



HAL
open science

Accelerating reading acquisition and boosting comprehension with a cognitive science-based tablet training

Cassandra Potier Watkins, Julien Caporal, Clément Merville, Sid Kouider, Stanislas
Dehaene

► **To cite this version:**

Cassandra Potier Watkins, Julien Caporal, Clément Merville, Sid Kouider, Stanislas Dehaene. Accelerating reading acquisition and boosting comprehension with a cognitive science-based tablet training. *Journal of Computers in Education*, 2020, 7 (2), pp.183-212. <10.1007/s40692-019-00152-6>. <hal-03015914>

HAL Id: hal-03015914

<https://hal.science/hal-03015914v1>

Submitted on 20 Nov 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

Journal of Computers in Education

Accelerating Reading Acquisition and Boosting Comprehension with a Cognitive Science Based Tablet Training

--Manuscript Draft--

Manuscript Number:	JLCE-D-19-00001	
Full Title:	Accelerating Reading Acquisition and Boosting Comprehension with a Cognitive Science Based Tablet Training	
Article Type:	Original Paper	
Funding Information:	Bpifrance (N° P3182-510089)	Not applicable
Abstract:	<p>Tablets and computers offer opportunities for learning, but their potential is only as great as the quality of the software they propose. Educational games must not only provide an engaging design, but also be based on principles from cognitive neuroscience and education research, and be evaluated in large-scale classroom tests. Here, we describe ELAN, an adaptive game that supports literacy acquisition through teaching and training phonics. It provides explicit systematic grapheme-phoneme correspondence instruction and reinforces full decoding through reading and spelling practice with 100% decodable text. The game also uses periodical lexical decision tasks to measure the transition from letter-by-letter decoding to fluent word recognition. The software was tested in a randomized control trial in 44 first grade classrooms (n=975 French children). Children who used ELAN software during the first term improved relative to two control groups, respectively using math software or no-tablet "business-as-usual" classrooms. Improvements were significant in reading fluency (one-minute word and pseudo-word reading) and sentence reading comprehension, consistent with the idea that improved decoding can help the child focus on understanding. These results emphasize the importance of early, explicit and systematic phonics training, and provide a new software tool to facilitate it.</p>	
Corresponding Author:	Cassandra Potier Watkins INSERM FRANCE	
Corresponding Author Secondary Information:		
Corresponding Author's Institution:	INSERM	
Corresponding Author's Secondary Institution:		
First Author:	Cassandra Potier Watkins	
First Author Secondary Information:		
Order of Authors:	Cassandra Potier Watkins	
	Julien Caporal	
	Clément Merville	
	Sid Kouider	
	Stanislas Dehaene	
Order of Authors Secondary Information:		
Author Comments:	<p>I am pleased to submit to you an original research article entitled Accelerating Reading Acquisition and Boosting Comprehension with a Cognitive Science Based Tablet Training for consideration.</p> <p>This article describes ELAN, a tablet reading program developed in collaboration between experts in serious-games (co-authors Julien Caporal and Clément Merville) and leading researchers in cognitive neuroscience (co-authors Sid Kouider and my thesis advisor, the principal investigator in this project, Stanislas Dehaene). For the last</p>	

ten years, Prof. Dehaene has led breakthrough research on how the brain learns to read and do math (Amalric & Dehaene, 2016; Dehaene et al., 2010) and has written several books on the subject (Dehaene, 2010b, 2010a). He has previously been involved in developing and testing math games (Wilson et al., 2009; Wilson et al., 2006). This article describes the development and testing of ELAN, a systematic explicit phonics application. The method was tested with 975 children. Results include:

- First-graders that used the software improved in 1 minute word and pseudo-word reading and short sentence comprehension.
- These benefits were made in comparison to a math application group and no-tablet control group.
- First-graders that benefitted from ELAN only did so when using the application early in learning. The tablet method, which focuses on phonics, did not help children as a review method.
- The students that benefitted from ELAN early in the year, continued to show better scores in reading at the end of the year — after four months of no longer using the program.

Our project combines understanding of how our brains learn and application to benefit school children by taking advantage of scenarios that tablets/computers can easily provide (e.g. adaptive training, immediate error-feedback and gamified drill and practice). More importantly, we used randomized control field testing to test our theories. We feel that the Journal of Computers in Education would be a good fit and help us to reach readers looking for innovative ways that technology that can support evidence based education practices.

Thank you for your consideration.

Sincerely,

Cassandra Potier Watkins
PhD Candidate, Frontières du Vivant
Cognitive Neuroimaging Unit, Neurospin

Accelerating Reading Acquisition and Boosting Comprehension with a Cognitive Science Based Tablet Training

Authors: Cassandra Potier Watkins^{a*}, Julien Caporal^c, Clément Merville^c, Sid Kouider^d, Stanislas Dehaene^{ab}

Affiliations :

^a INSERM, UMR992, CEA, NeuroSpin Center, University Paris Saclay, Gif-sur-Yvette, France

^b Collège de France, 11 place Marcelin Berthelot, 75231 Paris Cedex 05, France

^c Manzalab, 108 Rue Saint Honoré, 75001 Paris, France

^d Brain and Consciousness Group, Ecole normale supérieure, PSL University, EHESS, CNRS, Paris, France

*Corresponding author, cassandra.potier-watkins@cea.fr, +33(0)6.63.90.18.11, ORCID: 0000-0002-6588-0614

Abstract :

Tablets and computers offer opportunities for learning, but their potential is only as great as the quality of the software they propose. Educational games must not only provide an engaging design, but also be based on principles from cognitive neuroscience and education research, and be evaluated in large-scale classroom tests. Here, we describe ELAN, an adaptive game that supports literacy acquisition through teaching and training phonics. It provides explicit systematic grapheme-phoneme correspondence instruction and reinforces full decoding through reading and spelling practice with 100% decodable text. The game also uses periodical lexical decision tasks to measure the transition from letter-by-letter decoding to fluent word recognition. The software was tested in a randomized control trial in 44 first grade classrooms (n=975 French children). Children who used ELAN software during the first term improved relative to two control groups, respectively using math software or no-tablet “business-as-usual” classrooms. Improvements were significant in reading fluency (one-minute word and pseudo-word reading) and sentence reading comprehension, consistent with the idea that improved decoding can help the child focus on understanding. These results emphasize the importance of early, explicit and systematic phonics training, and provide a new software tool to facilitate it.

Keywords : reading; education; first-grade; tablets computers; intervention; cognitive science

Acknowledgements: Funding for this work was provided by Bpifrance, the public French bank under the mandate of the Government under the Programme Investissement d'Avenir (investing for the future).

Thanks to the many other people who contributed to this project: Lilliane Sprenger-Charolles, Gaëlle Olivier, Thibault Petit, Stéphanie Philippe, Bruce McCandliss and Isabelle Dénghien.

We are exceptionally grateful to our partner the Academy of Poitiers, in particular, Dominique Quéré (Délégué Académique au Numérique), Cedric Couvrat (Délégué académique au numérique Adjoint pour le 1er degré) and the school district team that tested and accompanied the participating classes. A very special thanks to the teachers that participated in this project and their pupils and families.

Declaration of Interest

We wish to confirm that there are no known conflicts of interest associated with this publication. We also wish to draw the attention of the Editor to the following facts:

- A 'Cooperation Accord' was signed between Manzalab, the labs and the BpiFrance to finance the development and testing of the ELAN software.
- Manzalab's (a company specializing in serious games) principle role was to develop the software ELAN as a serious-game and, if so desired, commercialize the game after testing.
- The principle roles of the two participating labs were to oversee adherence by the game software to principles from cognitive neuroscience and education research, test game ergonomics and test the ELAN software in a randomized control intervention study.
- Neither research labs received payment from Manzalab for their involvement in the project, nor are they privy to any financial gain from the dissemination or licensing of the application ELAN.

Research in this article was strictly conducted between INSERM and the school district of Poitiers, France. Testers from the school were blind to the intervention and progress made by students, while the researchers in the lab did not participate in testing. Only the 'global results' from the student's groups has been shared with Manzalab.

We confirm that we have given due consideration to the protection of intellectual property associated with this work and that there are no impediments to publication, including the timing of publication, with respect to intellectual property. In so doing we confirm that we have followed the regulations of our institutions concerning intellectual property.

1. Introduction

International evaluations in schools have reported that major investments in technology for the classroom do not necessarily lead to improvements in math and reading abilities (OECD, 2015). This is in part due to the lack of quality research-backed software that not only takes advantage of the experiences that technology can provide, but also incorporates scientific findings from cognitive neuroscience and education about how the brain learns. Previous research has shown that games that adhere to evidence backed pedagogy, gamify drill-and-practice, and use adaptive algorithms to provide individualized learning can lead to improvements in math (Räsänen, Salminen, Wilson, Aunio, & Dehaene, 2009; Schacter & Jo, 2016; Wilson, Dehaene, Dubois, & Fayol, 2009; Wilson, Revkin, Cohen, Cohen, & Dehaene, 2006) and reading (de Graaff, Bosman, Hasselman, & Verhoeven, 2009; Kyle, Kujala, Richardson, Lyytinen, & Goswami, 2013; Lyytinen, Ronimus, Alanko, Poikkeus, & Taanila, 2007; Saine, Lerkkanen, Ahonen, Tolvanen, & Lyytinen, 2011). The goal of the current project was to test the ELAN software, a tablet-based game designed to support reading acquisition in the classroom. The game's pedagogy is inspired from evidence in cognitive neuroscience and education research as to how children learn to read. The game focuses on phonics teaching and training, an essential stepping-stone to reading mastery as made evident by the meta-analyses of the National Reading Panel (Cunningham, 2001). ELAN was not designed to supplant the role of the teacher but aims to provide 'drill-and-practice' and tailored game difficulty to accelerate consolidation of the initial decoding stage of reading.

1.1. The reading brain

The goal of learning to read is to train the visual system to decode letter strings and provide a novel input into the already developed cortical areas for the spoken language comprehension, (Dehaene et al., 2010). Advances in functional MRI before and after learning to read have revealed that part of the left occipito-temporal pathway dedicated to visual recognition becomes sensitive to letter strings and develops increasingly efficient connections to regions, such as the planum temporale, specialized in the processing of speech sounds (Dehaene et al., 2010; Dehaene-Lambertz, Monzalvo, & Dehaene, 2018; Monzalvo & Dehaene-Lambertz, 2013). In the early stages of reading, the learner must effortfully convert each letter or group of letters (called a 'grapheme') into the corresponding sound unit (called a 'phoneme') while 'listening in' to understand the word. With practice, reading becomes

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

automatic, and words can be treated in their entirety, with all letters analyzed in parallel, permitting rapid access to the lexicon.

During the early decoding stages of reading, reading time increases linearly with word length, with a slope as high as ~200 ms/letter (Zoccolotti et al., 2005), reflecting a slow process of serial letter-by-letter reading. As reading automatizes, that slope decreases until all words between 3 and 8 letters are read equally fast (New, ferrand, pallier, & brybaert, 2006), except in young dyslexic readers (Zoccolotti et al., 2005). These results fit with a dual-route model of reading: the phonological path allows for a slow decoding of novel words, while the lexical path lends itself to fluent reading of familiar words (Coltheart, 2005). Both neural paths exist in the proficient reader, and the goal of reading instruction should therefore be to quickly establish them. According to meta-analyses by the National Reading Panel (Cunningham, 2001) the most efficient way to acquire literacy is phonics instruction, in both normally developing readers and children with increased risk for reading deficits.

