frontiers in Education

ORIGINAL RESEARCH
published: 20 April 2021
doi: 10.3389/feduc.2021.639294

Check for updates

Neurocognitive Predictors of Response to Intervention With GraphoGame Rime

*Angela Wilson[1], Henna Ahmed[1], Natasha Mead[1], Hannah Noble[1], Ulla Richardson[2], Mary A. Wolpert[3] and Usha Goswami[1]**

[1] Centre for Neuroscience in Education, Department of Psychology, University of Cambridge, Cambridge, United Kingdom, [2] Centre for Applied Language Studies, Faculty of Humanities and Social Sciences, University of Jyväskylä, Jyväskylä, Finland, [3] Faculty of Education, University of Cambridge, Cambridge, United Kingdom

This study explores the neurocognitive predictors of response to intervention with GraphoGame Rime (GG Rime), an adaptive software game designed to aid the learning of English phonics. A cohort of 398 children (aged 6–7 years) who had participated in a recent randomised controlled trial (RCT) of GG Rime in the United Kingdom were studied. Half were randomly assigned to play GG Rime and the other half were assigned to Business As Usual (BAU). A series of pretests were given prior to the intervention to all participants, designed to measure phonological awareness skills, executive function (EF) skills and the ability to synchronise finger tapping to a rhythmic beat. Rhythmic synchronisation has been linked to reading readiness and early reading attainment, and is related to phonological awareness. Individual differences prior to the intervention in all three types of measure were significantly associated with progression through the game. Gender was also important for progression through the game, with boys progressing significantly further than girls. Vocabulary was not a predictor of progression through the game. Playing time, rhythmic synchronisation, phonological skills and EF skills did not differ by gender. Once playing time and non-verbal cognitive ability were controlled, phonological awareness, EF, rhythmic synchronisation and gender all remained significant predictors of progression through the game. In further analyses comparing these predictors, their interactions and controlling for the autoregressor of prior responsiveness to phonics instruction, phoneme awareness and EF skills were the strongest unique predictors. Analyses with the whole cohort (analysing BAU and GG children independently) showed that all neurocognitive measures contributed to progress in reading and spelling over the school year. We conclude that individual differences in phonological skills and EF skills predict which children will benefit most from computer assisted reading interventions like GG Rime. Further, boys respond better to this computerised intervention than girls. Accordingly, to be maximally beneficial to poor readers, the supplementary use of GG Rime in addition to ongoing classroom literacy instruction could be especially targeted to boys, but should be accompanied by a focus on developing both oral phonological awareness and EF skills.

Keywords: reading software, phonological awareness, phonics instruction, executive function, tapping accuracy

OPEN ACCESS

Edited by:
Alessandro Antonietti, Catholic University of the Sacred Heart, Italy

Reviewed by:
Patricia Vadasy, Oregon Research Institute, United States
George Manolitsis, University of Crete, Greece

***Correspondence:**
Usha Goswami
ucg10@cam.ac.uk

Specialty section:
This article was submitted to Educational Psychology, a section of the journal Frontiers in Education

Received: *08 December 2020*
Accepted: *29 March 2021*
Published: *20 April 2021*

Citation:
Wilson A, Ahmed H, Mead N, Noble H, Richardson U, Wolpert MA and Goswami U (2021) Neurocognitive Predictors of Response to Intervention With GraphoGame Rime. Front. Educ. 6:639294. doi: 10.3389/feduc.2021.639294

INTRODUCTION

Digital learning apps such as computer-assisted reading interventions (CARIs) hold great promise for improving the literacy learning environments of young children (Bus et al., 2020). Indeed, Bus et al. (2020) describe such apps as potential "game changers" for literacy learning. Yet there has been little systematic research into whether this promise is being fulfilled. For example, there is very little research into individual differences between children that may affect their ability to benefit from educational technologies such as CARIs. It is known that boys access the internet more than girls and spend longer online, although such gender differences are less marked for younger children (Livingstone and Helsper, 2007). Accordingly, it is possible that boys may benefit more from digital learning apps such as CARIs. Here, we use a recent RCT into the effectiveness of a CARI for teaching English phonics, GraphoGame Rime (GG Rime, see Worth et al., 2018; Ahmed et al., 2020), to identify which factors best predict response to literacy intervention with a CARI. We assess both classic neurocognitive predictors of reading development such as phonological awareness (Bus and van Ijzendoorn, 1999), and skills that lie in the domain of executive functions (EFs), such as sustained attention and the ability flexibly to operate the software. By hypothesis, EF skills are required for efficient gaming. Underlying sensory/neural differences that predict how well children will develop phonological awareness may also affect who will benefit from CARIs. Rhythm production tasks such as tapping to a beat represent one such factor (Woodruff Carr et al., 2014; Rios-Lopez et al., 2019). Accordingly, three broad classes of neurocognitive predictor were examined in the current study, phonological awareness, EF and rhythmic synchronisation. Finally, gender was included as a predictor, to explore whether boys may benefit more from educational technologies than girls.

