earlier in good decoders than in poor decoders, and thus, the difference in the digital saccade length between large and small unblurring windows will increase earlier in the group of good than poor decoders.

As reported in Fig. 1B, for digital saccade length analysis, we observed an effect of window size: Digital saccades were significantly longer in the large window compared to the small one. Double interaction effects were also observed for window size x timepoint and window size x decoding ability. The difference in digital saccade length between large and small window conditions is observed only at the two last stages with the most robust effect at the end of second grade. The triple interaction analysis on timepoint x decoding ability x unblurring window size is marginally significant. The difference between large vs. small window is not significant at t1 for both good and poor decoders, but it reached significance for good decoders only at the end of first grade, and for both good and poor decoders at the end of second grade.

This triple interaction was also observed when we divided readers into good versus poor readers based on performance on the additional reading tests (see Supplementary Information Table S2, Table S3), suggesting that the interaction effect with decoding ability is robust. Contrary to what was observed in the average duration of digital fixation, the cost of crowding on saccade length was modulated both by children's' age and their decoding ability.

Average number of digital fixations

As reading skills improve we expected the small window to reduce access to parafoveal information, therefore inhibiting parafoveal benefit contray to the large window condition. As a result, we expected the number of digital fixations greater in the small window than in the large one. We also predicted that parafoveal preprocessing would be more important at the end of grade 2 than at the beginning and end of Grade 1. The difference between large and small window conditions on the number of digital fixations should increase at later stages. Within this context, the benefit of parafoveal preview may appear earlier in good decoders than in poor decoders, and the difference in the number of digital fixations between the large and small window conditions should increase earlier in the group of good decoders but not in the group of poor decoders. We observed an effect of the window size on the average number of digital fixations, fewer digital fixations were made in the large window condition compared to the small one (Fig. 1C). However, there found no interaction between unblurring window size and timepoint or decoding ability and no triple interaction suggesting that the effect of parafoveal preview in the small window condition had a similar cost on the number of digital fixations at all timepoints and decoding abilities.

Main effect	F-value	*p*-value	Contrast	b	SE	t-value	*p*-value
Digital fixation duration							
Stage	**129.39**	**< 0.001**	t1-t2	**379.23**	**75.55**	**5.02**	**< 0.001**
			t1-t3	**463.87**	**29.36**	**15.80**	**< 0.001**
			t2-t3	84.64	77.50	1.09	0.280
Decoding ability	**9.29**	**0.003**	good-poor	**-146.89**	**48.19**	**-3.05**	**0.003**
Unblurring window	**37.71**	**< 0.001**	large-small	**141.24**	**23.00**	**6.14**	**< 0.001**
Interaction effect							
Stage * decoding ability	**4.32**	**0.015**	t1:good - poor	**-263.62**	**53.62**	**-4.92**	**< 0.001**
			t2:good - poor	-132.99	118.68	-1.12	0.269
			t3:good - poor	-44.06	57.76	-0.76	0.446
Stage * unblurring window	**3.83**	**0.022**	t1:large-small	**218.85**	**30.34**	**7.21**	**< 0.001**
			t2:large-small	83.99	42.81	1.96	0.050
			t3:large-small	120.86	44.87	2.69	0.007
Unblurring window * decoding ability	0.42	0.514					
Stage * decoding ability * unblurring window	0.09	0.917					
Digital saccade length							
Stage	**344.24**	**<0.001**	t1-t2	**-1.27**	**0.41**	**-3.11**	**0.004**
			t1-t3	**-2.66**	**0.10**	**-26.22**	**<0.001**
			t2-t3	**-1.39**	**0.41**	**-3.37**	**0.002**
Decoding ability	0.56	0.457	good-poor	0.18	0.23	0.75	0.457
Unblurring window	**92.19**	**<0.001**	large-small	**0.75**	**0.08**	**9.60**	**<0.001**
Interaction effect							
Stage * decoding ability	**12.56**	**<0.001**	t1:good - poor	**-0.99**	**0.19**	**-5.11**	**<0.001**
			t2:good - poor	1.41	0.64	2.22	0.033
			t3:good - poor	0.10	0.20	0.49	0.626
Stage * unblurring window	**41.68**	**<0.001**	t1:large-small	-0.05	0.10	-0.53	0.598
			t2:large-small	**0.68**	**0.14**	**4.76**	**<0.001**
			t3:large-small	**1.62**	**0.16**	**10.41**	**<0.001**
Unblurring window * decoding ability	**16.05**	**<0.001**	good: large-small	**1.06**	**0.11**	**9.41**	**<0.001**
			poor: large-small	**0.44**	**0.11**	**4.05**	**<0.001**
Stage * decoding ability * unblurring window	2.51	0.081	t1:good: large-small	0.04	0.15	0.28	0.781
			t1:poor: large-small	-0.15	0.20	-1.10	0.271
			t2:good: large-small	**1.10**	**0.33**	**5.46**	**<0.001**
			t2:poor: large-small	0.26	0.20	1.27	0.203
			t3:good: large-small	**2.04**	**0.22**	**9.02**	**<0.001**
			t3:poor: large-small	**1.20**	**0.21**	**5.62**	**<0.001**
Digital saccade length							
Stage	**344.24**	**< 0.001**	t1-t2	**-1.27**	**0.41**	**-3.11**	**0.004**
			t1-t3	**-2.66**	**0.10**	**-26.22**	**< 0.001**
			t2-t3	**-1.39**	**0.41**	**-3.37**	**0.002**
Decoding ability	0.56	0.457	good-poor	0.18	0.23	0.75	0.457
Unblurring window	**92.19**	**< 0.001**	large-small	**0.75**	**0.08**	**9.60**	**< 0.001**
Interaction effect							
Stage * decoding ability	**12.56**	**< 0.001**	t1:good - poor	**-0.99**	**0.19**	**-5.11**	**< 0.001**
			t2:good - poor	1.41	0.64	2.22	0.033
			t3:good - poor	0.10	0.20	0.49	0.626
Stage * unblurring window	**41.68**	**< 0.001**	t1:large-small	-0.05	0.10	-0.53	0.598
			t2:large-small	**0.68**	**0.14**	**4.76**	**< 0.001**
			t3:large-small	**1.62**	**0.16**	**10.41**	**< 0.001**
Unblurring window * decoding ability	**16.05**	**< 0.001**	good: large-small	**1.06**	**0.11**	**9.41**	**< 0.001**
			poor: large-small	**0.44**	**0.11**	**4.05**	**< 0.001**
Continued							

Main effect	F-value	*p*-value	Contrast	b	SE	t-value	*p*-value
Stage * decoding ability * unblurring window	2.51	0.081	t1:good: large-small	0.04	0.15	0.28	0.781
			t1:poor: large-small	-0.15	0.20	-1.10	0.271
			t2:good: large-small	**1.10**	**0.33**	**5.46**	**< 0.001**
			t2:poor: large-small	0.26	0.20	1.27	0.203
			t3:good: large-small	**2.04**	**0.22**	**9.02**	**< 0.001**
			t3:poor: large-small	**1.20**	**0.21**	**5.62**	**< 0.001**
Digital number of fixations							
Stage	**57.44**	**< 0.001**	t1-t2	**40.92**	**9.93**	**4.12**	**< 0.001**
			t2-t3	**63.35**	**5.96**	**10.64**	**< 0.001**
			t1-t3	22.44	10.11	2.22	0.031
Decoding ability	**7.03**	**0.010**	Good-poor	**-19.22**	**7.25**	**-2.65**	**0.010**
Unblurring window	**20.70**	**< 0.001**	Large-small	**-22.01**	**4.84**	**-4.55**	**< 0.001**
Interaction effect							
Stage * decoding ability	2.66	0.075					
Stage * unblurring window	0.42	0.657					
Unblurring window * decoding ability	0.03	0.855					
Stage * decoding ability * unblurring window	0.21	0.809					

Table 1. LMM values (main and interactions effects) for sentences-level analysis (digital fixation duration, digital saccade length, number of digital fixations). Significant effects (absolute t value: |t| > 1.96, $p < 0.05$) are in bold. Utilizing Bonferroni adjustment with an adjusted α(Bonferroni) = 0.05/X, where X represents the number of comparisons contrasts from non-significant interactions are not reported.