1.2. Principles for phonics instruction

There are several key principles to a successful phonics instruction. Grapheme-phoneme correspondences should be taught explicitly (Castles, Rastle, & Nation, 2018; Cunningham, 2001). The explicit teaching of grapheme-phoneme correspondence should follow a systematic progression (Ehri, Nunes, Stahl, & Willows, 2001), meaning that grapheme-phoneme correspondences are taught in a carefully planned sequence, within a rational order that takes into account their frequency and their consistency. Systematic phonics software has been shown to provide better results than non-systematic software (de Graaff et al., 2009).

Practicing lesson-related words should also be a part of any systematic phonics instruction (Mason, McDaniel, & Callaway, 1974). Children should be presented with decodable text, i.e. using graphemes whose pronunciation has been previously taught. In spite of adequate phonics instruction, many children show difficulties in applying full decoding skills. These children have no trouble identifying the first letter of a word, but fail to apply their phonics knowledge to all successive letters in such a manner that they combine them into readable syllables. Reading activities that draw attention to the combinations of letters in a word help children apply the alphabetic principle to all letters (McCandliss, Beck, Sandak, & Perfetti, 2003). Learning to spell, which requires children to focus on each sound and its corresponding letter, also appears to improve decoding skills (Uhry & Shepherd, 1993).

1.3. Tablet based interventions for reading instruction

1
2 Tablet- or computer-based games may provide an excellent medium for helping
3 children to automatize their grapheme-phoneme decoding skills. A prime cross-language
4 example and an important source of inspiration for the present project is the GraphoGame, a
5 Finnish computer/tablet based game in which the primary goal is to incite the child to
6 automatize the associations between graphemes and phonemes (Richardson & Lyytinen,
7 2014). Functional brain imaging shows that when preschoolers play the GraphoGame for a
8 few hours, the neural circuits for reading begin to emerge (Brem et al., 2010). Playing the
9 GraphoGame helps children considered at risk for reading acquisition in both English (Kyle et
10 al., 2013) and Finnish (Saine et al., 2011) improve their decoding and encoding skills above
11 their respective control groups.
12
13
14
15
16
17
18
19
20

1.4. The ELAN software

21
22 Since the GraphoGame was at the time was not available in French, our goal for the
23 ELAN software was to provide similar phonics training for children in France. We also included
24 in ELAN a teaching section that explicitly introduces new grapheme-phoneme combinations
25 in a systematic order, thus allowing children to advance in the game independently of learning
26 in the classroom. ELAN also uses a lexical decision task to measure changes in reading fluency.
27 We now describe these points in turn.
28
29
30
31
32
33
34

35 In ELAN, the child must explore 20 different islands where he or she will learn 1 to 5
36 different grapheme-phonemes relationships. After learning a new grapheme-phoneme
37 lesson, automatization is reinforced through three mini-games that build on skills of syllable
38 and word reading. Passage from one island to the next is allowed once the child has completed
39 all mini-games with a minimum score of 80% correct responses and completed the island
40 “boss”. Children receive awards with each game won.
41
42
43
44
45
46

47 The systematic ordering of grapheme-phoneme correspondences was determined by
48 examining the frequency and consistency of all grapheme-phoneme correspondences in the
49 Manulex corpus, a corpus of 1.9 million words from French children’s books (Lété, Sprenger-
50 Charolles, & Colé, 2004). To further facilitate learning, we made several principled
51 adjustments based on the progression usually taught in phonics manuals and teachers
52 reports. In ELAN, the most frequent vowels are taught first. Regarding consonants, frequent
53 fricatives are instructed before stop consonants due to the greater facility in pronouncing such
54
55
56
57
58
59
60
61
62
63
64
65

1 phonemes for a prolonged duration in isolation, and therefore in teaching how they blending
 2 into syllables (e.g. 'ffff aaa makes fa'). The appendix presents the full order of grapheme-
 3 phonemes taught in ELAN.
 4

5 Within a given ELAN lesson, a grapheme-phoneme correspondence is introduced in
 6 three steps that are commonly used by teachers and supported by research as benefitting
 7 phoneme processing and grapheme memorization (Figure 1). In step 1, the child clicks the
 8 grapheme to hear and see a high-quality, noise-free video of an older child pronouncing the
 9 corresponding phoneme, with a zoom on the child's mouth and explicit articulation.
 10 Enhancing auditory signal-to-noise, for instance through auditory aids, is known to facilitate
 11 reading acquisition (Tallal, 2004). Pictures of distinct articulatory gestures underlying each
 12 phoneme facilitate the development of early reading (Boyer & Ehri, 2011; Castiglioni-Spalten
 13 & Ehri, 2003). In step 2, when the child clicks on the grapheme, a picture that starts with the
 14 phoneme is displayed and described (e.g. 'aaa, aaapple'). This type of exercise relies on the
 15 acrophonic principle, which is at the origins of alphabetic writing (each letter used to be the
 16 first letter of a common word, e.g. b = "beit" = a house). It is commonly used in the classroom
 17 and used here to bind the phoneme to a known word in order to facilitate memorization and
 18 to build phonemic awareness of letter sounds in words. The image is there only to ensure
 19 understanding of the spoken word. Children can repeat these two steps as much as needed.
 20 In step 3, the child traces the grapheme twice in both upper and lower case on the touch
 21 screen. Haptic exercises with letters, such as learning to write or to trace letters with the finger
 22 are useful activities which have been shown to improve reading skills, probably because they
 23 provide an additional motor code to support memory for grapheme (Bara & Gentaz, 2011;
 24 Bara, Gentaz, & Colé, 2007; Bara, Gentaz, Colé, & Sprenger-Charolles, 2004; Bara, Morin,
 25 Alamargot, & Bosse, 2016; Longcamp, Zerbato-Poudou, & Velay, 2005). Because gestures are
 26 completely different for visually similar letters such as b and d, tracing may also help surmount
 27 the mirror-letter confusions that all children experience, due to the fact that the visual cortex
 28 generalizes across viewpoints (Dehaene et al., 2010).
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52

53 *Fig1 Grapheme-phoneme learning in the ELAN software. The child is introduced to each grapheme-phoneme*
 54 *correspondence by seeing the grapheme and watching a video of another child pronouncing the sound (top images), seeing*
 55 *the grapheme and hearing the sound in a word with accompanying picture (bottom-left image) and writing the letter on the*
 56 *screen (bottom-right image).*
 57
 58
 59
 60
 61
 62
 63
 64
 65

1.4.1. Automatizing grapheme-phoneme association

1
2 Once the child has completed the lesson, the next goal is to overlearn the
3 correspondence to the point that it becomes automatic. ELAN uses several mini-games that
4 require children to make faster and faster responses to recently acquired grapheme-phoneme
5 relationships, similar to the GraphoGame (see Figure 2.a). In each game, children hear a
6 syllable and must click on the corresponding written stimulus on screen. Syllables can be a
7 single vowel sound, consonant-vowel (CV) or vowel-consonant (VC). Difficulty is created by
8 varying the distractors and the rate of targets and distractors. Distractors can be syllables with
9 the letters reversed, the vowel replaced or the consonant replaced (for example for the target
10 syllable 'la', distractors would include 'al', 'le' and 'ta').
11
12
13
14
15
16
17
18

19 Once children have learned a grapheme-phoneme pair, they immediately apply this
20 knowledge to reading and spelling small words (Figure 2.b). Finally, the third type of games in
21 ELAN asks children to read short sentences and show their understanding through an action,
22 for example, choosing a picture that corresponds to the written text, or showing the game
23 character the silent letters in a sentence (Figure 2.c). The algorithm for automatic word
24 selection in ELAN ensures that all words that are presented for spelling and reading are 100%
25 decodable.
26
27
28
29
30
31
32

33 *Fig2 Screenshots of various games. A, Letter and syllable games. After learning a given grapheme-phoneme correspondence,*
34 *children practiced automatizing it by playing various games inspired by the GraphoGame, in which they heard the grapheme*
35 *within syllables and had to find them written in upper and lower-case letters on screen. The number and type of distractors*
36 *and their rate of their appearance were automatically adapted to each child's performance. B, word building games required*
37 *children to hear a word and then spell it. Difficulty was adjusted to each child by varying word length and the number and*
38 *type of distractor letters. C, Decoding and comprehension games required children to read written phrases and sentences with*
39 *100% decodable text and demonstrate their understanding through an action (e.g. deciding whether the phrase was in the*
40 *singular or plural).*
41

42 1.4.2. Adapting to the child's level

43
44 ELAN's algorithm provides several types of adaptation to each child's level and errors.
45 The local support, common to adaptive gaming, addresses children's punctual errors: if the
46 child makes mistakes, an error bell and loss of a life-point ensue. If the child makes a second
47 error, the game character appears to remind the child of the game objective and highlights
48 the correct response. If the child responds correctly, the highlight is removed for the next
49 question, otherwise the game ends and another game proposes the same lesson through a
50 different strategy.
51
52
53
54
55
56

57 Game difficulty is also adaptive. As the child makes more correct responses, the game
58 requires quicker responses. For each game, five levels are set up. These five levels of difficulty
59
60
61
62
63
64
65

1 control, according to the mini game environment, the number of distractors present, the
2 number of locations the child must monitor in order to find the target, and the time the target
3 is visible. Two successive mini-game wins or losses cause the level of difficulty to change.
4 Requiring faster and faster responses in environments of growing difficulty would be difficult
5 to reproduce with traditional pen and paper tests and is intended to promote automaticity.
6
7

8
9 Finally, the choice of targets and distractors also adapts to each child. The child is
10 required to constantly maintain a performance level superior to 75% correct for all the
11 grapheme-phoneme pairs learned. If across several games, the child makes several mistakes
12 on the same grapheme-phoneme, its score may decrease below this criterion, causing it to
13 become reintroduced as a target. The knowledge score of each grapheme-phoneme pair is
14 based on the last 5 responses to the pair. At the onset of each mini-game, the game software
15 chooses appropriate targets for the child. This pool of syllable or word items is chosen from
16 the current lesson (80% of possible target items), previously learned items (10%) and items
17 containing a grapheme-phoneme correspondence that tipped below the 75% correct
18 response threshold (10%). If all items are above threshold, then random previously learned
19 items are presented. Using this technique allows for repeated spaced practice, which is known
20 to benefit long-term retention (for reviews, see (Carpenter, Cepeda, Rohrer, Kang, & Pashler,
21 2012; Cepeda, Pashler, Vul, Wixted, & Rohrer, 2006).
22
23
24
25
26
27
28
29
30
31
32
33
34

35 Distractor stimuli are also carefully chosen according to the target and the child's
36 performance history. While only previously learned grapheme-phoneme correspondences
37 can act as distractors (nothing is presented to the child that has not been taught in the game),
38 these are chosen according to their orthographic or phonemic proximity to the target. The
39 visual similarity matrix was adapted from another study of similarity judgements of letter pairs
40 (Boles & Clifford, 1989). The phoneme similarity matrix was adapted from French listeners'
41 confusions in discriminating phonemes in CVC pseudo-words presented in natural background
42 noise (Meyer, Dentel, & Meunier, 2013). Only vowels are presented as vowel distractors and
43 consonants as consonant distractors. Half of distractors were chosen based on the letter
44 shape matrix and half based on the phoneme similarity matrix. Initial items are chosen (when
45 possible) to be $>.75$ in similarity to the target (1 being the matrix diagonal, and 0 being
46 maximally different items). If the player chooses an erroneous distractor, then the
47 corresponding threshold on distractor similarity is lowered.
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1.4.3. Measuring the word length effect

As mentioned in section 1.1, a marker of efficient reading is the disappearance of the word length effect, all written words being ultimately processed equally fast regardless of their length (between ~3 and ~8 letters). The challenge in ELAN is to measure reading affected by word length without requiring the child to actually read aloud (given that high-quality child-oriented speech recognition was unfortunately not available at the time of programming), as is customary. ELAN uses a lexical decision game, presented as a “boss” allowing passage from island to island (see section 1.4), requiring the children to send ‘real words’ to a green buoy and ‘invented words’ to a red buoy to measure the players reading time. In this game, the child is presented with 16 words and 16 pseudo-words (with a single letter change or inversion), 2-5 letters long, in random order. The challenge must be accomplished in under three minutes (indicated to the child by the sun that crosses the sky). This game thus emphasizes fast word recognition and allows us to measure the identification time for words and pseudo-words. Through this game, we aimed to evaluate the reading slope, i.e. the amount of increase in reading time with each additional letter. As described below, this strategy was successful: identification time initially increased linearly with each additional letter, and that slope decreased with automatization.