No studies of neurocognitive predictors of response to intervention with CARIs are yet in the literature (to our knowledge). However, response to intervention has become an important part of the education lexicon (e.g., Linan-Thompson et al., 2006; van der Kleij et al., 2017). In the context of literacy learning environments, response to intervention is the degree to which a child who has been identified as being at risk for reading difficulties and who has been provided with a literacy intervention has then benefitted from this intervention. A range of studies has explored factors that may contribute to individual differences in response to literacy interventions (see meta-analysis by Nelson et al., 2003, and the literature synthesis by Al Otaiba and Fuchs, 2002). The majority of studies exploring response to literacy interventions by younger learners have reported that phonological awareness is the most important factor (Vaughn et al., 2003; Compton et al., 2006; Vellutino et al., 2006, 2008; Fletcher et al., 2011; Denton et al., 2013; Catts et al., 2015; van der Kleij et al., 2017). Furthermore, these studies typically reported a linear relationship with response to intervention. This suggests that the cognitive abilities underlying reading can be placed on a continuum, with lower phonological skills limiting the child's ability to respond to a literacy intervention in a linear manner (see Vellutino et al., 2006). Accordingly, the first research question addressed in the current study was whether pre-existing phonological awareness skills predicted response to intervention with GG Rime.

Executive functions are usually described as domain-general control processes that permit children to complete goal-directed activities (Miciak et al., 2019), and hence are rarely measured in standard reading intervention studies (i.e., non-CARI studies). EFs in children comprise three broad domains, working memory, inhibition and attentional flexibility (Hughes, 1998), and performance on EF measures has been associated with reading outcomes in the classroom (Yeniad et al., 2013). One prior study of which we are aware examined the effect of EF measures as factors determining a child's response to a literacy intervention (Miciak et al., 2019). This study reported that the EF skills of adequate versus inadequate responders to the intervention did not differ, despite the inclusion of six different EF measures spanning working memory, cognitive flexibility, planning and inhibitory skills. The authors had hypothesised that children with different levels of EF might respond differently to reading interventions, with good EF potentially compensating to some extent for skill deficits. This was not the case (Miciak et al., 2019). Since the literacy interventions studied by Miciak et al. (2019) were not delivered by computer, however, it is possible that EF may predict response to intervention with a CARI. Accordingly, the second research question addressed in the current study was whether pre-existing EF skills predicted response to intervention with GG Rime.

Both rhythm perception and rhythm production have been shown to be related to reading development in children (Goswami, 2011; Huss et al., 2011; Goswami et al., 2013). Rhythmic non-speech beat perception measured prior to schooling predicts reading development, letter-sound knowledge and phonological awareness from 3 to 6 years (Corriveau et al., 2010; Ozernov-Palchik et al., 2018). Rhythm production is typically measured by asking children to drum, clap or tap in time to a rhythmic beat. Individual differences in clapping, drumming and tapping tasks are reliably related to reading development across languages (e.g., Dellatolas et al., 2009 [French]; Bonacina et al., 2018 [English]; Lundetrae and Thomson, 2018 [Norwegian]). Rhythmic synchronisation is also related to pre-reading skills across languages, for example letter-name knowledge (Rios-Lopez et al., 2019) and phonological awareness (Woodruff Carr et al., 2014). The third research question addressed in the current study was whether pre-existing rhythmic synchronisation skills predicted response to intervention with GG Rime.