Fixation maps

Figure 2 shows the individual digital fixation maps as a function of timepoints for a poor decoder (left) (decoding score = 32/85 at t_1, 76/85 at t_3) with long and frequent digital fixations and short length saccades, and a good decoder (decoding score = 79/85 at t_1 and 84/85 at t_3) with shorter digital fixations and longer digital saccades (Fig. 2). We observe the decrease in the number of digital fixations and the duration in fixations, as well as the increase in the length of the digital saccades between the two time points. In addition, we show the effects of the window condition: the small one characterized by shorter digital saccades, particularly at t_3, while the large by longer digital fixations durations in the early stages. These differences are more important for poorer decoders than for good decoders.

Is reading fluency influenced by the unblurring window size and decoding skills?

We predicted that reading fluency could be influenced by the level of decoding skills and window size, and that this effect might change with age. At a later stage, parafoveal benefit should become more relevant for fluency. At the beginning of Grade 1, fluency is expected to be mainly influenced by decoding skills but not by the window size. At later stages, reading fluency should be influenced by decoding skills and window size. Mean and SEMs for reading fluency as a function of unblurring window size are reported in Table 2.

We used the following LMM model (timepoints t_1, t_2, t_3):

reading fluency at t_X*~unblurring window size*decoding ability at* t_X*+ (1|subject)*.

At the beginning of Grade 1, we found an effect of decoding ability on reading fluency, with good decoders reading the text more fluently than poor decoders. However, as expected, neither the size of the unblurring window nor the interaction between decoding ability and window size explained reading fluency. LMM models on reading accuracy and reading time yielded similar results (see Supplementary Information). At the end of first grade, we observed a significant effect on unblurring window size, decoding ability, and an interaction between unblurring window size and decoding ability. At this stage, children are more fluent in the large window condition compared to the small one and good decoders are more fluent than poor decoders. Moreover, good decoders are more fluent in the large window condition compared to the small one, which is not the case for poor decoders. At the end of second grade, only the effect of unblurring window size is significant: children at this stage are more fluent in the large unblurring window than in small one but decoding skills no longer distinguish children on the reading fluency task.

Discussion

In this study, we investigated the effect of parafoveal crowding in young beginning readers (first to second grade) using a novel procedure (digital tracking) in which blurred words and sentences were presented to children who were asked to read them by unblurring the text with the index finger. Finger movements were recorded in two blur conditions of small and large windows. Compared with the small window condition, the large window condition was designed to induce a crowding effect due to the display of greater information in parafoveal vision, while this effect was reduced when parafoveal processing was inhibited.

We hypothesized that in first-grade beginning readers, children will be more affected by crowding. Therefore, blurring parafoveal letters as presented in the small window condition would decrease the cost of crowding and

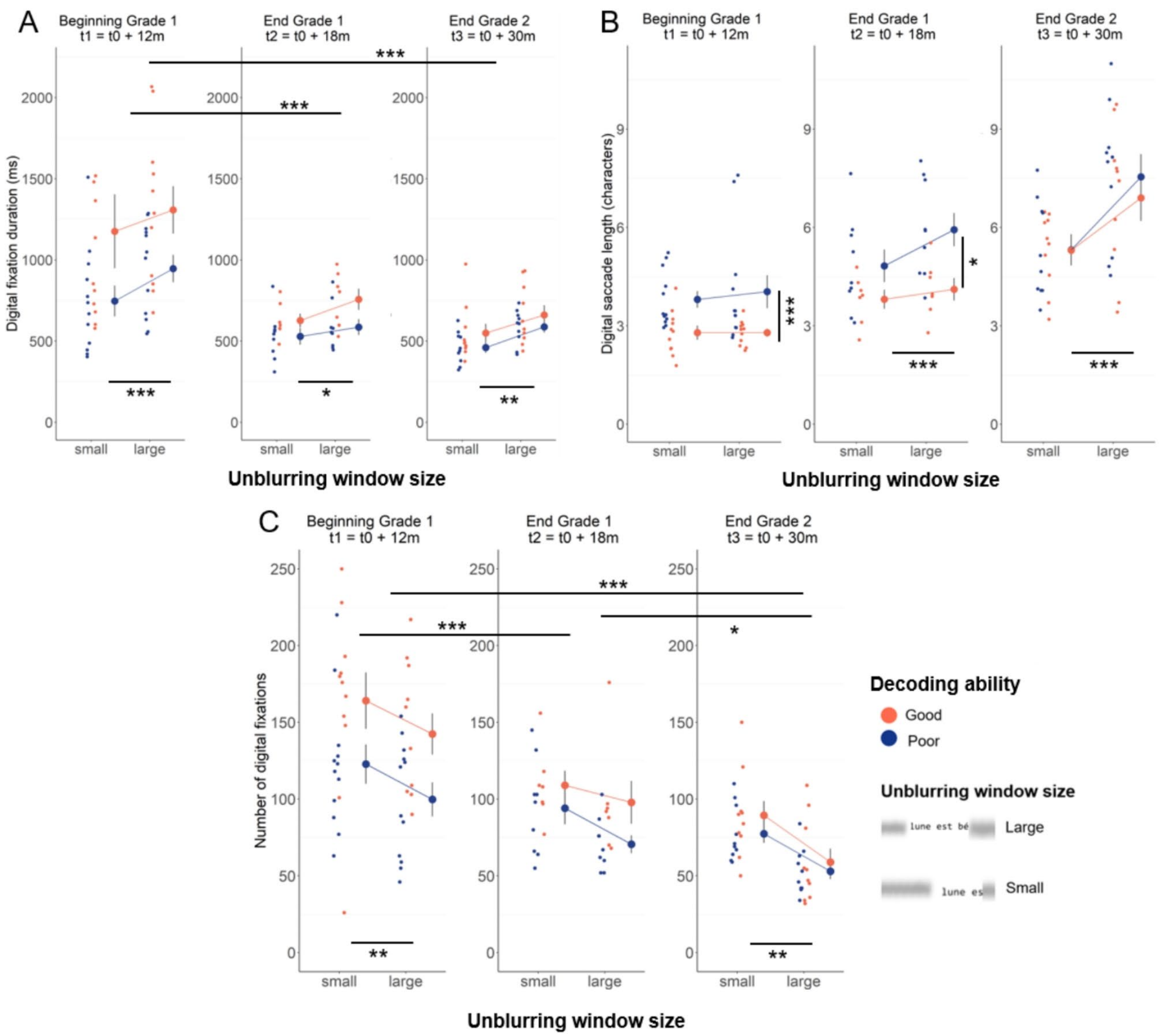

Fig. 1. Effects of unblurring window size on digital saccades as a function of decoding abilities and timepoints. Results for finger movement analysis at sentence-level for each time point (t1-left, t2-middle, t3-right), according to unblurring window condition (small vs. large) and decoding skills (good decoders in blue vs. poor decoders in red). (**A**) Average duration of digital fixations in ms. (**B**) Digital saccade length in number of characters. (**C**) Number of digital fixations made on the whole 40-words-long text. Significance level: ***:$p < 0.001$; **:$0.001 < p < 0.01$; *:$0.01 < p < 0.05$; .: $0.05 < p < 0.1$.

help learning to read more efficiently. We expected temporal measurements to differ between conditions, with shorter temporal digital fixation times only in the small window condition.