2. Experimental Study

The ELAN software was designed in several successive stages, each involving piloting with first-grade children as well as adult crash tests. Once the games and their parameters were set up, we tested it in a large test with 1st grade French schools, contrasting it with both a low-level no-tablet control and with an active control (children playing a previously designed number game, the NumberCatcher). We now describe how this experimental study was organized. The project was approved by the ethical committee CERES (comité d’éthique pour la recherche en santé) on June 28, 2016.

2.1. Materials and Methods

2.1.1. Participants

Testing of the ELAN software was done in collaboration with the public school district of Poitiers, a region in central-west France. School district personnel in charge of the project were asked to provide approximately 1000 children from classes representing the region. Teachers were approached by district employees, and 53 teachers agreed to participate, for a

1
2
3
4
5
6
7
8
9
10
total of 53 classrooms from 45 schools and 36 towns (ranging in population from >80,000 to <3,000 inhabitants). Representative of these demographics, 44 classrooms agreed to integrate our tablet intervention with their reading instruction curriculum, while 9 control classes (group labelled “control”) followed a “business-as-usual” curricula for the entire school year. Note that the tablet/no-tablet variable was not randomly assigned, as it was not possible to force teachers to one of these two groups against their will.

11 2.1.2. Procedure

12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
At the start of the school year, teachers sent home with students a letter explaining the project, accompanied by a ‘letter of opposition’ for parents who did not wish for their child to participate. In classrooms participating in the tablet intervention, the project followed a randomized cross-over design (see Figure 3). Students were randomly assigned to groups of 3 to 8 children playing the same game at the same time in the classroom’s designated ‘tablet zone’. The only request made by the lab was that teachers check that groups be heterogenous in ability. These teacher-determined small groups were then randomized by the lab into either a group that played ELAN for the first half of the year, then the control math game for the second half of the year (group 1, labelled “read/math”); or, conversely, the math game first, then ELAN (group 2, labelled “math/read”). At the start of the year, 975 children were announced by teachers as participating in the project. 417 were assigned to read/math, 357 to math/read and 201 to control. Whether a group was to play ELAN or the control math game was only announced to teachers on day one of the game intervention.

39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
All children were pretested from September 20th to October 11th 2016, except for 5 children who were tested in the first week of the intervention. This was followed by two weeks of school vacation, and a week to install the application. The two tablet intervention groups then began the first intervention session as soon as the games were installed. The first period lasted from November 7th 2016 to until the end of the first week of February 2017, for a possible 11 weeks of game play (this excludes the 2 weeks of Christmas break). The first posttest (labelled posttest A) began January 30th and ended February 18th 2017. After posttest A, children had two weeks of winter vacation, then switched games. At this point, schools were sent an updated version of ELAN correcting some bugs. Session B commenced as soon as the game was updated. Session B lasted from March 13th to June 9th 2017. This second period was one week longer (excluding the 2 weeks of spring vacation), allowing teachers to make up for lost time due to four 3-day weekends during this period. The second posttest

1 (labelled posttest B) began June 12th and ended June 30th 2017. Teachers were asked that, as
2 much as possible, children be allowed to make up for lost days. Our goal was for all children
3 to play a minimum of 480 minutes (8 weeks, each with three 20-minute periods).
4

5 The Control game used for the intervention was Number Catcher
6 (www.thenumbercatcher.com). This game, a sequel to a previous game developed by our lab
7 (Number Race: Wilson et al., 2009; Wilson, Dehaene, et al., 2006; Wilson, Revkin, et al., 2006),
8 was created for children between ages 5-10. The goal of Number Catcher is to train basic
9 concepts of number, arithmetic, and base-10 understanding, and to cement the links between
10 symbols and quantities by requiring quicker and quicker responses. No research had been
11 conducted on the benefits of the game for strengthening number understanding at the time
12 of the study.
13
14
15
16
17
18
19
20
21

22 *Fig3 All children took reading and math tests at the beginning, middle and end of the year (respectively termed pretest,*
23 *posttest A and posttest B). In a classic crossover design, during the first half of the school year, children in the Intervention*
24 *groups (Group 1 and Group2) played with either ELAN reading software or the math NumberCatcher software. After*
25 *posttest A, children switched to the other software for the second half of the year. Control classes (Group 3) did not have*
26 *tablets. Children played in small groups using the same game: NumberCatcher (top picture) or ELAN (bottom picture).*
27
28
29
30

31 Teachers were instructed not to change their own teaching method in light of the tablet
32 intervention. Adding the tablet atelier was a challenge to incorporate for many teachers, but
33 reports from the academy found this change to be overall positive, as it gave more time to
34 take care of other children during this period. The games were created for autonomous play
35 and teachers did not interact with the children while they played, unless asked for help by the
36 student. All teachers reported that, while a subgroup of children were playing with the tablet,
37 they continued their usual learning program with the rest of the class or provided other
38 specialized work groups. When playing ELAN or Number Catcher, children worked individually,
39 wearing headphones, but could communicate and help each other since all children in a group
40 played the same game.
41
42
43
44
45
46
47
48

49 Most of the participating classes did not have access to wifi, so a child's progression in
50 the game was stored in the application and only occasionally uploaded to our servers. This
51 required that a child always play on the same tablet, logging in by entering a private code. To
52 share the child's progression data with the lab, teachers were asked to connect to the school
53 wifi and upload the data to the lab server. Periodic reminders were sent to all actors in the
54 project reminding them to upload data.
55
56
57
58
59
60
61
62
63
64
65

2.1.3. Pre and Post Tests

1
2 Except for a vocabulary test that was only administered at pretest, all tests were given
3
4 at all three test periods. Forty employees from the Poitiers school district individually tested
5
6 each child. All tests were done on paper. Twenty-three members of this team had attended a
7
8 test training day organized by the lab, and were in charge of training other testers. These
9
10 school employees were often in contact with teachers but would only be knowledgeable of a
11
12 student's group if they purposefully asked, which they were instructed not to do during
13
14 testing. Given that most children would not yet have had formal reading training at pretest,
15
16 we emphasized testing known predictive measures as well as measures of reading ability.
17
18 Although this was not the primary goal of this study, we also included two number knowledge
19
20 tests to provide a minimal evaluation of the results for our control group.

21 **Passive Vocabulary Test (60 items)**

22
23 We used a standardized French vocabulary test (TVAP, Deltour and D, 1990). A word
24
25 was said by the experimenter. The child was asked to choose the best corresponding image
26
27 from a choice of six pictures. Two points were awarded for the correct response, 1 point for a
28
29 close response, 0 for all other responses, for a total possible score of 60.

30
31 **Phoneme (24 items) and Syllable Suppression (10 items)**

32
33 These two tests were taken from EVALEC, a batterie of French predictive reading
34
35 acquisition measures (Pourcin, Sprenger-Charolles, El Ahmadi, & Colé, 2016). The child was
36
37 asked to suppress the first phoneme/syllable of an pseudo-word pronounced by the
38
39 experimenter. Ten syllable suppression questions, then twelve consonant-vowel-consonant
40
41 and twelve consonant-consonant-vowel pseudo words were said to the child. For each type
42
43 of question there were two trials with feed-back. The child's percentage correct was taken
44
45 separately for syllable suppression and phoneme suppression.

46
47 **Letter Knowledge (52 items)**

48
49 Children were asked to give the name and sound of all 26 letters. The child's
50
51 percentage correct out of 52 items was collected. Letters were presented on paper, one-by-
52
53 one and out of order in lower-case.

54
55 **One minute word and pseudo word reading (2 x 1 minute reading)**

56
57 The child was asked to read a maximum number of words followed by a maximum
58
59 number of pseudo-words, each in one minute (Gentaz, Theurel, & Sprenger-Charolles, 2017).
60
61 Children were presented with 35 words and 30 pseudo-words at pretest and posttest A. 60
62
63
64
65

1 words were presented at posttest B. Words and pseudo-words were all either mono- or bi-
2 syllabic. All the presented words were frequent and taken from the Manulex children’s word
3 data-base (Lété et al., 2004). Pseudo-words were matched to the words in syllable structure
4 and orthographic difficulty . If a child was unable to do these two tests, reading less than four
5 words from the first ten words, then he or she was not asked to continue to the more difficult
6 reading tests that followed. To stabilize the variance, children’s mean combined score for
7 these two tests was analyzed, after checking that they provided consistent results.
8
9

10 **Non-sense text reading**

11 The child was asked to read a text of 265 words, all of which are frequent. The text,
12 however, is non-sense and thus requires that the child avoid reading words by context. The
13 text, known as the Alouette test, has been standardized with French primary school children
14 and is used to evaluate reading age, diagnose dyslexia (Lefavrais, 2005) and predict future
15 reading difficulties (Bertrand, Fluss, Billard, & Ziegler, 2010; Cavalli, Duncan, Elbro, Ahmadi, &
16 Colé, 2017).
17

18 **Sentence comprehension (8 items)**

19 Eight items were adapted from the French standardized ECOSSE test and Reading
20 Evaluation Test (Gentaz et al., 2017; Lecocq, 1996). The child had to read aloud a sentence
21 such as “the man is eating an apple”, then choose the correct corresponding image from a
22 choice of four pictures, of which the distractors made inferred a semantic or syntactic change
23 to the sentence. Items were changed from posttest A to posttest B. The sentences in this test
24 were in order of increasing syntactic difficulty. The percentage correct score was calculated
25 for each child. Items from pre and posttest A were not changed as 57% of children did not
26 attempt this test and the average score was below 1 response correct at pretest.
27

28 **Symbolic Number Comparison (number of CR in 30 seconds)**

29 In the SYMP test of number comparison, children were asked to cross out the
30 numerically larger of two Arabic one- digit numbers in thirty seconds (Brankaer, Ghesquière,
31 & Smedt, 2017). 56 pairs of numbers were presented on a page. Four trials with feedback
32 were provided before beginning. Children did the same test at each of the three tests. The
33 maximum number of correct responses was collected from each child.
34

35 **Single-Digit Subtraction (12 items)**

36 Single-digit subtraction was used instead of addition, because subtraction provided a
37 better measure of a child’s ability to do arithmetic as answers are not typically memorized as
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1
2
3
4
5
6
7
8
9
10
with the addition table. Children saw the subtractions, presented one-by-one in p - q format, and heard them orally. There were 18 questions in total: six -1 subtrahend questions, six small subtrahend questions (-2, -3, -4) and six large subtrahend questions (-6, -7, -8). The test was stopped if the child had less than three correct answers on the first 8 questions. Item order was changed from posttest A to posttest B. The percentage correct score was calculated for each child.