Finally, despite the importance of measuring autoregressor effects in longitudinal investigations, only one response to intervention study that we are aware of included an autoregressor (i.e., controlled for the child's previous response to literacy instruction before computing the effects of different neurocognitive predictors on children's current response to literacy interventions, Catts et al., 2015). In this study, which included 366 children, both phonological awareness and letter knowledge remained significant predictors of response to a literacy intervention when the autoregressor of prior responsiveness was controlled. The fourth hypothesis addressed in the current study was whether pre-existing phonological awareness, EF and rhythmic synchronisation skills would predict response to intervention with GG Rime after participants' prior sensitivity to phonics tuition had been controlled via an autoregressor. Finally, in prior response to intervention studies with younger readers, individual differences in intelligence have not typically predicted response to literacy interventions (Vellutino et al., 2006). However, in some studies, vocabulary has been a significant predictor of response to reading interventions (Compton et al., 2006; Vellutino et al., 2006). These factors were also measured prior to the intervention.

To address these questions, we utilised data from an RCT of the phonics software GG Rime (Worth et al., 2018). We studied the effects of three types of neurocognitive predictor (phonological awareness, EF, and rhythmic synchronisation) on children's response to intervention with this CARI. In prior publications based on this dataset, it has been established that GG Rime is as effective as BAU regarding progress in literacy (Worth et al., 2018), and that GG Rime children who are adequate responders make greater gains in phonic decoding skills than BAU children (non-word reading; Ahmed et al., 2020). The neurocognitive measures administered before this RCT commenced have not yet been studied, and these data are examined here. All children in the original RCT (Worth et al., 2018) had been identified as at risk for reading difficulties, as all children were in their second year of reading tuition in the United Kingdom (aged on average 6–7 years) and had failed the statutory test of phonic decoding skills that is administered in United Kingdom schools at the end of the first year of reading tuition (at age 5–6 years), the Phonics Check. A number of the children also had English as an Additional Language (EAL, N = 108). We defined response to intervention as the degree to which a child assigned to the GG Rime intervention had been able to progress through the game (see Ahmed et al., 2020). As all the children were in their second year of reading instruction and had not been progressing as expected under universal tuition, both the BAU control group and the GG Rime group would be classified as Tier 2 children in the United States. In the United Kingdom children who fail the Phonics Check are given extra literacy support. The GG Rime group received individualised computerised instruction as part of this extra support, typically being left to play GG Rime solo in a corner of the classroom or in the school library. The BAU control group received phonics tuition from a classroom teacher or teaching assistant for matched lengths of time, either in small groups or one-on-one, using a range of literacy materials, for example Reading Recovery. Accordingly, the GG Rime group received the computerised training instead of this varied phonics tuition, typically at the same time during the school day, most usually during the Literacy Hour that is a daily event in United Kingdom primary schools. Ahmed et al. (2020) reported that the children who progressed further through GG Rime and who thus could be considered adequate responders had significantly better literacy outcomes regarding learning phonic decoding skills than children in the control group who were receiving BAU. The key measure indexing phonic decoding skills in Ahmed et al. (2020) was the TOWRE PDE (Test of Word Reading Efficiency, Phonetic Decoding Efficiency, Torgesen et al., 1999).

The study by Ahmed et al. (2020) analysed the TOWRE data for the "top half" of GG Rime players only (designated adequate responders). The game takes players through 25 streams of phonic knowledge, and the mean progression point reached by the whole GG Rime group participating in the RCT was Stream 16, level 5, just over half-way through the game (range Stream 2, level 2 to stream 25 level 7, the final level). Ahmed et al. reported that children who showed progression through GG Rime that was above the average level attained by the entire GG Rime group of 195 children showed significant advantages in phonic learning (TOWRE PDE measure) compared to the BAU children. Further, these effects were particularly strong for various sub-groups of players, for example children in particularly poorly performing schools as designated by United Kingdom Government inspectors, or children who were younger when they received the intervention. For these sub-group analyses, the gains in phonic decoding skills conferred by GG Rime persisted over the school summer vacation, hence demonstrating long-term benefits of a CARI. However, Ahmed et al. (2020) did not analyse data from those GG Rime players whose progression through the game was below average (inadequate responders), and they did not analyse the neurocognitive predictor measures that had been administered to the entire cohort of 398 children before the intervention began. These analyses are provided in the current report.