Moreover, since at this stage there is no benefit from parafoveal preview[9], we expected no effect of the window size on spatial measures (digital saccade length, number of fixations, digital landing position). As predicted, in beginners, saccade length was similar between the two window conditions, in line with studies showing poor parafoveal benefit in young readers[9,10]. It's interesting to note that for good decoders, the average length of digital saccades was 4 characters (6.4 mm), while for poor decoders it was 3 characters (4.8 mm), approximately the size of the small window. This suggests that the children targeted a nearby position despite the fact that in the large window condition, more information could be processed in parafoveal vision. The failure to use parafoveal processing limits the number of units processed at once, thus resulting in a greater number of fixations in order to be able to handle an optimal amount of information. We suggest that novice readers do not benefit from parafoveal preview during reading and that, on the contrary, this processing induces a visual cost. Furthermore, although children may have the choice of reading while ignoring information in parafoveal vision, simple passive exposure to this content is likely to trigger a large number of fixations, thus disrupting reading processes in foveal vision.

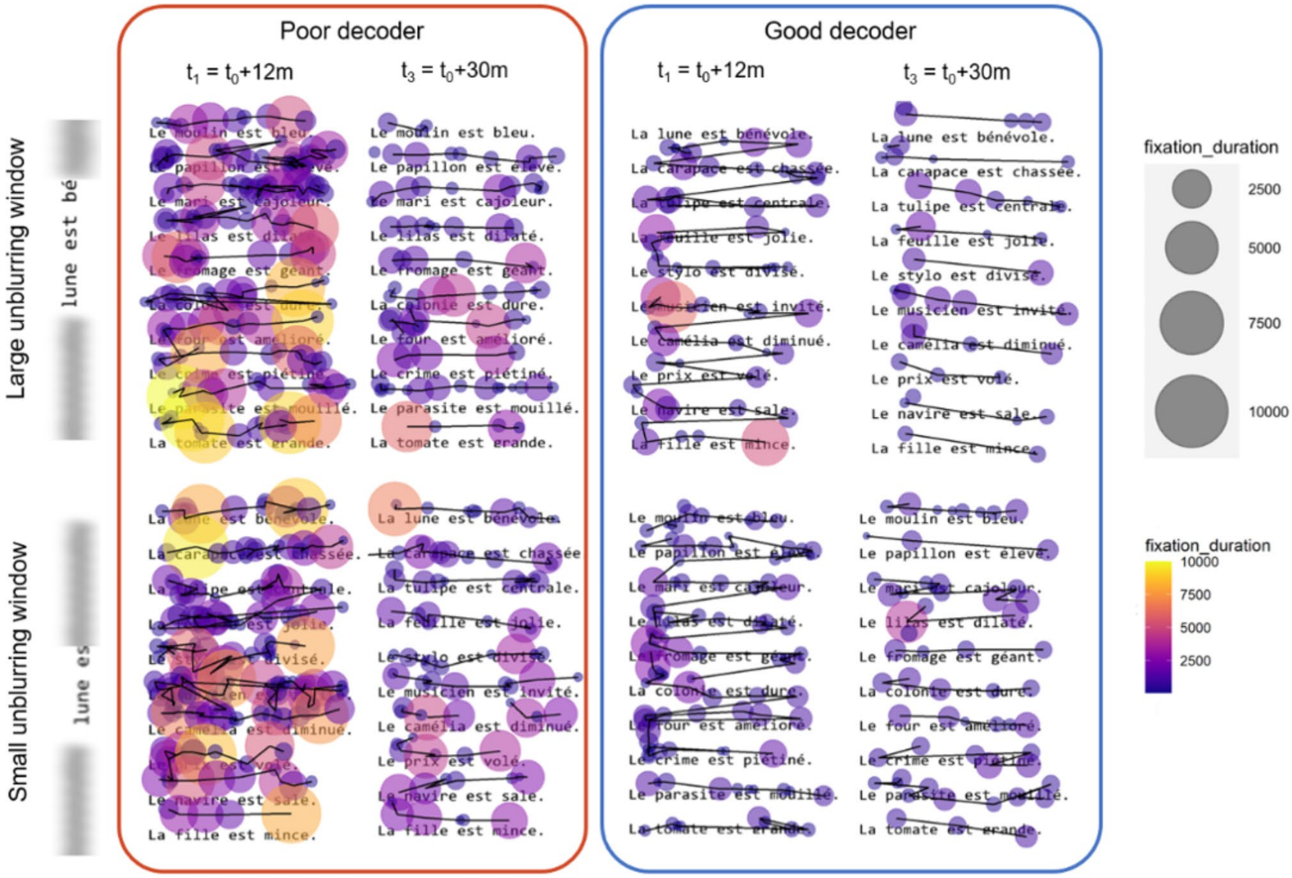

Fig. 2. Patterns of digital movements in poor and good readers across grade timepoints. Digital fixation maps of a poor (left) and a good (right) decoder at t_1 (beginning of Grade 1) and at t_3 (end of second grade). Decoding scores: poor decoder = 32/85 at t_1 and 76/85 at t_3, small blur; good decoder = 79/85 at t_1 and 84/85 at t_3. Circles represent digital fixations whose durations are proportional to the circles' radius. Color gradient indicates fixation duration (longer durations are represented with warmer color). Segments correspond to the length of saccades and color variation represents the order of digital fixations in a given trial.

Unblurring window size	t1 (Beginning Grade 1)	t2 (End Grade 1)	t3 (End Grade 2)
Large	13.52 (2.52) wpm	25.89 (3.12) wpm	37.67 (3.01) wpm
Small	12.58 (2.11) wpm	22.71 (2.19) wpm	30.65 (2.51) wpm
p-values	0.798	0.018	< 0.001

Table 2. The effect of unblurring window size on words reading and grade timepoints. Mean (SEMs) fluency (correct words read per minutes, wpm) for different unblurring window sizes and timepoints (t_1, t_2, t_3).

As for the temporal measurements, as expected, the digital movement patterns at t1 (average duration of digital fixation, duration of first digital fixation and duration of gaze on word) revealed a negative effect of the large window condition, which we interpret as the result of crowding, in line with what has been shown previously in the eye movements literature[21].

Finally, our results show that the negative effect of crowding in beginners during the first year of school appears to be independent of children's reading decoding abilities or general cognitive performance.

Thus, in line with eye movement studies, our results corroborate the hypothesis that parafoveal preprocessing induces crowding that plays a negative role in learning to read. To our knowledge, this is the first time that the cost of crowding has been demonstrated in young beginning readers using digital movements as a proxy for reading performance. We conclude that parafoveal previewing has a cost and interferes with the process of learning to read in the first year of school.

Future studies should confirm our findings in a larger population and examine whether other factors, such as non-linguistic material, also trigger crowding or whether word knowledge mitigates it instead. For example, it would be interesting to compare two large window conditions, one with normal letters and the other containing letters replaced by 'x', as in the moving window paradigm[1] in order to dissociate the basic visual effect from the effect linked to linguistic information.

Overall, these results suggest that if exposure to parafoveal information is limited early in the first year, the negative effect of crowding for learners will be reduced. Some compensatory methods have been designed to provide a crowding-free reading environment for people with loss of foveal vision and those with dyslexia[44–46]. However, this problem has not been addressed in young neurotypical readers. Although our results need to be replicated in a wider population, they raise questions about how to address the problem of reading difficulties at school arising from the basic phenomenon of visual crowding. Our digit-tracking method can be applied as a compensatory tool for children at the beginning of Grade 1, and we have shown that learning to read with suppressed parafoveal information stimulates reading acquisition in children[47].

Interestingly, at later stages (t2 and t3), as children improve their reading, they are less at risk of peripheral crowding and begin to benefit from parafoveal preview, as shown in the large window condition. We predicted that the trajectory of finger movements in this case would be like that of an expert adult reader (longer saccades, less fixations and landing finger position nearer to the word center) and we also anticipated that the removal of the parafoveal area would be detrimental in this case. Our results confirm this hypothesis showing that at t_2 (end of first grade), there were no differences between the two window conditions on temporal digital measures. The effect of crowding is therefore considerably reduced at this stage, probably because children are better at decoding orthographic units of words and syllables[48]. With regard to the length of the digital saccade, we observed two different patterns linked to the children's decoding skills: good decoders produced longer saccades in the large condition than in the small condition, while digital saccades of poor decoders were of similar length in both conditions. Our interpretation is that, in the large window condition, the good decoders, benefitting from parafoveal information, process the written units in a single fixation and focus simultaneously on the following units at distant locations.