11 12 2.1.4. Participant attrition

13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
The goal of this project was to evaluate the possible benefits in reading acquisition for children using the ELAN software. To measure this with confidence, we had set several requirements for inclusion in the research analysis. The first requirement was that all participants complete all three tests so as to ensure an accurate comparison of progression made, given that all children would undoubtedly progress thanks to normal class training. With each test, we lost a number of children due to absence on the testing day, children leaving schools, or testers forgetting to report id information on the tests. From an original count of 975 participants, 728 children completed all three tests (910, pretest; 867, posttest A and 728 posttest B). The large loss between posttest A and posttest B was due to two classes that left the project and three classes that used incorrect posttest B tests, they had re-used posttest A tests.

35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
We also lost participants due to a software bug that damaged datafiles, causing progression data to be lost and sending many players back to square one in the game. If the child reported to the teacher this anomaly, we were able reinitialize the child's game. Unfortunately, not all children reported losing their progression and this problem was not always detected quickly enough. In order to keep a participant in our analysis, we required a normal game progression, meaning: 1) the child had spent sufficient time using the game (the threshold criteria was set at a minimum of 330 minutes of play) and 2) progression had been unhampered by problems with the software. The number of participants remaining from the intervention groups after these criteria was 350 (113 read/math; 103 math/read).

52
53
54
55
56
57
58
59
60
61
62
63
64
65
To examine the baseline balance between the two intervention and control groups, we compared the three groups, two-by-two, in terms of baseline demographic variables and a composite z-score of the four predictive reading acquisition tasks at pretest (vocabulary, syllable awareness, phoneme awareness and letter knowledge). An ANOVA of this composite score revealed no significant difference between the two intervention groups ($F(1, 214) = .55$,

p = 0.46. A significant difference was however found between the read/math and control groups ($F(1, 245) = 5.70, p = 0.02$), but not between the math/read and control group ($F(1, 235) = 2.47, p = 0.12$). Seven children with the lowest composite predictor scores in the control group were removed to ensure equal groups, see Table 1 for baseline information of the three groups.

Table 1 Baseline characteristics at pretest.

mean (sd)	read/math	math/read	control	group difference
N	113	103	127	
N = Male	61	53	63	$\chi^2(2)=.46, p=0.80$
Age in months at pretest	75.8 (3.5)	75.7 (3.8)	75.6 (3.5)	$F(2, 340)=.287, p=0.75$
Time using ELAN, minutes	528 (114)	508 (139)	-	$F(1, 214)=1.36, p=0.25$
vocabulary, x=60	41.06 (5.25)	39.60 (5.57)	40.50 (5.80)	$F(2, 340)=1.89, p=0.15$
syllable suppression, %	52% (41%)	55% (40%)	48% (38%)	$F(2, 340)=1.14, p=0.32$
phoneme suppression, %	46% (28%)	42% (29%)	39% (30%)	$F(2, 340)=1.80, p=0.17$
letter knowledge, %	73% (18%)	74% (19%)	71% (20%)	$F(2, 340)=.979, p=0.38$
composite predictor	0.08 (0.68)	0.01 (0.73)	-0.07 (0.77)	$F(2, 340)=1.20, p=0.30$

2.1.5. Statistical analysis

Given our 2-period cross-over design, for each task, we first analyzed responses from the children randomized into the read/math and math/read groups in a mixed analysis of variance with factors of intervention group (2 levels: read/math versus math/read) and test period (3 levels: pretest, posttest A, posttest B), corresponding to the randomized part of the experimental design. We then did a larger mixed analysis of variance with 3 levels for the group factors (the 2 intervention groups plus the control group) and the same test period factor. For the tests where reading was required by the child, to factor out pre-schooling differences amongst children, we added a co-variable which was a composite predictor of the pretest scores on the pre-reading knowledge tests (vocabulary, syllable suppression, phoneme suppression and letter knowledge). As shown in Table 2, this covariable was useful since there were significant correlations between the predictive reading acquisition tests and the reading tests.

In a crossover test, we expected to see a 2-step pattern of improvement: children in the read/math group should improve in reading tests at posttest A relative to the math/read

intervention and control groups; reading scores should increase for the math/read group relative to the control group at posttest b. An effect of the training is therefore expressed as a significant interaction between the within-participant factor of test time point and the between-participant factor of group. Given a significant interaction effect, post-hoc tests were done at each test period to untangle changes in progress made by each group.

Table 2 Correlations (r^2) between reading predictors at pretest and the two posttests.

	Predictors at Pretest			
	vocabulary	syllable suppression	phoneme suppression	letter knowledge
Posttest A				
syllable suppression	0.07	0.20	0.14	0.08
phoneme suppression	0.06	0.25	0.37	0.23
letter knowledge	0.04	0.10	0.16	0.22
word and pseudo-word reading	0.04	0.21	0.32	0.23
non-sense text reading	0.05	0.17	0.21	0.14
sentence comprehension	0.04	0.07	0.11	0.05
Posttest B				
syllable suppression	0.08	0.14	0.09	0.10
phoneme suppression	0.07	0.19	0.24	0.17
letter knowledge	0.08	0.10	0.12	0.20
word and pseudo-word reading	0.02	0.16	0.19	0.14
non-sense text reading	0.03	0.11	0.12	0.08
sentence comprehension	0.06	0.04	0.07	0.04
(df = 341) , r^2 scores superior to 0.03 are significant at $p < 0.001$.				

3. Results

Figures 4 and 6 show the evolution of performance across the three test periods for each of the read/math, math/read, and control groups, respectively for the reading and math tests.

3.1. Syllable Suppression

The analysis was done on the percentage of correct responses from the 10 items. The intervention group ANOVA revealed, as expected, a significant main effect of test session ($F(2, 428) = 94.8, p < 0.001$), consistent with an improvement over time. No group x test interaction was found ($F(2, 428) = .35, p = 0.705$). Similar results were found when including the control

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

group in our analysis: there was a significant main effect of test session ($F(2, 680) = 176, p < 0.001$), but no group x test interaction ($F(4, 680) = 1.59, p = 0.176$).

3.2. Phoneme Suppression

The analysis was done on the percentage of correct responses combining both CV and CCV phoneme suppression tests, for 24 questions total. Again, the intervention group ANOVA showed a significant main effect of test session ($F(2, 428) = 230, p < 0.001$). All children progressed equally in the task as was made evident by the absence of a group x test interaction ($F(2, 428) = .415, p = 0.66$). Similar results were found when including the control group in our analysis: there was a significant main effect of test session ($F(2, 680) = 390, p < 0.001$), but no group x test interaction ($F(4, 680) = .333, p = 0.856$).

3.3. Letter Knowledge

The percentage correct for the child's combined letter name and letter sound score was collected in this test. The intervention group ANOVA showed a significant main effect of test session, ($F(2, 428) = 286, p < .0001$), but no interaction between group and test period ($F(2, 428) = 1.15, p = 0.316$). Similar results were found when adding the control group: all three groups made progress over time, as shown by a main effect of test session ($F(4, 680) = 441, p < 0.001$), in absence of an interaction ($F(2, 680) = 1.44, p = 0.22$).

3.4. One-minute word and pseudo word reading

We analyzed the mean score for word and pseudo word reading combined. We were notified by testers that on several occasions the one-minute time limit was not observed. These instances were signaled on the tests and the scores of these children removed from analysis, leaving a smaller number of participants per group (read/math, $n = 99$; math/read, $n = 94$; control, $n = 123$). Both groups in the intervention group analysis improved over the three test periods, as revealed by a main effect of test period ($F(2, 382) = 1304, p < 0.001$). An effect of the reading predictor covariate was also significant ($F(1, 190) = 81.37, p < 0.001$). The group x test interaction approached significance ($F(2, 382) = 2.61, p = 0.075$). When adding the control group, significant effects were found for the main effect of test period ($F(2, 626) = 2164, p < 0.001$), the composite reading predictor ($F(1, 312) = 176, p < 0.001$) and, most importantly, the group x test interaction $F(4, 626) = 2.96, p = .019$. A post-hoc ANOVA at pretest showed no difference between the three groups at pretest $F(2, 313) = 0.45, p = 0.64$. An ANCOVA, using the predictive measures covariable, showed that at posttest A the general

1 progress made by all three groups was no longer equal ($F(2, 313) = 4.44, p < 0.01$) with the
2 read/math group reading significantly more items per minute than the math/read group
3 ($F(1,190) = 7.19, p < 0.01$) and the control group ($F(1, 219) = 6.89, p < 0.01$). In quantitative
4 terms, after 3 ½ months of game play, the group that had used ELAN improved by .26 standard
5 deviations over the control group, and .27 standard deviations over the math software group.
6
7

8
9 The advantage obtained by the read/math group continued into posttest B ($F(2, 312)$
10 = 4.61, $p = 0.01$), with the read/math group continuing to read more items on average than
11 the math/read ($F(1, 190) = 4.55, p = 0.03$) and the control group ($F(1, 219) = 9.60, p < 0.01$). In
12 quantitative terms, at the end of the school year, the group that had used ELAN in the first
13 session, after 4 months of no longer playing ELAN, maintained an advantage of 0.34 standard
14 deviations over the control group, and a 0.27 standard deviation advantage over the math
15 software group.
16
17
18
19
20
21

22 3.5. Non-sense text reading

23 Children that were unable to read more than four words of the first ten words in the
24 one-minute word reading task were not asked to continue with the more difficult text reading
25 tasks. It was also reported that while some children read more than four words on the one-
26 minute reading task, this was done with great difficulty and the child refused to continue to
27 the more difficult reading tasks. As a consequence, the number of children taking the following
28 reading tests were smaller (pretest: read/math = 70, math/read = 45, control= 64; posttest A:
29 read/math = 112, math/read = 98, control = 121; posttest B: read/math = 113, math/read =
30 100, control = 126). Children unable to do these tests were scored at '0' in our analysis.
31
32

33 Analyses of the intervention group ANCOVA revealed a significant main effect of
34 improvement across the three tests ($F(2, 428) = 739, p < 0.001$) and an effect of the composite
35 pretest predictor ($F(1, 213) = 42.7, p < 0.001$). No group x test interaction was found ($F(2, 428)$
36 = 2.39 $p = .09$). When including the control group, again, there was a main of effect of test
37 period ($F(2, 680) = 1181, p < 0.001$), and a significant effect of the composite pretest predictor
38 ($F(1, 339) = 116, p < 0.001$), but no group x test interaction ($F(4, 680) = 1.70 p = 0.15$).
39
40

41 3.6. Sentence reading comprehension

42 In this test, children read 8 sentences and, for each, selected the picture depicting their
43 meaning among 4 possible choices. The intervention group ANCOVA showed a significant
44 main effect of test period ($F(2, 428) = 596, p < 0.001$) and a significant effect of the composite
45
46
47
48
49
50
51
52
53

pretest predictor ($F(1, 213) = 44.38, p < 0.001$). Most importantly, a significant group \times test period interaction was also found ($F(2, 428) = 3.75, p = 0.02$). Post-hoc tests at each time point revealed no difference between the two intervention groups at pretest ($F(1, 214) = .74, p = 0.39$), but a significant difference at posttest A ($F(1, 214) = 11.49, p < 0.001$) and posttest B ($F(1, 214) = 6.41, p = 0.012$).