The RCT was funded by the Education and Neuroscience scheme, a collaboration between the United Kingdom Education Endowment Foundation (EEF) and United Kingdom Wellcome Trust that was set up to enable RCTs to test promising educational interventions based on educational neuroscience. The scheme allocated independent evaluators to selected projects, and the independent evaluator selected the children for the GG Rime RCT, allocated them to the participant groups, and selected the efficacy measures. The full evaluation of GG Rime is available at https://educationendowmentfoundation.org.uk/projects-and-evaluation/projects/graphogame-rime/. The primary outcome measures were chosen by the independent evaluator and comprised the New Group Reading Test (NGRT) and the Single Word Spelling Test (SWST, both GL Assessment). These tests were administered by independent test administrators (provided by the independent evaluators) who were blind to group status, within a month of the intervention ending. The current authors administered the TOWRE. The TOWRE consists of two subtests measuring speeded decoding of words (SWE, Sight Word Efficiency) and non-words (PDE). The EEF report (Worth et al., 2018) concluded that the improvements made by the 195 children playing GG Rime solo as assessed by the NGRT and SWST were equivalent to the improvements made by the 196 children receiving BAU. It also concluded that the teachers and teaching assistants involved found the GG Rime intervention easy to set up and to implement for their children. The report noted that teachers, senior leaders and pupils considered GG Rime highly engaging, motivational and enjoyable.

In the current study, we use progression through the game as the primary outcome measure of response to the intervention

(Ahmed et al., 2020), and we also provide further analyses of the entire cohort (BAU + GG Rime children) in order to evaluate the strength of the chosen neurocognitive predictors regarding the literacy outcome measures. The main research questions are whether phonological awareness skills, EF skills and rhythmic synchronisation abilities will predict response to intervention with GG Rime. Secondary research questions are whether boys may benefit from this CARI more than girls, and whether vocabulary and non-verbal intelligence will also be associated with response to intervention with GG Rime. Finally, we ask whether the selected neurocognitive predictors are predictive of progress in literacy for the whole cohort of 391 children.

MATERIALS AND METHODS

Design

The RCT (Worth et al., 2018) was a two-armed pupil-RCT, carried out over two consecutive school years with a relatively large number of pupils (398 randomised; 391 in the final analyses presented here). The data presented here include primary outcome data for pupils from all 15 schools involved in the trial. Less than ten per cent of participating pupils had missing data and the independent evaluators considered that this attrition was likely to be unbiased (page 5, Worth et al., 2018). Training, technical support and some delivery support (e.g., fixing school firewall problems) for GG Rime was provided by the current authors. In addition to the primary outcome measures (NGRT, SWST), a process evaluation used case-study visits, telephone interviews and analysis of data on pupils' usage of the game to capture the perceptions and experiences of participating teaching staff and pupils (Worth et al., 2018). All children gave their assent prior to testing, and the study was reviewed by the Psychology Research Ethics Committee of the University of Cambridge.

Participants

Three hundred and ninety eight Year 2 children aged between 6 and 7 years old participated in the study, all of whom were eligible for inclusion because they had failed the Phonics Check at the end of Year 1 (scoring 31 or less, a cut-off decided by the independent evaluators). This threshold was chosen to target the programme at struggling readers and to ensure that a consistent selection threshold was applied across all the schools involved. Some children who failed the Phonics Check had EAL. At the end of the study, following some movement and drop-out, data from 391 children were available for analysis. Pretest data for the GG Rime children ($N = 195$) and the control BAU group ($N = 196$) are presented in **Table 1**, which shows that the two groups did not differ significantly on any measure prior to receiving the intervention. Although the non-word reading task (TOWRE PDE) approached significance ($p = 0.052$), the advantage was for the BAU group. Allocation of the children to either the GG Rime or control groups was carried out by the independent evaluator to ensure full randomisation. As EAL was not considered as a factor during randomisation, there were slightly more EAL children in the BAU group (GG Rime, $N = 48$; BAU, $N = 60$). However, their mean Phonic Check scores were comparable

TABLE 1 | Group characteristics expressed as mean and (S.D.) for GG Rime children and the "Business As Usual" control group.