On the other hand, the small window condition, by preventing parafoveal pre-processing, forces them to decode using short digital saccades, similar to those observed in poor decoders.

At the end of second grade (t_3), we observed longer fixation durations for the large window condition but, no difference for the duration of the first digital fixation as well as for their total duration. At the same time, the number fixations within words (see Supplementary Information 3.1.3) was significantly lower in the large than in the small window condition.

No effect of decoding ability was observed, since both good and poor decoders were able to benefit from parafoveal preview and made longer saccades in the large compared to the small window condition. Also, the differences in saccade length between the two window conditions do not appear to be related on children's general cognitive performance (Supplementary Information).

A large body of research has already investigated the role of parafoveal information in reading. Most of these studies were based on the moving window task and the gaze contingent paradigm[2] where the parafoveal information was manipulated by replacing letters with meaningless characters (i.e., a string of 'x's to replace letters). This approach has been questioned[49,50] as the parafoveal preview (even if masked by the string of 'x's) induces a cost that could interact with parafoveal processing. Therefore, the 'x's mask paradigm cannot be considered a valid baseline for inferring the magnitude of parafoveal benefit. Marx and colleagues used another method, called the "incremental boundary paradigm", which involves degrading parafoveal letters by the random displacement of black pixels[51]. They showed that when 20% of pixels were degraded, children experienced a reduced parafoveal cost compared to letters that were minimally or not at all degraded. However, this result has been called into question[52], with the argument that the degradation of parafoveal letters always induces involves a cost associated to the phenomenon of *change awareness*, (i.e., the attentional resources activated when, at the onset of the fixation, readers are surprised by the change from the initial text to a degraded preview. In this context, we can suggest that our digit-tracking method, by fully masking the parafoveal field without causing unexpected stimulus changes, can attenuate the interfering effect of parafoveal information. A study comparing the effects of all these methods for handling parafoveal material could prove important in the future for developing a more effective approach to learning to read.

Regarding the results on reading fluency, we found that they were similar at t_1 between the small and large windows, but not at t_2 and t_3 where we observed a better performance in the large window condition as shown by the lengthy digital saccades. This suggests that parafoveal previewing is beneficial and that the cost of crowding is reduced. For good decoders' fluency performance in the large window was significantly higher than in the small condition, while for poor decoders fluency didn't differ between the two conditions. At t_3, the difference between good and poor decoders in saccade length is no longer significant, with both groups showing longer saccades in the large window condition, a performance similar to what found also for reading fluency. Reading in the large window condition facilitates fluency, regardless of subjects' decoding ability. Again, these results are consistent with the study by Sperlich and colleagues, who show that parafoveal preprocessing begins between Grade 2 and Grade 3[9].

Our findings for the three developmental stages support the idea that removing parafoveal information to reduce crowding, consolidate learning and automaticity of reading in beginning readers. Crowding is a phenomenon that affect the integration of a written target with neighboring elements[53]. For children who are still in the process of learning to read, crowding can slow down their ability to recognize the basic elements of reading. Therefore, the removal of parafoveal letters at an early stage, although not improving reading fluency (we observed that reading fluency is similar in both large and small window conditions) it might facilitate familiarity and faster decoding of writing units.

A limitation of our study is that, beyond reading fluency, we provide no information on parafoveal preprocessing and other dimensions of reading such as reading comprehension. For example, is not clear how reducing the window size at the beginning of first grade may have a negative effect on reading comprehension.

In summary, our study provides evidences on the changes of parafoveal crowding and parafoveal benefit during learning to read. Our findings are in line with previous research on eyes movements and the problem of parafoveal interference in novice readers, and highlight the cost of crowding during the early reading acquisition process. These processes are closely linked as selective digital saccades (a strategy developed along with reading skills and parafoveal processing[34,39,54]), helps to reduce crowding at the target location[55]. In addition, familiarity with string characters also reduces the crowding effect[48]. In conclusion, we have demonstrated that learning to read is part of a trade-off process between parafoveal cost and parafoveal benefit, and that their weights during learning to read evolves with the acquisition of reading competences. Finally, our results also underline the importance of digital movements analysis as a window to reveal children's learning processing stage and possible difficulties in learning to read.

Materials and methods

Participants

Data were collected from two groups of children, each tested at two time points separated by an 18-months interval (Fig. 3). All children had normal or corrected-to-normal vision. No other sensory impairment was reported or detected. The study was conducted in accordance with the Declaration of Helsinki. Prior to the study, parents received information of the purpose of the study and gave their written informed consent for their child to participate. The experimental procedures were approved by the French National ethic committee (CPP n° 3574, 2017-A03065-48). All families were informed and agreed to the European regulation procedures on data protection.

Group 1

For Group 1, 19 children (Fig. 3) took part in the study at the beginning of their last year of kindergarten (t_0, pre-test phase) in November-December 2019 for general cognitive evaluations (Table 3); and at the end of their