Similarly, the ANCOVA including the control group also revealed a significant main effect of test ($F(2, 680) = 880, p < 0.001$), an effect of the composite pretest predictor ($F(1, 339) = 95.95, p < 0.001$), and a group \times test interaction ($F(4, 680) = 2.85, p = .023$). A post-hoc ANOVA at pretest showed no difference between the groups at the start of the year ($F(2, 339) = .47, p = 0.65$). By posttest A, however, the general progress made by all three groups was no longer equal, ($F(2, 339) = 7.92, p < 0.001$) with the read/math group answering more questions correct than either the math/read group ($F(1, 213) = 12.34, p < 0.001$) or the control group ($F(1, 237) = 11.75, p < 0.001$). No difference was found between the math/read and control groups ($F(1, 227) = 0.03, p = 0.87$). In quantitative terms, after 3 ½ months of game play, the group that had used ELAN improved by .39 standard deviations over the control group, and .41 standard deviations over the math software group

The group differences continued into the posttest B ($F(2, 339) = 3.49, p = 0.032$), with the read/math group continuing to answer more items correct than either the math/read ($F(1, 213) = 6.96, p < 0.01$) or the control group ($F(1, 237) = 3.92, p < 0.05$). No difference was found between the children that used ELAN in the second period and the control group ($F(1, 227) = .55, p = 0.46$). At the end of the school year, the group that had used ELAN in the first session, after 4 months of no longer playing ELAN, maintained an advantage of 0.22 standard deviations over the control group, and a 0.31 standard deviations advantage over the math software group.

Fig4(a-f) Performance improvements pertaining to reading acquisition. Panels indicate performance in tests of (a) syllable awareness; (b) phoneme awareness (combined score of CV and CVV suppression in pseudo-words); (c) letter knowledge (combined score of letter name and sound knowledge); (d) word and pseudo-word reading (average number of items read in one minute); (e) non-sense text reading (total number of words correctly read in a non-sense text within a 3-minute time limit); (f) sentence comprehension (reading a short sentence and choosing the correct image out of 4 semantically or syntactically related pictures). All tests improved across time. The p -value indicates the significance of the 3×3 group \times time interaction. Pairwise differences that were significant at $p < 0.05$ are highlighted (RM= read/math group; MR= math/read group; C=control group).

3.7. Measuring the evolution of the word length effect

ELAN was developed to help automatize the child's knowledge of grapheme-phoneme correspondences and word reading. The game pushes the child to respond faster by increasing game difficulty and requiring faster and faster responses. As mentioned in section 1.4.3, we used a lexical decision task at the end of each game level to evaluate how reading speed varied with word length.

Our main goal was to measure if the child's word reading speed became progressively less affected by word length over the course of the game, in other words, whether ELAN could be used to measure when students move from slow decoding to parallel word reading. Therefore, we only looked at data from children that had passed at least the first 4 game levels, meaning that these children had learned all the vowels, fricative graphemes and the most common occlusive consonants. 21 children from the read/math group and 44 children from the math/read group were included in the following analysis. The software measured the time elapsed between the presentation of a target word or pseudo-word and the child's decision. We removed outlier trials with response times below 325 milliseconds (less than 1% of trials) or above a fixed cutoff, defined for a given cell of the design as being 2 standard deviations from the mean of the within-participants medians in that cell (6 % of trials). The medians of the remaining correct response times were entered into a mixed ANOVA with participant as the random factor, intervention as the between factor and word length (number of letters, from 2 to 5), test number (test numbers 1 through 4), and lexical status (word or pseudo-word) as within factors. All effects are reported as significant at $p < 0.05$. Degrees of freedom are adjusted to reflect the fact that some cells had missing data for some participants.

The mean of the median response times for each intervention group at each test period and for all lengths are shown in Figure 5. There was a significant length effect: as expected, response times increased as the number of letters increased from 2 to 5 ($F(3, 175) = 190.21, p < 0.001$). We also found a main effect of lexical status ($F(1, 51) = 243.81, p < 0.001$) indicating that responses were overall faster for words than for pseudo-words, and a length x lexical status interaction ($F(3, 180) = 35.39, p < 0.001$), due to the fact that the slope of the length effect was shallower for words than for pseudo-words. Finally, all of these effects were qualified by a length x test number ($F(9, 557) = 13.06, p < 0.001$) and a triple interaction of length x test number x lexical status ($F(9, 543) = 7.12, p < 0.001$). As seen in Figure 5, over the

1
2
3
4
5
6
7
8
9
four successive tests, the length effect became shallower for words (length x test number interaction restricted to words only ($F(9, 548) = 21.15, p < 0.001$), whereas this improvement was also present but less pronounced for pseudo-words ($F(9, 562) = 4.79, p < 0.001$). In agreement with this interpretation, at test number 4, the length effect remained significant for pseudo-words ($F(3, 188) = 36.63, p < 0.001$) but no longer for words ($F(3, 178) = 0.95, n.s.$).

10
11
12
13
14
15
16
17
18
All of the above effects were true within both the math/read and the read/math groups. There were no interactions involving this intervention factor. Only a main effect of intervention was observed ($F(1, 51) = 5.52, p = 0.03$), reflecting the fact that the math/read group responded faster, in line with the fact that these children took the tests later in the year than the read/math group.

19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
In summary, the lexical decision task concurred with paper tests in indicating that subjects in group read/math, who started with ELAN software, gained in reading acquisition. Superficially, the absence of any interaction with intervention in the lexical decision task seemed to suggest that the intervention did not have an effect on the evolution of the word length effect, and that the improvements were solely due to time elapsed at school. However, it should be remembered that the children in the read/math group were tested within the first half of the year. Thus, their fast progression and the disappearance of the word length effect by test 4 are remarkable and do suggest that the software could have had an effect over and above the improvements due to schooling itself.

37
38
39
40
41
42
43
44
45
46
47
To evaluate this, we compared the performance of the read/math group at test 4 with that of the math/read group at test 1. The rationale was that if effects were due solely to time at school, the math/read group should exhibit a better performance on its test 1 (because it was taken later in the year and used grapheme-phoneme correspondences that children would already be very familiar with), whereas if the software intervention was the primary driver of performance, then the converse should hold.

48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
We entered the corresponding median response times into a mixed ANOVA with subject as the random factor, intervention as a 2-level between factor and word length (number of letters, from 2 to 5) and lexical status (word or pseudo-word) as within factors. There were main effects of word length $F(3, 186) = 68.72, p < 0.001$ and lexical status $F(1, 60) = 96.70, p < 0.001$. Most importantly, the length x intervention was significant ($F(3, 186) = 3.50, p = 0.017$), confirming that, indeed, the read/math group, although tested at an earlier time, had a shallower length effect than the math/read group. Furthermore, the triple

1 interaction intervention x length x lexical status was significant ($F(3, 179) = 6.14, p = 0.001$).
 2 This was due to the fact that intervention reduced the length effect down to non-significance
 3 for words only for the read/math group (length x intervention interaction, restricted to words:
 4 ($F(3, 180) = 5.46, p = 0.001$) and to pseudo-words ($F(3, 188) = 2.76, p = 0.04$)). In detail, for the
 5 read/math group, no effect of word length was found when reading words on test 4 ($F(3, 53)$
 6 $= .94, n.s.$), while the length remained significant for pseudo-word reading ($F(3, 59) = 18.19, p$
 7 < 0.001). For the math/read group, a significant effect of word length was found at test 1 for
 8 words ($F(3, 127) = 36.67, p < 0.001$) and for pseudo-words ($F(3, 129) = 44.36, p < 0.001$).
 9

10
 11
 12
 13
 14
 15 In summary, these results demonstrate that the children who used ELAN in the first
 16 session ceased to be affected by word length at an earlier time than children in the math/read
 17 group. These results lend support to ELAN as a catalyst to parallel reading.
 18
 19
 20
 21
 22

23 *Fig5 Measuring the speed of reading and its improvement over time using a simple lexical decision game. Passing the lexical*
 24 *decision task was required at the end of each level to proceed to the next. Children had to decide if a word presented was*
 25 *either a real word or an invented word. 32 words were presented: four words of 2, 3, 4 or 5 letters for both words and pseudo*
 26 *words. All of the words were decodable from the lessons completed by the child. We measured reading response time by word*
 27 *length for both words and pseudo words of the two different groups. Group 1 played ELAN during session A of the intervention.*
 28 *Group 2 played ELAN during session B of the intervention.*
 29
 30

31 3.8. Number Comparison

32
 33
 34 The intervention group analysis on the number of comparison pairs correctly solved in
 35 one minute revealed a main effect of test period ($F(2, 428) = 116, p < 0.001$), but no significant
 36 group x test interaction ($F(2, 428) = .92, p = 0.40$). The 3-group analysis yielded a significant
 37 interaction, however: ($F(4, 680) = 3.25, p = 0.01$). Post-hoc analysis revealed no difference
 38 between the 3 groups at pretest ($F(2, 340) = .32, p = 0.73$). At posttest A, there continued to
 39 be no difference between the two intervention groups ($F(1, 214) = .14, p = 0.71$), but both the
 40 read/math and the math/read were superior to the control group (respectively $F(1, 238) =$
 41 $9.45, p < 0.01$; and $F(1, 228) = 12.54, p < 0.001$). At posttest B, there continued to be no
 42 difference between the two intervention groups ($F(1, 214) = .54, p = 0.46$), but the differences
 43 relative to the control group remained (respectively $F(1, 238) = 9.28, p < 0.001$; and $F(1, 228)$
 44 $= 5.79, p = 0.02$). These results point to a possible tablet effect as all children that used the
 45 tablets significantly improved over the no-tablet control group.
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65

3.9. Subtraction

The child's percentage of correct responses in subtraction was entered into an intervention group ANOVA, which revealed a significant main effect of the test period ($F(2, 428) = 253, p < 0.001$), but no group \times test interaction ($F(2, 428) = .314, p = 0.73$). In the 3-group analysis, the main effect of test was significant ($F(2, 680) = 417, p < 0.001$), but again, no group \times test interaction was found ($F(4, 680) = .27, p = 0.90$).