	GG Rime	BAU control	$t(1,194)$	p
N	195	196		
Age (years; months)	6;7 (3.2)	6;7 (3.4)	0.5	0.64
Phonics Check Y1	17.9 (8.8)	18.9 (9.3)	1.1	0.26
WISC Blocks	8.7 (2.7)	8.8 (3.1)	0.3	0.77
WISC Similarities	9.2 (3.20	9.2 (3.1)	0.03	0.98
BPVS Vocabulary	88.3 (10.6)	90.0 (10.5)	1.5	0.14
PhAB Rhyme	89.9 (11.8)	88.6 (12.0)	1.1	0.29
PhAB Phoneme	91.0 (9.1)	92.3 (9.2)	1.4	0.17
WISC DSF	6.2 (1.8)	6.3 (1.7)	1.0	0.31
WISC DSB	3.7 (1.8)	4.0 (1.7)	1.7	0.10
H&F Congruent	69.9 (24)	70.4 (23)	0.2	0.84
H&F Incongruent	38.8 (24)	41.5 (23)	1.1	0.26
Sync Score Trial 1	2.9 (1.4)	3.0 (1.5)	0.7	0.49
Sync Score Trial 2	3.0 (1.4)	3.1 (1.4)	0.3	0.79
NGRT pretest raw	8.6 (4.3)	8.9 (5.2)	0.6	0.53
TOWRE SWE pretest raw	18.11 (10.4)	18.9 (11.0)	0.7	0.49
TOWRE PDE pretest raw	8.2 (4.8)	9.2 (5.3)	2.0	0.052

WISC, Wechsler Intelligence Scale for Children, standardised mean score = 10; BPVS, British Picture Vocabulary Scales, standardised mean score = 100; PhAB, Phonological Awareness Battery, standardised mean score = 100; DSF, Digit Span Forwards, raw score; DSB, Digit Span Backwards, raw score; H&F Congruent, Hearts and Flowers task, % correct in the Congruent condition; H&F Incongruent, Hearts and Flowers task, % correct in the Incongruent condition; Sync Score, tapping synchronisation score out of 6; NGRT, raw score on National Group Reading Test; SWE, raw score on the TOWRE Sight Word Efficiency Scale; PDE, raw score on the TOWRE Phonic Decoding Efficiency Scale.

(GG Rime = 17.9; BAU = 18.5). Pre-testing with the selected baseline and neurocognitive predictor measures described below was carried out blind by the current authors, prior to being informed of the randomisation.

Pre-intervention Baseline Assessments

A series of baseline measures were administered to the entire cohort prior to the intervention commencing. These assessments were given over the Autumn term of the 3-term United Kingdom school year, and measured the children's basic verbal and non-verbal cognitive skills, receptive vocabulary, phonological memory and reading.

(i) WISC Similarities Test

The Similarities task is one of the core sub-tests of the WISC IV (Wechsler Intelligence Scales for Children, Wechsler, 2016) and is a measure of verbal intelligence. The child is given two words orally that name common objects or concepts and is asked to say how they are similar. There are 23 items and the maximum raw score is 44. A scaled score between 1 and 19 can then be computed, with the mean scaled score being 10.

(ii) WISC Block Design Test

The Block Design task is a measure of non-verbal intelligence. For this task a child must recreate, within a specified time limit, the design of a hand-made model or picture in a stimulus book, using red and white blocks. There are 14 items with a total maximum score of 68. A scaled score between 1 and 19 can then be computed, with the mean scaled score being 10.

(iii) WISC Digit Span Forwards Test

The Digit Span task has both a forwards and a backwards component, and the forwards component measures phonological short-term memory. For the Digit Span Forwards (DSF) task, the child repeats numbers in the same order as they are said aloud by an administrator. The task comprises pairs of trials at different span lengths, with a total of eight trials. Hence the maximum score is 16.

(iv) British Picture Vocabulary Scales

The child views sets of four pictures provided in a stimulus book, and on each trial the administrator names one picture. The child must point to the correct match. The maximum raw score is 168 and a standard score can be computed with a mean of 100, S.D. 15 (Dunn et al., 2009).

(v) Reading

The children completed two standardised assessments of reading. The first was the New Group Reading Test Level IB (NGRT, GL Assessment). This is an untimed multiple choice test with three sub-sections, Phonics, Sentence Comprehension and Passage Comprehension. In the Phonics section (15 items), children find the word which rhymes with a target from a multiple choice selection, for example selecting 'ocean' to rhyme with 'motion' (both irregular spellings), or complete word endings and word beginnings by ticking the stem which best completes a target word, for example selecting 'al' to complete the stem 're' (to make 'real,' thereby artificially dividing a vowel digraph). In the sentence completion section (18 items), children read sentences and then choose the word which best fits a gap in the sentence ('She put the book – [under] her bed'). In the passage comprehension section (10 items), children read a passage independently and then answer multiple choice questions. There are five alternative choices for every item in each section of the NGRT, hence chance responding is 20% (8.6 items). Children are not required to read aloud when completing the NGRT. The second standardised measure was the Test of Word Reading Efficiency (TOWRE, Torgesen et al., 1999). The TOWRE consists of two subtests measuring speeded decoding of words (SWE, Sight Word Efficiency) and non-words (PDE, Phonetic Decoding Efficiency). In each case, children are required to read aloud from a list of items graded in difficulty as many items as possible in 45 s, as quickly and as accurately as they can.