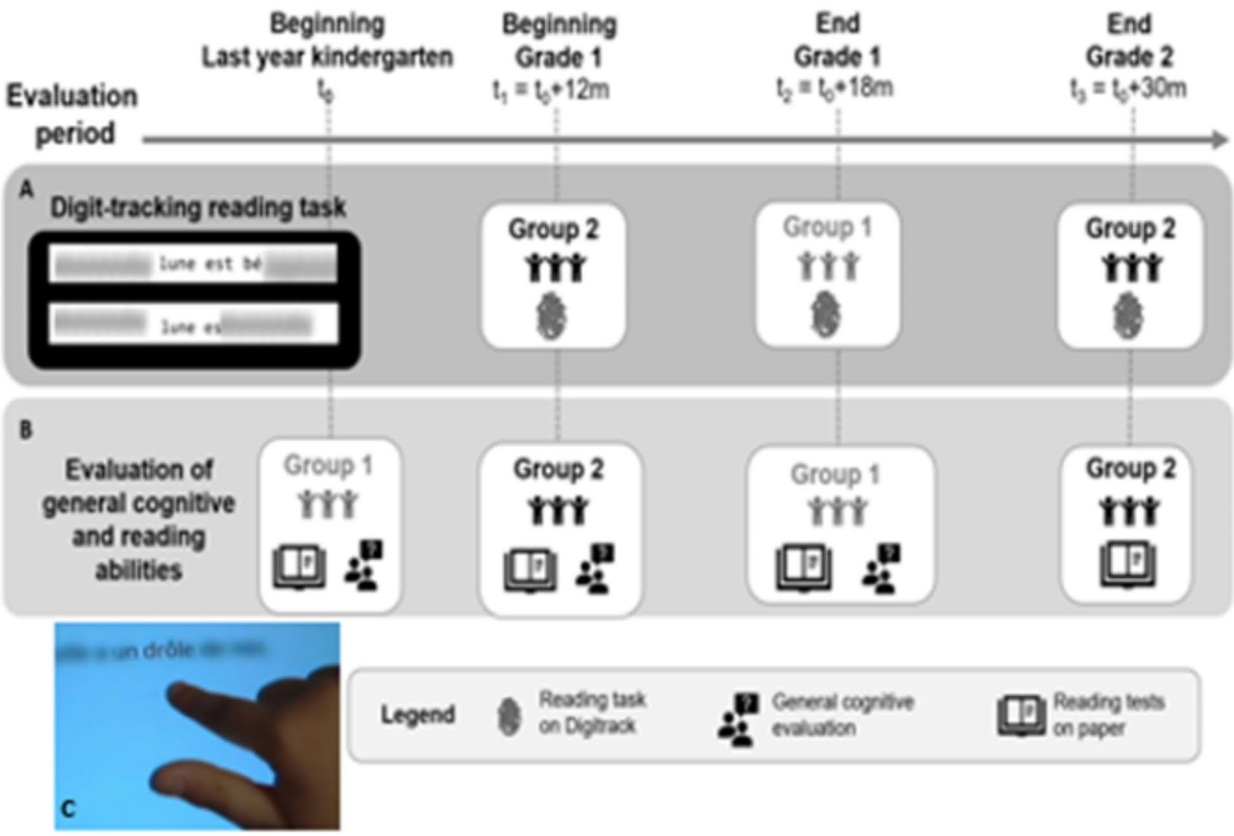

Fig. 3. Experimental procedure. (**A**) An example of digit-tracking reading task (upper left) assigned to Group 1 at the end of Grade 1 (t_2) and to Group 2 (black) at two time points: the beginning of first grade (t_1) and at the end of second grade (t_3). Written texts were presented either in large unblurring window (black box, upper line) or small window (black box, lower line). (**B**) Children's general cognitive abilities were assessed at pre-test (t_0 for Group 1 and t_1 for Group 2, except for digit-span, (see part 2.2.). General cognitive abilities) and their reading skills were assessed on paper at post-test (t_2 for Group 1 and t_3 for Group 2). C. Example of a child unblurring the text using digit-tracking in the small window condition (the large window is twice as larger in length).

		t0	t1	t2	t3
		Mean (SD)	**Mean (SD)**	**Mean (SD)**	**Mean (SD)**
Age		5.47 *(0.33)*	6.49 *(0.28)*	6.90 *(0.33)*	7.89 *(0.47)*
Reading test					
Digit-tracking reading					
	Accuracy (%)	NC	69.26 *(20.33)*	85.96 *(14.57)*	91.71 *(7.60)*
	Reading duration (s)	NC	176.74 *(72.67)*	97.73 *(28.53)*	73.25 *(23.94)*
	Fluency (correct words/minute)	NC	13.05 *(10.41)*	24.21 *(10.81)*	33.32 *(11.63)*
Letter and syllable reading		22.82 *(11.73)*	49.82 *(19.01)*	72.67 *(10.06)*	78.23 *(6.45)*
Rapid Automatized Naming					
	Images	36.69 *(13.80)*	48.93 *(16.45)*	NC	NC
	Letters	45.25 *(14.42)*	68.31 *(18.56)*	NC	NC
Visual attention span					
	Letters read	34.81 *(11.70)*	55.22 *(11.10)*	NC	NC
Alouette (Meaningless text)					
	Accuracy (%)	NC	NC	82.02 *(12.87)*	84.56 *(10.25)*
	Reading duration (s)	NC	NC	180 *(0)*	179.04 *(4.37)*
	Fluency (correct words/minute)	NC	NC	28.91 *(13.75)*	42.56 *(17.34)*
Mr Petit (Meaningful text)					
	Accuracy (%)	NC	NC	84.41 *(17.28)*	91.04 *(10.79)*
	Reading duration (s)	NC	NC	60 *(0)*	60 *(0)*
	Fluency (correct words/minute)	NC	NC	36.33 *(23.17)*	70.36 *(31.14)*
Pseudowords					
	Accuracy (%)	NC	NC	64.83 *(25.62)*	72.70 *(19.98)*
	Reading duration (s)	NC	NC	149 *(45.90)*	118.95 *(46.62)*
	Fluency	NC	NC	13.47 *(8.24)*	19.88 *(10.29)*
Other cognitive tests					
Digit Span	Direct	5.18 *(1.04)*	6.54 *(1.23)*	6.20 *(1.05)*	6.64 *(1.40)*
	Indirect	3.81 *(1.63)*	4.77 *(1.08)*	4.87 *(0.72)*	5.36 *(1.26)*
Vocabulary		9.82 *(3.18)*	11.31 *(4.79)*	NC	NC
Matrix		10.00 *(2.30)*	9.36 *(2.96)*	NC	NC
Flanker test	Incong-Cong difference (ms)	96.22 *(85.89)*	121.87 *(107.66)*	NC	NC

Table 3. Means and (SDs) of reading and general cognitive performances of each group at different time points.

Grade 1 (t_2, post-test phase) in May-June 2021 for reading tasks and for a digit-span evaluation. Data from one child was excluded because he had difficulty in understanding tasks instructions and he showed no sign of motivation during the test. Hence, data on reading performance of 18 children (8 girls) at t_2 were kept for the analysis (Mean age at pre-test = 5.47 years, $SD = 0.33$).

Group 2
For Group 2, 26 children in Grade 1 took part in the pre-test phase (t_1) in November – December 2019 and 25 of them participated in the post-test phase (t_3), eighteen months later in May-June 2021 (Fig. 3). The data set (pre-test and post-test) of three children was excluded: two of them had reading difficulties in the untimed reading tasks, and one had severe reading difficulties linked to fatigue and lack of motivation. In addition, pre-test data from 3 other children were excluded because they could only spell few letters without being able to read the word. Finally, post-test data of three additional children were also excluded for reasons not linked to reading competences (one for moving away, one for disturbing behavior and one for being very distracted during testing). In summary, we computed data of 23 children in Group 2 (10 girls, mean age at pre-test = 6.49 years, $SD = 0.28$), 16 among them had both pre-test and post-test data, 3 with only pre-test and 3 with only post-test. We decided to include subjects' performance with missing data since the linear mixed models used for our statistical analysis (see part 2.4. Statistical analysis) allow to handle this issue with robustness and optimized statistical power[56,57].

Procedure
Our procedure included 4 testing sessions (t_0, t_1, t_2, t_3) (Fig. 3): the last three sessions involved the experimental manipulation with the digit-tracking reading task (see Sect. 2.1.2). Two groups of children were assessed at two time points separated by an 18-month interval. Both groups were also tested on reading skills (at each timepoint for the BELO test and at t_2 for Group 1 and t_3 for Group 2 for the other reading tests) and on general cognitive abilities (t_0 for Group 1 and t_1 for Group 2). Reading tests were used to divide children into good and poor

readers within each group at each time point, in order to assess whether inter-group variations in reading skills is reflected in finger movement measurements as well. We will report the analysis of good versus poor decoders (based on BELO performance– letter and syllable reading test) in the main results. Analysis of good versus poor readers based on the three reading tests (meaningful text reading, meaningless text reading, pseudoword reading) is reported in Supplementary Information. General cognitive abilities were used to disentangle the effects linked to linguistic competences or potential cognitive factors such as fluid intelligence, selective attention, working memory, vocabulary or other visuo-attentional processes.

Role of test-retest effects
Because the study describes the evolution of digital behavior when learning to read, one must insure that the effect observed cannot be attributed to simple test-retest effects. Our study includes two different groups in which subjects' performance was examined longitudinally in order to assess whether changes in reading can be attributed to the learning phase or to exposure to the experimental material per se. We argue that if, test-retest effect, could have had an influence, only the data from t_1 and t_3 could have shown this bias because data acquired at t_2 came from another group of children. It is therefore very unlikely that children from group 2 had memories of the specific sentences that they had read once 18 months ago. Therefore, if group 1 measures from t_2 are intermediate values between those from t_1 and t_3 for group 2, it will be more parsimonious to link the observed changes in finger kinematics to their progression in reading abilities in general, rather than to a memory trace of these short texts.

Reading assessment
Reading task with digit-tracking Each child read 20 sentences on the tablet laying on the table in front of them. Touching the screen with the finger caused an unblurring space inside which the text was displayed in clear. For each child, the task was composed of two sets of 10 French sentences: a set where the size of the unblurring window was large and another set where it was small. The large window displayed 11 to 12 characters while the small one 4–5 characters. The order of presentation for each set was counterbalanced across children (one group of children saw the first text in the large unblurring window while the other in the small). Prior to each trial, children were submitted to a training phase where they unblur and read aloud two sentences visually presented similar to the ones used in the main experiment.