Fig6 Performance improvements in numerical tests. Two tests were used to measure possible effects of the control game intervention: (a) Symbolic number comparison (% correct in deciding which of two Arabic numerals is larger); (b) Subtraction (% correct). Same format as figure 4. All tests improved across time. The p-value indicates the significance of the 3x3 group \times time interaction. Pairwise differences that were significant at $p < 0.05$ are highlighted (RM= read/math group; MR= math/read group; C=control group).

4. Discussion

4.1. Summary of results

The progress of the read/math group at posttest A, which continued to posttest B on the one minute reading test and sentence comprehension, coupled with the evidence from the lexical decision task for a reduced length effect for this group give strong support for ELAN as a tool to aiding classroom reading instruction and accelerating first grade reading acquisition. The observed effect sizes of 0.26 - 0.41 are considered as small to medium, but more importantly, the benefits seem to extend four months beyond the training period. Interestingly, these results pertain to those children that only used the software in the initial learning stages of reading. Given the short duration of the training period (read/math $\bar{x}=8.8$ hrs, $sd=1.9$; math/read $\bar{x}=8.5$ hrs, $sd=2.3$), it is likely that improved results would be obtained with prolonged and/or more intense exposure.

Our results are consistent with previous studies showing that early explicit systematic phonics provides the fastest route to reading acquisition (Cunningham, 2001; Ehri et al., 2001). This study also gives strength to a recent observational report of 2,500 first-grade students in France, which showed that reading ability improved faster in classes where a larger number of grapheme-phoneme correspondences were taught early in the year (Goigoux, 2016). Phonics methods are often criticized for not supporting comprehension due to the limiting constraints of word choice during the early lessons in methods that strictly adhere to 100% decodable text. In this context, however, it is particularly important to emphasize that sentence reading comprehension improved in the read/math group relative to both active and no-tablet control groups, even though our software was entirely focused on grapheme-

1 phoneme teaching and 100% decoding practice. While we cannot disentangle the benefits of
2 using ELAN from other instructional materials used by the participating teachers, this aspect
3 of our results would seem to support the early use of phonics to help children improve access
4 to comprehension. It is compatible with the classical equation that defines reading
5 comprehension as the product of decoding and oral comprehension (Gough & Tunmer, 1986;
6 Hoover' & Gough, n.d.; Lervåg, Hulme, & Melby-Lervåg, 2018). In other words, once a text is
7 decoded the reader uses the same mechanisms that he or she would apply to hearing
8 language. By facilitating a faster automatization of decoding skills, the reader can focus on
9 what they have read. Early intense practice of phonics maybe key to alleviating the inefficient
10 difficulty of the phonological path of slow decoding so that the child can quickly access the
11 lexical path of fluent reading for understanding. The goal for developing ELAN was, in fact,
12 that the software help automatize the rote task of grapheme-phoneme correspondence
13 memorization and allow teachers to focus on developing oral comprehension and reading
14 aloud.

15
16
17
18
19
20
21
22
23
24
25
26
27 While it is disappointing that the NumberCatcher game did not have a more specific
28 influence on number comparison, and had no detectable effect on subtraction abilities, the
29 results suggest instead that tablet software itself may have a generic effect on speeded
30 decision tasks, as both tablet groups improved in the number comparison test. Numerous
31 studies on 'first person shooter games', video games that require the player to make
32 combative actions from a first person perspective, improve a number of cognitive functions,
33 including reaction time to 'go signals' (Colzato, Wildenberg, Zmigrod, & Hommel, 2013) and
34 use of attentional resources in simple tasks of flanker compatibility and spatial localization
35 (Green and Bavelier, 2006, 2003). While neither ELAN or Number Catcher would be classified
36 as 'action games', they both require players to make faster and faster actions in environments
37 with many distractors. Thus, while our isolated result should be taken with caution, given that
38 the benefit in number comparison was only obtained relative to a non-randomized no-tablet
39 control group, future study of more general cognitive benefits of educational games that
40 require timed responses in noisy environments is warranted.

4.2. Limitations

41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

to such an extent that only 47% of them experienced a game intervention adapted to their progress and level. While it is virtually impossible to create bug-free software without broad field testing, improved piloting was warranted in this case. Children were sometimes stymied from advancing in the game due to poor ergonomics that were not discovered until a close examination of the children’s game progression.

A second limitation to this study concerns the absence of results with students that used ELAN in session B. ELAN adapts to the child’s response rate, but requires completion of all lessons to move through the game. It is highly likely, given the improved reading scores for all children at posttest A, that for the math/read group, the grapheme-phoneme correspondences presented during the second session were mainly a review of already well-learned material in the class. Ideally, ELAN should use an adaptive mechanism that not only provides tailored game difficulty, but adjusts more quickly in order to specifically train the grapheme-phoneme correspondences that the child has not learned or needs to practice.

A third limitation to this study is that we measured the benefits of autonomous use of ELAN without taking into account other factors known to influence reading acquisition, such as socio-economic status on academic achievement (for a review of the effects of SES on academic achievement: (Sirin, 2005) and the effects of materials and methods already in use by the participating teachers (Braibant, & Gerard, 1996). The benefits of ELAN might have been partially dependent on the activities already in place by the teacher. For example, it is unknown to us which teachers were already using systematic phonics methods and how many grapheme-phoneme correspondences were taught in the class. If ELAN were to be used in the classroom, it would be important to know if the software itself is sufficient in phonics teaching or if it should be used to provide repetition of the teacher’s lesson.

4.3. Conclusion

In a recent review on the science of learning to read, it has been proposed that the “reading wars” debate between whole-word methods and phonics needs to end (Castles et al., 2018). Enough evidence has been accumulated to support explicit systematic phonics and our focus should be on developing guiding principles for “balanced instruction” that acknowledge the complexities of learning that underlie the individual child’s progress from decoder to reading expert. This requires combining an understanding of the biological and cognitive aspects of learning to read, respecting the evidence-based practices already known and furthering research on the bricks of best-practice that build a reading method from start

to finish. In the case of this project, it appears that intense and focused learning of decoding skills may be more effective when taught early in learning. We also demonstrate that software that focuses on the repetition of decoding skills, and that adapts to the child's level, can indeed lead to beneficial effects for young readers. More importantly, this work builds onto the evidence that addressing 'which' strategies and 'when' should now be the focus of reading research.

Bibliography

- Bara, F., & Gentaz, E. (2011). *Haptics in teaching handwriting: The role of perceptual and visuo-motor skills*. *Human Movement Science*, 30(4), 745–759. <https://doi.org/10.1016/j.humov.2010.05.015>
- Bara, F., Gentaz, E., & Colé, P. (2007). *Haptics in learning to read with children from low socio-economic status families*. *British Journal of Developmental Psychology*, 25(4), 643–663. <https://doi.org/10.1348/026151007X186643>
- Bara, F., Gentaz, E., Colé, P., & Sprenger-Charolles, L. (2004). *The visuo-haptic and haptic exploration of letters increases the kindergarten-children's understanding of the alphabetic principle*. *Cognitive Development*, 19(3), 433–449. <https://doi.org/10.1016/j.cogdev.2004.05.003>
- Bara, F., Morin, M.-F., Alamargot, D., & Bosse, M.-L. (2016). *Learning different allographs through handwriting: The impact on letter knowledge and reading acquisition*. *Learning and Individual Differences*, 45, 88–94. <https://doi.org/10.1016/j.lindif.2015.11.020>
- Bertrand, D., Fluss, J., Billard, C., & Ziegler, J. C. (2010). *Efficacité, sensibilité, spécificité : comparaison de différents tests de lecture*, Abstract. *L'Année psychologique*, 110(2), 299–320. <https://doi.org/10.4074/S000350331000206X>
- Boles, D. B., & Clifford, J. E. (1989). *An upper- and lowercase alphabetic similarity matrix, with derived generation similarity values*. *Behavior Research Methods, Instruments, & Computers*, 21(6), 579–586. <https://doi.org/10.3758/BF03210580>
- Boyer, N., & Ehri, L. C. (2011). *Contribution of Phonemic Segmentation Instruction With Letters and Articulation Pictures to Word Reading and Spelling in Beginners*. *Scientific Studies of Reading*, 15(5), 440–470. <https://doi.org/10.1080/10888438.2010.520778>
- Braibant, J. M., & Gerard, F.-M. (1996). *Savoir lire : question(s) de méthodes ?* *Bulletin de psychologie scolaire et d'orientation*, 7-45.
- Brankaer, C., Ghesquière, P., & Smedt, B. D. (2017). *Symbolic magnitude processing in elementary school children: A group administered paper-and-pencil measure (SYMP Test)*. *Behavior Research Methods*, 49(4), 1361–1373. <https://doi.org/10.3758/s13428-016-0792-3>
- Brem, S., Bach, S., Kucian, K., Kujala, J. V., Guttorm, T. K., Martin, E., ... Richardson, U. (2010). *Brain sensitivity to print emerges when children learn letter–speech sound correspondences*. *Proceedings of the National Academy of Sciences*, 107(17), 7939–7944. <https://doi.org/10.1073/pnas.0904402107>
- Carpenter, S. K., Cepeda, N. J., Rohrer, D., Kang, S. H. K., & Pashler, H. (2012). *Using Spacing to Enhance Diverse Forms of Learning: Review of Recent Research and Implications for*