Pre-intervention Neurocognitive Assessments

A series of hypothesis-driven neurocognitive predictor measures were administered to the entire cohort prior to the intervention commencing. Assessments intended to measure three broad factors were administered, namely phonological awareness (rhyme and phoneme levels), executive function/attention (WISC backwards digit span, Hearts and Flowers task) and motor rhythmic synchronisation to the beat (a tapping measure administered on a computer).

(i) Phonological Awareness

Two tests from the United Kingdom standardised Phonological Assessment Battery (PhAB) were used to assess phonological awareness skills in the children (Frederickson et al., 1997; GL Assessment).

a. The *Rhyme* Test was used to assess the children's ability to identify the rime in single syllable words. Participants listened to an administrator say three words and then chose which two of the three words ended with the same shared sound (e.g., big, *hiss, miss*). There were three practice items, followed by 12 items in Part 1 of the test and 9 items in Part 2. If a child did not answer nine or more items correctly in Part 1 they did not proceed to Part 2. The total possible maximum score was 21. A standard score can be computed with a mean of 100 and S.D. 15.
b. The *Spoonerisms* Test was used to make an assessment of the children's ability to isolate phonemes in single syllable words and then recombine them to make new words. In Part 1 of the test the children were asked to replace the first sound of a word with a new sound, in order to make a new word. For example 'cat' with a 'f' makes 'fat.' In Part 2 of the test the children were asked to exchange the first sounds of two words to make two new words. For example 'lazy dog' makes 'daisy log.' A score was given for each one of the two words they correctly produced on each item. Parts 1 and 2 began with three practice items, followed by 10 test items, with a time limit of 3 min. Children of all ages attempted Part 1, but only those aged 7 or older carried on to Part 2 of the assessment. For these children the total maximum score was 30 (Part 1 maximum score of 10 + Part 2 maximum score of 20). A standard score can be computed with a mean of 100 and S.D. 15.

(ii) Executive Function/Attention

Two tasks were used to assess executive function and attention skills.

a. WISC *Digit Span Backwards* Test. For the Digit Span Backwards (DSB) task, the child repeats numbers in the reverse order to that read aloud by the administrator. The DSB task contains eight pairs of trials of increasing span length, enabling a total score of 16. The DSB task has proved to be a useful measure of executive function in young children (Lipsey et al., 2017).
b. *Hearts and Flowers* Task (Davidson et al., 2006). This is a graded and computerised executive function task measuring cognitive flexibility and inhibitory control. The game is presented on a laptop with two buttons (1,0) set up for recording the child's responses. The stimuli are presented trial-by-trial to the left or right of the screen in the shape of either a heart or a flower icon. The appearance of a heart indicates that the child should make a response on the same side as the heart icon. The appearance of a flower indicates that the child should respond on the opposite side to the flower icon. The stimuli are represented for 750 ms and the inter-stimulus intervals were set at 500 ms. Each child attempted three

blocks of trials, congruent, incongruent and mixed. Each block began with a set of instructions and a short set of practice items, followed by 20 experimental trials. The initial congruent block consisted of heart trials, in which the heart icon appeared to the left or right of the screen, and all responses were to the same side as the picture. The second incongruent block consisted of flower trials, in which the flower icon appeared to the left or right of the screen, and all the responses were on the opposite side to the picture. The final block was a mixed one, where congruent (heart) and incongruent (flower) icons both appeared to the left or right of the screen, requiring the child to remember the rules of the game (heart = same side and flower = opposite side) and respond accordingly. For the current sample, the mixed task at this ISI proved too difficult, with very few children able to answer before the next trial commenced. Mean performance for the GG Rime group was only 24% correct (S.D. 15%), with 22% correct (S.D. 14%) for BAU. Hence we do not report the mixed trial data further.