The structure of the sentences was fixed as follow (French words in italic): Determinant (*le*, *la*, etc.) + noun + "is" (*est*) + adjective. For example, "*La feuille est jolie*" (Meaning: The leaf is pretty). Adjectives and nouns were paired on number of letters, number of syllables and the complexity of graphemic or syllabic structure (see Supplementary Information for the complete material used in the task). The children read sentences made up of words from the lexicon that were grammatically correct but had no meaning. We instructed children not to pay any attention to the meaning of the words.

Sentences were presented in Consolas font, size 11, a fixed-width font. Hence, a given physical length in standard measure contained the same number of letters, regardless of their identity. The width of each character corresponds to 12 pixels (px), which equals to 1.66 mm (mm).

For stimuli presentation and finger trajectory data collection, we used an algorithm by the Psychtoolbox from Matlab Software (version 8.1). Stimuli were presented at a viewing distance of 30 cm on a Dell latitude tablet (1920 × 1280 px) with a screen size of 26.56 × 16.60 centimeter (cm). The sampling rate was 60 Hz (Hz). The original text was blurred using a Gaussian blur filter (standard deviation of 20 pixels). The size of the large unblurring window was 80 × 30 px, and the small unblurring window had a size of 35 × 30 px. The distance between finger touch position and the center of the unblurring window was 80 px (in Y).

Reading tests on paper *Letter and syllable reading (at t_0, t_1, t_2, t_3)* we assessed letter and syllable reading using the French standardized test for reading and orthography evaluation "Batterie d'évaluation de l'écriture et de l'orthographe"[58]. Children had to read aloud different orthographic elements until 5 incorrect answers were given in a row. The total score for each child was the total number of elements correctly read, with a maximum of 85 points (Table 3).

Pseudowords reading (at t_2 or t_3): PW reading was performed using the first 45 of the 90 pseudowords used by Bosse et al., (2007) Children were presented with a sheet with the pseudoword aligned in the column and asked to read the word aloud as quickly and as accurately as possible. Reading time and errors were collected. Reading speed (pseudowords read per minute) and accuracy (pseudowords correctly read/pseudowords read) were calculated (Table 3).

Text reading (at t_2 or t_3) Evaluation of meaningless text was assessed using "Alouette R" test[59] (Table 3). This 265-word text is widely used to assess reading skills in French language. Children are told to read as quickly and accurately as possible the text, within 3 min time limit. Fluency was assessed as the number of words correctly read within 3 min, and accuracy rate as the number of words correctly read among the total number of words read. This test was used at each pre and post-training test, as well as the reading evaluation at the end of the school year (Table 3).

Evaluation of meaningful text reading (at t_2 or t_3) was performed using the French standardized reading test "Monsieur Petit" test[60]. Children were presented with a text and asked to read it as quickly and accurately as possible during 1 min. The number of words correctly read per minute (fluency) and errors were counted. This test was used for the evaluation at the end of the school year (Table 3).

General cognitive abilities

Fluid intelligence (at t_0 or t_1) The age-standardized Matrix task from the WPPSI-IV[61] was used to assess fluid intelligence. The child had to choose a missing drawing amongst several to complete each matrix of drawings in a logical manner. Prior testing, example items were presented to the child with feedbacks. 26 matrices were presented in total. Those who were between 5 and 6 years old begun at the fourth item, and those more than 6 years old at the seventh item (previous items were considered as correct). The test is stopped when the child gave 3 incorrect answers in a row. The child earned 1 point for each correct response. The maximum raw score was 26, which was standardized following the test guidelines (maximum standard score = 20) (Table 3).

Selective attention (at t_0 or t_1) Selective attention was assessed using the child-friendly version of the Flanker fish test[62] based on the basic Eriksen flanker task[63]. Stimuli were presented using Inquisit 5 software. The child saw five fishes on the screen, and had to press the keys to indicate the orientation of the fish at the center (left or right). In congruent trials, the orientation of all five fishes was the same, whilst in incongruent trials, the orientation of the fish at the center was opposite to that of the other four fishes. Beforehand, to familiarize with the task, children had two training blocks with 20 trials each. The time limit for each trial was 3000 ms. After the training blocks, children completed three blocks of 40 trials. Individual measures of selective attention were obtained by subtracting mean RT for valid congruent trials was to mean RT of incongruent trials (Table 3).

Working memory (at t_0, t_1 and t_2) Children performed the Digit span task from the WISC-IV battery[64] in which the experimenter read out loud a string of numbers to the children. They then had to repeat the sequence in a direct or reverse order (Table 3). Children began with strings of two digits, if they repeated correctly two strings in a row, the next two trials would contain one additional number for each string. If children missed two trials in a row, the test is interrupted. The raw score was computed for both orders with a maximum of 16 points for each and converted to age-standardized scores. Since the standard score of this test is only available for children older than 6 years-old, and some of the children in the group 2 did not reach this age at t_0, we assessed this task again at t_2.

Vocabulary (at t_0 and t_1) Vocabulary skills were assessed using the Vocabulary subtest from the WPPSI-IV[61]. The experimenter asked children to describe different words, from concrete to abstract ones (e.g. "Can you tell me what is a doctor?"). Oral responses were noted and each answer was evaluated (from 0 to 2 points) using criteria defined in the Manual Scoring of the battery. The maximum raw score goes up to 43 and was converted to an age-standardized score (maximum standard score = 20) (Table 3).

Rapid Automatized Naming (at t_0 and t_1) : Rapid Automatized Naming was assessed using the battery ODEDYS2[65]. Children were presented two sheets (A4 format) of 25 items, the first one with letters and the second with images. Items were displayed in a 5 × 5 array: in each trial, 5 items (letter or image) were repeated 5 times over each line in a quasi-random order. Children had to name sequentially the items as quickly as possible. Before the assessment, the experimenter showed the children the 5 items to ensure they can name them correctly and if not, instructed them how to name them. The experimenter noted the duration of the task and the number of errors. Naming fluency was calculated for each trial by dividing the number of correctly named items to the duration of the task in minutes (Table 3).

Visual attention span (at t_0 and t_1) Visual attention span was assessed using the Global report test[66], on the same tablet used for digit-tracking reading tasks described above. At each trial, children first saw a fixation cross indicating the onset of the trial. Then, 4 consonant letters appeared briefly during 200 ms. The letters were displayed in Geneva police, size 24, with a distance of 1 cm between letters. Children were then asked to report the letters they just saw. Displayed letters were chosen carefully to avoid forming graphemes (e.g. 'gr', 'ph'). The visual attention span score was computed as the total number of letters reported across 20 trials (max = 80) (Table 3).

Preprocessing of finger kinematics data

Data were pre-processed and processed using homemade Python 3.7 and R (version 3.6.2) scripts[67,68]. Finger touch coordinates were collected at 60 Hz sampling rate (every 16.7 ms), as percentage of screen (width for X-coordinates and height for Y-coordinates), then converted into coordinates pixels. For each child and trial, we computed an individual median speed (after excluding all null values). This individual median speed threshold was used as a cut-off between **digital fixations** (i.e. when the finger moves slower than the median speed of the individual) and **digital saccades** (i.e. when the finger moves faster than the median speed of the individual). The median speed was chosen as threshold value for empirical reason: it is the threshold value that maximizes the number of fixations detected across a large range of speed thresholds. The *digital fixation duration* is computed as the total duration before the next digital saccade (in milliseconds); the digital *position of fixation* (X, Y) is computed as the barycenter of all the position belonging to this digital fixation. If two fixations with a distance lower than 5.8 pixels in X and 6.4 pixels in Y, they were then merged. The *number of digital fixations* is computed as the number of contiguous points streak where the finger moves slower than the median speed during the trial or over the stimulus in the region of interest.