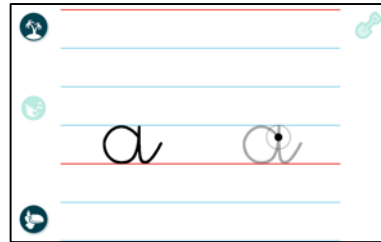
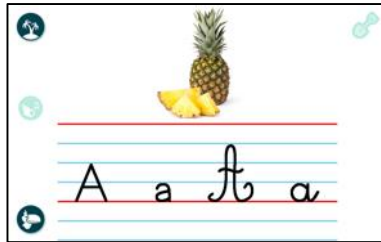
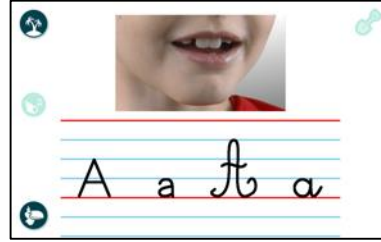
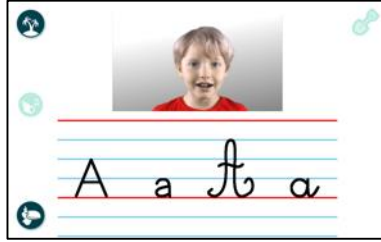
- Instruction. *Educational Psychology Review*, 24(3), 369–378.
<https://doi.org/10.1007/s10648-012-9205-z>
- 1
2
3 Castiglioni-Spalten, M. L., & Ehri, L. C. (2003). Phonemic Awareness Instruction: Contribution
4 of Articulatory Segmentation to Novice Beginners' Reading and Spelling. *Scientific*
5 *Studies of Reading*, 7(1), 25–52. https://doi.org/10.1207/S1532799XSSR0701_03
- 6
7 Castles, A., Rastle, K., & Nation, K. (2018). Ending the Reading Wars: Reading Acquisition From
8 Novice to Expert. *Psychological Science in the Public Interest*, 19(1), 5–51.
9 <https://doi.org/10.1177/1529100618772271>
- 10
11 Cavalli, E., Duncan, L. G., Elbro, C., Ahmadi, A. E., & Colé, P. (2017). Phonemic—Morphemic
12 dissociation in university students with dyslexia: an index of reading compensation?
13 *Annals of Dyslexia*, 67(1), 63–84. <https://doi.org/10.1007/s11881-016-0138-y>
- 14
15 Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., & Rohrer, D. (2006). Distributed practice in
16 verbal recall tasks: A review and quantitative synthesis. *Psychological Bulletin*, 132(3),
17 354–380. <https://doi.org/10.1037/0033-2909.132.3.354>
- 18
19 Coltheart, M. (2005). Modeling Reading: The Dual-Route Approach. In M. J. Snowling & C.
20 Hulme (Eds.), *The science of reading: A handbook* (pp. 6–23). Malden: Blackwell
21 Publishing. <https://doi.org/10.1002/9780470757642.ch1>
- 22
23 Colzato, L. S., Wildenberg, W. P. M. van den, Zmigrod, S., & Hommel, B. (2013). Action video
24 gaming and cognitive control: playing first person shooter games is associated with
25 improvement in working memory but not action inhibition. *Psychological Research*,
26 77(2), 234–239. <https://doi.org/10.1007/s00426-012-0415-2>
- 27
28 Cunningham, J. W. (2001). The National Reading Panel Report. *Reading Research Quarterly*,
29 36(3), 326–335. <https://doi.org/10.1598/RRQ.36.3.5>
- 30
31 de Graaff, S., Bosman, A. M. T., Hasselman, F., & Verhoeven, L. (2009). Benefits of Systematic
32 Phonics Instruction. *Scientific Studies of Reading*, 13(4), 318–333.
33 <https://doi.org/10.1080/10888430903001308>
- 34
35 Dehaene, S., Pegado, F., Braga, L. W., Ventura, P., Filho, G. N., Jobert, A., ... Cohen, L. (2010).
36 How Learning to Read Changes the Cortical Networks for Vision and Language. *Science*,
37 330(6009), 1359–1364. <https://doi.org/10.1126/science.1194140>
- 38
39 Dehaene-Lambertz, G., Monzalvo, K., & Dehaene, S. (2018). The emergence of the visual word
40 form: Longitudinal evolution of category-specific ventral visual areas during reading
41 acquisition. *PLOS Biology*, 16(3), e2004103.
42 <https://doi.org/10.1371/journal.pbio.2004103>
- 43
44 Deltour, J.-J., & D, H. (1990). *Test de vocabulaire actif et passif pour enfants de 3 à 5 ans -*
45 *TVAP*. Issy-les-Moulineaux: Ed. scientifiques et psychologiques.
- 46
47 Dye, M. W. G., Green, C. S., & Bavelier, D. (2009). Increasing Speed of Processing With Action
48 Video Games. *Current Directions in Psychological Science*, 18(6), 321–326.
49 <https://doi.org/10.1111/j.1467-8721.2009.01660.x>
- 50
51 Ehri, L. C., Nunes, S. R., Stahl, S. A., & Willows, D. M. (2001). Systematic Phonics Instruction
52 Helps Students Learn to Read: Evidence from the National Reading Panel's Meta-
53 Analysis. *Review of Educational Research*, 71(3), 393–447.
54 <https://doi.org/10.3102/00346543071003393>
- 55
56 Gentaz, E., Theurel, A., & Sprenger-Charolles, L. (2017). EVALUER LES CAPACITÉS DE LECTURE
57 CHEZ LES ENFANTS DE 6-7 ANS, 134.
- 58
59 Goigoux, R. (2016). Apprendre à lire et à écrire au cours préparatoire : enseignements d'une
60 recherche collective. *Revue française de pédagogie. Recherches en éducation*, (196), 5–
61 6.
62
63
64
65

- 1 Gough, P. B., & Tunmer, W. E. (1986). *Decoding, Reading, and Reading Disability. Remedial and Special Education*, 7(1), 6–10. <https://doi.org/10.1177/074193258600700104>
- 2 Green, C. S., & Bavelier, D. (2003). Action video game modifies visual selective attention, 423,
3 6.
- 4 Green, C. S., & Bavelier, D. (2006). Effect of action video games on the spatial distribution of
5 visuospatial attention. *Journal of Experimental Psychology: Human Perception and*
6 *Performance*, 32(6), 1465–1478. <https://doi.org/10.1037/0096-1523.32.6.1465>
- 7 Hoover, A., & Gough, B. (n.d.). *The simple view of reading*, 34.
- 8 Kyle, F., Kujala, J., Richardson, U., Lyttinen, H., & Goswami, U. (2013). Assessing the
9 Effectiveness of Two Theoretically Motivated Computer-Assisted Reading Interventions
10 in the United Kingdom: GG Rime and GG Phoneme. *Reading Research Quarterly*, 48(1),
11 61–76. <https://doi.org/10.1002/rrq.038>
- 12 Lecocq, P. (1996). *L'ECoSse: une épreuve de compréhension syntaxico-sémantique*. Presses
13 Universitaires du Septentrion, Lille.
- 14 Lefavrais, P. (2005). *Alouette-R, Test d'analyse de la lecture et de la dyslexie*. Paris.
- 15 Lervåg, A., Hulme, C., & Melby-Lervåg, M. (2018). Unpicking the Developmental Relationship
16 Between Oral Language Skills and Reading Comprehension: It's Simple, But Complex.
17 *Child Development*, 89(5), 1821–1838. <https://doi.org/10.1111/cdev.12861>
- 18 Lété, B., Sprenger-Charolles, L., & Colé, P. (2004). MANULEX: a grade-level lexical database
19 from French elementary school readers. *Behavior Research Methods, Instruments, &*
20 *Computers: A Journal of the Psychonomic Society, Inc*, 36(1), 156–166.
- 21 Longcamp, M., Zerbato-Poudou, M.-T., & Velay, J.-L. (2005). The influence of writing practice
22 on letter recognition in preschool children: A comparison between handwriting and
23 typing. *Acta Psychologica*, 119(1), 67–79.
24 <https://doi.org/10.1016/j.actpsy.2004.10.019>
- 25 Lyttinen, H., Ronimus, M., Alanko, A., Poikkeus, A.-M., & Taanila, M. (2007). Early
26 identification of dyslexia and the use of computer game-based practice to support
27 reading acquisition. *Nordic Psychology*, 59(2), 109–126. <https://doi.org/10.1027/1901-2276.59.2.109>
- 28 Mason, G., McDaniel, H., & Callaway, B. (1974). Relating Reading and Spelling: A Comparison
29 of Methods. *The Elementary School Journal*, 74(6), 381–386.
- 30 McCandliss, B., Beck, I. L., Sandak, R., & Perfetti, C. (2003). Focusing attention on decoding for
31 children with poor reading skills: Design and preliminary tests of the word building
32 intervention. *Scientific Studies of Reading*, 7(1), 75–104.
- 33 Meyer, J., Dentel, L., & Meunier, F. (2013). Speech Recognition in Natural Background Noise.
34 *PLOS ONE*, 8(11), e79279. <https://doi.org/10.1371/journal.pone.0079279>
- 35 Monzalvo, K., & Dehaene-Lambertz, G. (2013). How reading acquisition changes children's
36 spoken language network. *Brain and Language*, 127(3), 356–365.
37 <https://doi.org/10.1016/j.bandl.2013.10.009>
- 38 New, B., Ferrand, L., Pallier, C., & Brysbaert, M. (2006). Reexamining the word length effect in
39 visual word recognition: New evidence from the English Lexicon Project. *Psychonomic*
40 *Bulletin & Review*, 13(1), 45–52. <https://doi.org/10.3758/BF03193811>
- 41 OECD. (2015). *Students, Computers and Learning*. OECD Publishing.
42 <https://doi.org/10.1787/9789264239555-en>
- 43 Pourcin, L., Sprenger-Charolles, L., El Ahmadi, A., & Colé, P. (2016). Reading and related skills
44 in Grades 6, 7, 8 and 9: French normative data from EVALEC. *Revue Européenne de*
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

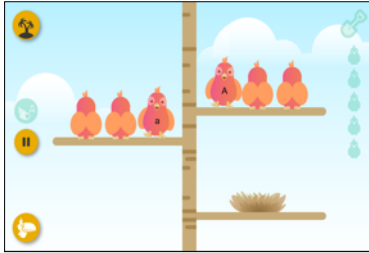
- 1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
- Psychologie Appliquée/European Review of Applied Psychology*, 66(1), 23–37.
<https://doi.org/10.1016/j.erap.2015.11.002>
- Räsänen, P., Salminen, J., Wilson, A. J., Aunio, P., & Dehaene, S. (2009). Computer-assisted intervention for children with low numeracy skills. *Cognitive Development*, 24(4), 450–472. <https://doi.org/10.1016/j.cogdev.2009.09.003>
- Richardson, U., & Lyytinen, H. (2014). The GraphoGame Method: The Theoretical and Methodological Background of the Technology-Enhanced Learning Environment for Learning to Read. *Human Technology*, 10(1). <https://doi.org/10.17011/ht/urn.201405281859>
- Saine, N. L., Lerkkanen, M.-K., Ahonen, T., Tolvanen, A., & Lyytinen, H. (2011). Computer-Assisted Remedial Reading Intervention for School Beginners at Risk for Reading Disability: Computer-Assisted Reading Intervention. *Child Development*, 82(3), 1013–1028. <https://doi.org/10.1111/j.1467-8624.2011.01580.x>
- Schacter, J., & Jo, B. (2016). Improving low-income preschoolers mathematics achievement with Math Shelf, a preschool tablet computer curriculum. *Computers in Human Behavior*, 55, 223–229. <https://doi.org/10.1016/j.chb.2015.09.013>
- Sirin, S. R. (2005). Socioeconomic Status and Academic Achievement: A Meta-Analytic Review of Research. *Review of Educational Research*, 75(3), 417–453. <https://doi.org/10.3102/00346543075003417>
- Tallal, P. (2004). Improving language and literacy is a matter of time. *Nature Reviews Neuroscience*, 5(9), 721–728. <https://doi.org/10.1038/nrn1499>
- Uhry, J. K., & Shepherd, M. J. (1993). Segmentation/Spelling Instruction as Part of a First-Grade Reading Program: Effects on Several Measures of Reading. *Reading Research Quarterly*, 28(3), 219–233. <https://doi.org/10.2307/747995>
- Wilson, A. J., Dehaene, S., Dubois, O., & Fayol, M. (2009). Effects of an Adaptive Game Intervention on Accessing Number Sense in Low-Socioeconomic-Status Kindergarten Children. *Mind, Brain & Education*, 3(4), 224–234. <https://doi.org/10.1111/j.1751-228X.2009.01075.x>
- Wilson, A. J., Dehaene, S., Pinel, P., Revkin, S. K., Cohen, L., & Cohen, D. (2006). Principles underlying the design of. *Behavioral and Brain Functions*, 2(1), 19.
- Wilson, A. J., Revkin, S. K., Cohen, D., Cohen, L., & Dehaene, S. (2006). An open trial assessment of "The Number Race", an adaptive computer game for remediation of dyscalculia. *Behavioral and Brain Functions*, 2(1), 1.
- Zoccolotti, P., De Luca, M., Di Pace, E., Gasperini, F., Judica, A., & Spinelli, D. (2005). Word length effect in early reading and in developmental dyslexia. *Brain and Language*, 93(3), 369–373. <https://doi.org/10.1016/j.bandl.2004.10.010>