The *digital saccade length* is the distance between two consecutive fixations, computed as the differences in millimeters between coordinates of the fixation at the end and at the beginning of the saccade. A digital saccade was defined as progressive if the X-coordinate of the last point is greater than the X coordinate of the first point and regressive otherwise. As children made long return-sweeps (regressive digital saccades made while jumping to the next line), the average saccade length might be overestimated. To overcome this problem, we excluded all digital saccades whose projected length in Y-axis was superior to 50 pixels or 6.77 mm - the distance between the bottom of the line N and the top of the line $N+1$.

Statistical analysis
Data were analyzed using R software (version 3.6.2, R Development Core Team, 2018) and the *lmer* function from the *lme4* package (version 1.1.-10)[69]. For each finger kinematic variable, we look at the effect of timepoint (3 levels: 1st grade-beginning vs. 1st grade-end vs. 2nd grade-end), of unblurring window size (small vs. large) and of decoding skills at the moment of testing (good vs. poor decoder). Similar analysis in which we use other reading scores (Pseudowords, meaningless text and meaningful text reading) assessed at t_2 or t_3 to determine good versus poor readers are also reported in Supplementary Information. Linear Mixed Models (LMM) were preferred over repeated-measures ANOVA as it can handle missing data, violations of distributional assumptions[70], heterogeneous sample sizes across or within groups and interindividual variations[56,57]. For clarity and readability, we reported main effects and interactions of LMMs using type III F-test output with Satterthwaite's method of estimating degrees of freedom. Graphics are built using ggplot2 package[71]. We established a significance level of $p<0.05$ for all statistical analyses, except when Bonferroni corrections were applied to post-hoc comparisons to control for Type I errors.

Data availability
The datasets used and/or analysed during the current study available from the corresponding author on reasonable request.

Received: 27 December 2023; Accepted: 8 October 2024
Published online: 24 October 2024

References
1. Rayner, K. The Perceptual Span and Peripheral cues in Reading. *Cogn. Psychol.* **7**, 65–81 (1975).
2. Schotter, E. R., Angele, B. & Rayner, K. Parafoveal processing in reading. *Atten. Percept. Psychophys*. **74**, 5–35 (2012).
3. Henderson, J. M. & Ferreira, F. Eye movement control during reading: fixation measures reflect foveal but not parafoveal processing difficulty. *Can. J. Exp. Psychol.* **47**, 201–221 (1993).
4. White, S. J., Rayner, K. & Liversedge, S. P. The influence of parafoveal word length and contextual constraint on fixation durations and word skipping in reading. *Psychonomic Bulletin & Review* 12, 466–471 (2005). (2005).
5. Choi, W., Lowder, M. W., Ferreira, F. & Henderson, J. M. Individual differences in the perceptual span during reading: evidence from the moving window technique. *Atten. Percept. Psychophys*. **77**, 2463–2475 (2015).
6. Veldre, A. & Andrews, S. Parafoveal preview benefit is modulated by the precision of skilled readers' lexical representations. *J. Exp. Psychol. Hum. Percept. Perform.* **41**, 219–232 (2015).
7. Veldre, A. & Andrews, S. Lexical quality and eye movements: individual differences in the perceptual span of skilled adult readers. *Q. J. Exp. Psychol. (Hove)*. **67**, 703–727 (2014).
8. Chace, K. H., Rayner, K. & Well, A. D. Eye movements and phonological parafoveal preview: effects of reading skill. *Can. J. Exp. Psychol.* **59**, 209–217 (2005).
9. Sperlich, A., Schad, D. J. & Laubrock, J. When preview information starts to matter: development of the perceptual span in German beginning readers. *J. Cogn. Psychol.* **27**, 511–530 (2015).
10. Sperlich, A., Meixner, J. & Laubrock, J. Development of the perceptual span in reading: a longitudinal study. *J. Exp. Child. Psychol.* **146**, 181–201 (2016).
11. Norgett, Y. & Siderov, J. Foveal crowding differs in children and adults. *J. Vis.* **14**, 1–10 (2014).
12. Bondarko, V. M. & Semenov, L. A. Visual acuity and the crowding effect in 8- to 17-year-old schoolchildren. *Hum. Physiol.* **31**, 532–538 (2005).
13. Bouma, H. Interaction effects in parafoveal letter recognition. *Nature*. **227**, 520–521 (1970).
14. Whitney, D. & Levi, D. M. Visual crowding: a fundamental limit on conscious perception and object recognition. *Trends Cogn. Sci.* **15**, 160–168 (2011).
15. Manassi, M., Sayim, B. & Herzog, M. H. When crowding of crowding leads to uncrowding. *J. Vis.* **13**, 1–10 (2013).
16. Anderson, E. J., Dakin, S. C., Schwarzkopf, D. S., Rees, G. & Greenwood, J. A. The neural correlates of crowding-induced changes in appearance. *Curr. Biol.* **22**, 1199–1206 (2012).
17. Pelli, D. G. et al. Crowding and eccentricity determine reading rate. *J. Vis.* **7**, 20–20 (2007).
18. Legge, G. E., Mansfield, J. S. & Chung, S. T. L. Psychophysics of reading XX. Linking letter recognition to reading speed in central and peripheral vision. *Vis. Res.* **41**, 725–743 (2001).
19. Bertoni, S., Franceschini, S., Ronconi, L., Gori, S. & Facoetti, A. Is excessive visual crowding causally linked to developmental dyslexia? *Neuropsychologia*. **130**, 107–117 (2019).
20. Martelli, M., Di Filippo, G., Spinelli, D. & Zoccolotti, P. Crowding, reading, and developmental dyslexia. *J. Vis.* **9**, 14–14 (2009).
21. Bouma, H. & Legein, C. P. Foveal and parafoveal recognition of letters and words by dyslexics and by average readers. *Neuropsychologia*. **15**, 69–80 (1977).
22. Chakravarthi, R., Rubruck, J., Kipling, N. & Clarke, A. D. F. characterizing the in-out asymmetry in visual crowding. *J. Vis.* **21**, 1–14 (2021).
23. Shechter, A. & Yashar, A. Mixture model investigation of the inner–outer asymmetry in visual crowding reveals a heavier weight towards the visual periphery. *Scientific Reports 2021 11:1* 11, 1–12 (2021).
24. Moll, K. & Jones, M. Naming fluency in dyslexic and nondyslexic readers: Differential effects of Visual Crowding in Foveal, Parafoveal, and Peripheral Vision. *Q. J. Experimental Psychol.* **66**, 2085–2091 (2013).
25. Blythe, H. I. Developmental changes in Eye Movements and Visual Information Encoding Associated with Learning to Read. *Curr. Dir. Psychol. Sci.* **23**, 201–207 (2014).
26. Blythe, H. I. & Joseph, H. S. S. L. Children's Eye Movements during Reading. *Oxf. Handb. Eye Movements*. https://doi.org/10.1093/oxfordhb/9780199539789.013.0036 (2012).
27. Rayner, K. Eye movements and attention in reading, scene perception, and visual search. *Q. J. Experimental Psychol.* **62**, 1457–1506 (2009).
28. Hawelka, S., Gagl, B. & Wimmer, H. A dual-route perspective on eye movements of dyslexic readers. *Cognition* **115**, 367–379 (2010).
29. Hawelka, S. & Wimmer, H. Impaired visual processing of multi-element arrays is associated with increased number of eye movements in dyslexic reading. *Vis. Res.* **45**, 855–863 (2005).