Appendix, Order of grapheme-phonemes taught in ELAN

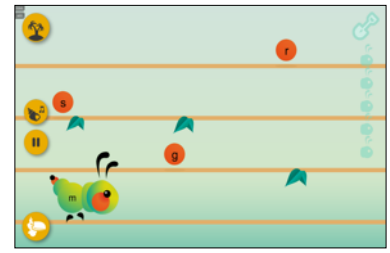
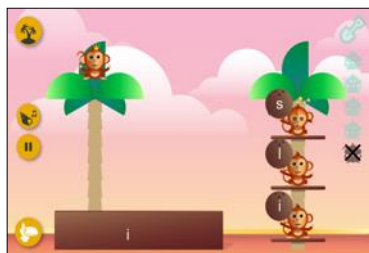
Lesson Number	Grapheme Orthography	Phoneme Code	Lesson Number	Grapheme Orthography	Phoneme Code
1	A, a	a	31	et, est	e,E
2	E, é	e	32	ent	silent
3	I, i	i	33	an, am	@
4	O, o	o	34	en, em	@
5	U, u	y	35	on, om	§
6	E, e	*	36	in, im, ym	5
7	Y, y	i	37	à, â	a
8	L, l	l	38	au, eau	O
9	M, m	m	39	è, ê, ë	E
10	S, s	s	40	oeu	9
11	R, r	R	41	gn	N
12	Ou	u	42	eConsonantC onsonant	E
13	silent letters (e, s, t, x)	silent	43	ai, ei	E
14	Jj,	Z	44	ez er	e
15	F, f	f	45	et, fin d'un mot	E
16	V, v	v	46	Hh	silent
17	eu	2,9	47	ph	f
18	Z, z	z	48	ç, ce, ci, cy	s
19	ch	S	49	ge, gi, gy	Z
20	Nn, un	n,1	50	Xx	ks
21	T, t	t	51	silent end letters (all)	silent
22	D, d	d	52	ain, aim	5
23	C, c	k	53	ien	j5
24	P, p	p	54	eConsonant	E
25	Qq, qu	k	55	ti	s
26	B, b	b	56	Yy	j
27	Gg, gu	g	57	il, ill	ij
28	K, k	k	58	il, ill	j
29	s	z	59	ch	k
30	es	E	60	W, w	w



a. Syllable games




b. Word building games



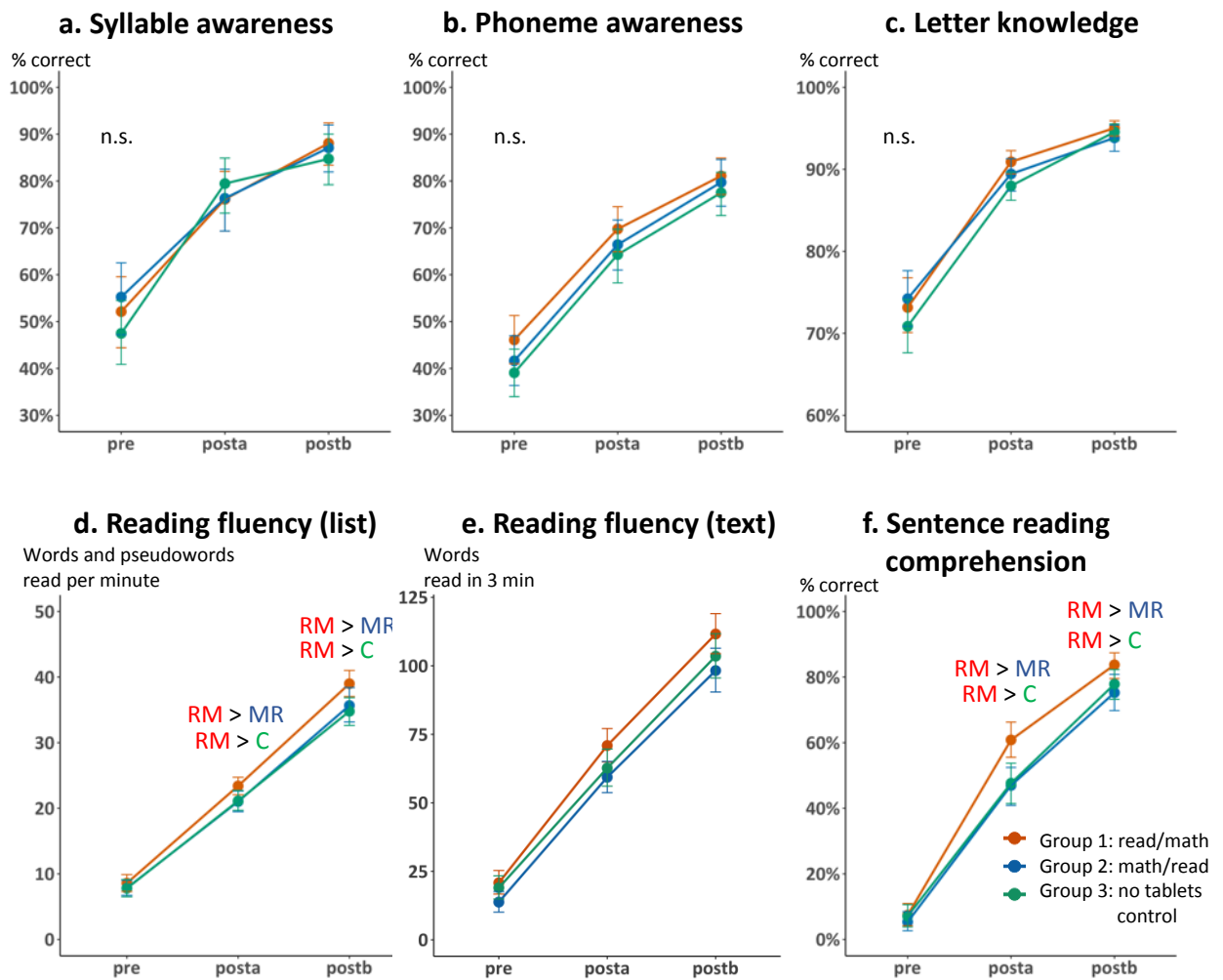
c. Comprehension and attention to French morphology



Study Design				
sept.	oct. – jan.	feb.	mar.-may	june
pretest	session A	posttest A	session B	posttest B
Group 1)	Reading (Elan)		Math	
Group 2)	Math		Reading (Elan)	
Group 3)	no tablets		no tablets	



Pictures taken during a class visit.

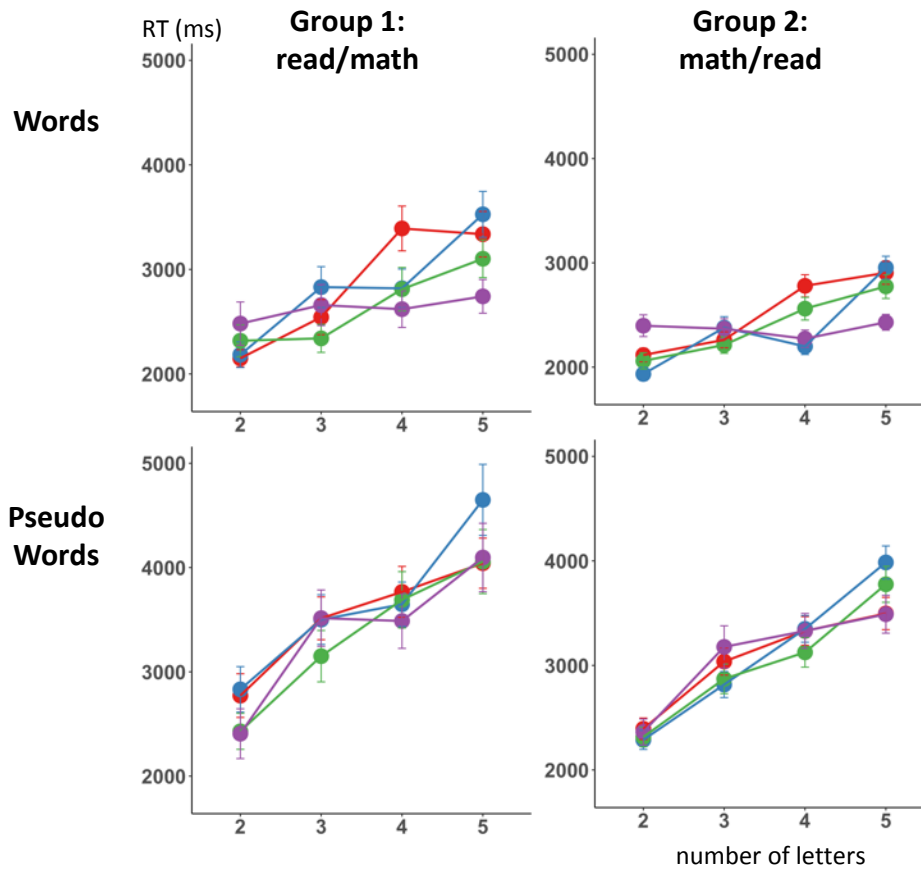


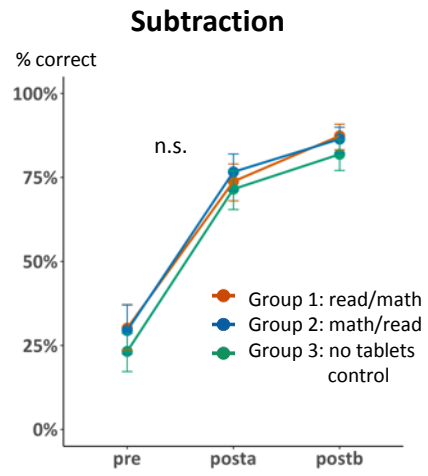
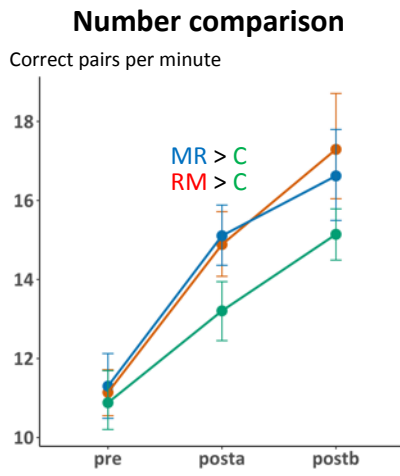


Lexical Decision Game

Successive Tests

- 1
- 2
- 3
- 4





Cassandra Potier Watkins is a second year PhD student, under the supervision of Stanislas Dehaene, at the Neurospin Cognitive Neuroimaging lab in France. Her doctoral research investigates the common threads that allow our brains to master cultural skills like reading and math, and in turn, how learning changes the brain. She takes a multidisciplinary approach that encompasses the fields of cognitive neuroscience and on-the-field education research. The goal of her work is advance the science of how we learn and develop applicable solutions for an evidence-based education. She is an associate member of the scientific council for the French Education Ministry.