30. Barton, J. J. S., Hanif, H. M., Eklinder Björnström, L. & Hills, C. The word-length effect in reading: a review. *Cogn. Neuropsychol.* **31**, 378–412 (2014).
31. Rayner, K. Visual attention in reading: Eye movements reflect cognitive processes. *Mem. Cognit*. **5**, 443–448 (1977).
32. Henderson, J. M. & Ferreira, F. Effects of foveal processing difficulty on the perceptual span in reading: implications for attention and eye movement control. *J. Exp. Psychol. Learn. Mem. Cogn.* **16**, 417–429 (1990).
33. Bricolo, E., Salvi, C., Martelli, M., Arduino, L. S. & Daini, R. The effects of crowding on eye movement patterns in reading. *Acta Psychol. (Amst)*. **160**, 23–34 (2015).
34. Rayner, K. & Bertera, J. H. Reading without a fovea. *Sci. (1979)*. **206**, 468–469 (1979).
35. Johnson, R. L. & Dunne, M. D. Parafoveal processing of transposed-letter words and nonwords: evidence against parafoveal lexical activation. *J. Exp. Psychol. Hum. Percept. Perform.* **38**, 191–212 (2012).
36. Pollatsek, A., Lesch, M., Morris, R. K. & Rayner, K. Phonological codes are used in Integrating Information Across Saccades in Word Identification and Reading. *J. Exp. Psychol. Hum. Percept. Perform.* **18**, 148–162 (1992).
37. Miellet, S. & Sparrow, L. Phonological codes are assembled before word fixation: evidence from boundary paradigm in sentence reading. *Brain Lang.* **90**, 299–310 (2004).
38. Ashby, J., Treiman, R., Kessler, B. & Rayner, K. Vowel processing during silent reading: evidence from eye movements. *J. Exp. Psychol. Learn. Mem. Cogn.* **32**, 416–424 (2006).
39. Inhoff, A. W. Parafoveal Processing of Words and Saccade Computation during Eye fixations in Reading. *J. Exp. Psychol. Hum. Percept. Perform.* **15**, 544–555 (1989).
40. Vitu, F., O'Regan, J. K. & Mittau, M. Optimal landing position in reading isolated words and continuous text. *Percept. Psychophys.* **47**, 583–600 (1990).
41. Juhasz, B. J., White, S. J., Liversedge, S. & Rayner, K. Eye movements and the use of parafoveal word length information in reading. *J. Exp. Psychol. Hum. Percept. Perform.* **34**, 1560–1579 (2008).
42. O'Regan, J. K., Lévy-Schoen, A., Pynte, J. & Brugaillère, B. Convenient fixation location within isolated words of different length and structure. *J. Exp. Psychol. Hum. Percept. Perform.* **10**, 250–257 (1984).
43. Lio, G., Fadda, R., Doneddu, G., Duhamel, J. R. & Sirigu, A. Digit-tracking as a new tactile interface for visual perception analysis. *Nat. Commun.* **10**, 1–13 (2019).
44. Duranovic, M., Senka, S. & Babic-Gavric, B. Influence of increased letter spacing and font type on the reading ability of dyslexic children. *Ann. Dyslexia*. https://doi.org/10.1007/s11881-018-0164-z (2018).
45. Wallis, S., Yang, Y. & Anderson, S. J. Word Mode: a crowding-free reading protocol for individuals with macular disease. *Sci. Rep.* **8**, 1–10 (2018).
46. Zorzi, M. et al. Extra-large letter spacing improves reading in dyslexia. *Proc. Natl. Acad. Sci.* https://doi.org/10.1073/pnas.1205566109 (2012).
47. Nguyen, V. C. L., Lio, G., Perret, T., Gomez, A. & Sirigu, A. A digital interface to speed up learning to read in 1st-grade and to collect finger movement related to reading skills. https://psyarxiv.com/3nqhu/ doi: (2022). https://doi.org/10.31234/OSF.IO/3NQHU
48. Huckauf, A. & Nazir, T. A. How odgcrnwi becomes crowding: stimulus-specific learning reduces crowding. *J. Vis.* **7**, 1–12 (2007).
49. Kliegl, R., Hohenstein, S., Yan, M. & McDonald, S. A. How preview space/time translates into preview cost/benefit for fixation durations during reading. *Q. J. Experimental Psychol.* **66**, 581–600 (2013).
50. Hutzler, F., Schuster, S., Marx, C. & Hawelka, S. An investigation of parafoveal masks with the incremental boundary paradigm. *PLoS One*. **14**, 1–26 (2019).
51. Marx, C., Hawelka, S., Schuster, S. & Hutzler, F. An incremental boundary study on parafoveal preprocessing in children reading aloud: parafoveal masks overestimate the preview benefit. *J. Cogn. Psychol.* **27**, 549–561 (2015).
52. Vasilev, M. R., Slattery, T. J., Kirkby, J. A. & Angele, B. What are the costs of degraded parafoveal previews during silent reading? *J. Exp. Psychol. Learn. Mem. Cogn.* **44**, 371–386 (2018).
53. Pelli, D. G., Palomares, M. & Majaj, N. J. Crowding is unlike ordinary masking: distinguishing feature integration from detection. *J. Vis.* **4**, 1136–1169 (2004).
54. Bellocchi, S., Massendari, D., Grainger, J. & Ducrot, S. Effects of inter-character spacing on saccade programming in beginning readers and dyslexics. *Child Neuropsychol.* **25**, 482–506 (2018).
55. Harrison, W. J., Mattingley, J. B. & Remington, R. W. Eye movement targets are released from visual crowding. *J. Neurosci.* **33**, 2927–2933 (2013).
56. Magezi, D. A. Linear mixed-effects models for within-participant psychology experiments: an introductory tutorial and free, graphical user interface (LMMgui). *Front. Psychol.* **6**, 2 (2015).
57. Boisgontier, M. P. & Cheval, B. The anova to mixed model transition. *Neurosci. Biobehav Rev.* **68**, 1004–1005 (2016).
58. Pech-Georgel, C. BELO-Batterie d'évaluation de lecture et d'orthographie. (2006).
59. Lefavrais Alouette-R. Preprint at (2005).
60. Juton, A. & Lequette, C. Evaluation de la fluence en lecture, du CP au lycée. http://www.cognisciences.com/IMG/pdf/evaluation_de_la_fluence_en_lecture2017-2.pdf
61. Wechsler, D. Wechsler Preschool and Primary Scale of Intelligence | Fourth Edition (WPPSI-IV). (2012).
62. Christ, S. E., Kester, L. E., Bodner, K. E. & Miles, J. H. Evidence for selective inhibitory impairment in individuals with Autism Spectrum Disorder. *Neuropsychology*. **25**, 690–701 (2011).
63. Eriksen, B. A. & Eriksen, C. W. Effects of noise letters upon the identificaion of a target letter in nonsearch task. *Percept. Psychophys.* **16**, 143–149 (1974).
64. Wechsler, D. Wechsler Intelligence Scale for Children | Fourth Edition. (2003).
65. Jacquier-Roux, M., Valdois, S., Zorman, M., Lequette, C. & Pouget, G. Outil de DÉpistage des DYSlexies. 1–71 Preprint at (2005).
66. Bosse, M. L., Tainturier, M. J. & Valdois, S. Developmental dyslexia: the visual attention span deficit hypothesis. *Cognition*. **104**, 198–230 (2007).
67. Van Rossum, G. & Drake, F. L. Jr Python 3 Reference Manual. Preprint at (2009).
68. R Development Core Team. A Language and Environment for Statistical Computing. *R Foundation for Statistical Computing* vol. 2 (2018). https://www.R-project.org Preprint at.
69. Bates, D., Mächler, M., Bolker, B. M. & Walker, S. C. Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* **67**, 1–48 (2015).
70. Schielzeth, H. et al. Robustness of linear mixed-effects models to violations of distributional assumptions. *Methods Ecol. Evol.* **11**, 1141–1152 (2020).
71. Wickham, H. *Ggplot2* (Springer International Publishing, 2016). https://doi.org/10.1007/978-3-319-24277-4

Acknowledgements

We would like to thank the children and parents for their participation in this research and the teachers for their help throughout the study.

Author contributions

AS, AG and VCLN designed the study; VCLN tested subjects; TP and VF provided the software. VCLN, AG, AS wrote the manuscript.

Declarations

Competing interests
The CNRS and the University of Lyon hold a patent on the method employed in this study. AS is affiliated to CNRS and AG to the University of Lyon. All other Authors do not declare any competing interests.

Additional information
Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-024-75706-5.

Correspondence and requests for materials should be addressed to A.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.