(iii) Motor Synchronisation to a Beat

This task was included as prior research has suggested that it may provide a simple index of individual differences in neural entrainment to rhythmic inputs (Colling et al., 2017) and that it is a predictor of reading readiness (Woodruff Carr et al., 2014). Using a laptop and headphones, children were asked to tap along to a metronome beat using the spacebar at a tempo of 120 bpm (2 Hz) for 30 s. Responses were recorded automatically. The task was given twice to enable familiarity. Before each trial, children were instructed that they would hear a rhythmic beat, and that they should tap along to the beat until it ended. They were told not to worry if they missed a beat and to keep on tapping. They were asked to press 'ENTER' when they were ready to begin each trial. The stimuli were programmed using Audacity software and Presentation software was used to capture the timing of children's taps. For scoring, the first three taps and the final tap of each trial were discarded. Following our prior work with this task (Cumming et al., 2015), a synchronisation score was then computed for each child, with the range of possible scores being from 0 (very poor synchronisation) to 6 (very good synchronisation).

Post-intervention Assessments

Following the intervention with GG Rime, which was given over the Spring term of the 3-term United Kingdom school year, reading and spelling were assessed during the Summer Term by both the independent evaluators and by the current authors.

Reading

The independent evaluators administered the New Group Reading Test Level IB. The authors re-administered the TOWRE, using both the SWE and the PDE sub-scales.

Spelling

The Single Word Spelling Test (GL Assessment) was administered post-intervention only, by the independent evaluators. This was a spelling to dictation task which was untimed. The items on the test begin with relatively simple words like 'on,' 'it,' and 'up,' and then progress to more difficult words like 'shout' and 'team.'

GraphoGame Rime Intervention

GraphoGame Rime was administered during the Spring term of the United Kingdom school year, which runs from mid-January to mid-April. GG Rime is now a gaming App available from a Finnish educational technology company, Grapho Group Oy. The RCT assessed outcomes for pupils who were intended to spend 10–15 min each day for 12 weeks of the school Spring term playing GG Rime on a computer during literacy lessons in a quiet corner of the classroom. Their progress in reading was compared to BAU pupils from the same classes who received direct literacy tuition during these lessons. GG Rime provides highly repetitive and individualised intervention aimed at developing phonics skills in young learners. The game is based on the intrasyllabic unit of the rime, comprising the vowel sound and any subsequent consonants (e.g., st – AMP; cl – OCK), thought to be an important psycholinguistic unit for English-speaking children (Treiman et al., 1995). The player hears auditory targets consisting of either sounds or words and has to match these auditory targets to visual targets (letters and sequences of letters) displayed on the screen of a computer, tablet or mobile phone. The letters and letter sequences are displayed as part of different games played by the child's avatar, for example popping balloons, or helping frogs catch bugs. Children progress through a series of graduated game streams (total streams 25), each of which has multiple levels (ranging from 5 to 9 levels). To keep motivation levels high, children are rewarded with tokens at the end of each level within a stream, which they save up and then spend in a "shop." The shop sells kit for their avatar. There are also word formation games to encourage spelling skills, in which children are presented with boxes containing letters or onset and rime patterns (onsets correspond to any phonemes before the vowel in a syllable) and are asked to put them into the correct order to spell target words (e.g., c – at). As the game is adaptive, the exact letters and letter sequences practiced by different players will vary depending on speed of progression through the game. Overall the game teaches grapheme-phoneme correspondences, but using methods based on rhyme families.

GraphoGame Rime uses a success criterion of at least 80% for each level before children are able to move onto the next level. If a child fails to achieve 80% accuracy on a level, they are given individualised extra training levels in which the computer automatically selects targets that the child knew and contrasts them with targets that the child did not know. The words for GG Rime were recorded by a female speaker who had a British accent. The teaching sequence in GG Rime is based primarily on orthographic rime units. Children are introduced to single letter-sound correspondences (e.g., C, A, T, N), which are then blended into orthographic rime units (-at, -an), and then into CVC words (c-at, c-an). For example, in Stream 1 a small set of seven single phonemes and graphemes are introduced (C, S, A, T, P, I, N), and the children are told "Let's put these sounds together to make rime units." The children are then told "Now let's put another sound in front of the rime units you have just played with," and CVC words like *cat* and *tin* are created by showing blending of *c+at* and *t+in*. The children are also reinforced on the grapheme-phoneme correspondences (GPCs) in these CVC words ("The sounds in *tin* are *t*, *i*, *n*"). Subsequently, orthographic rimes that are not also real words are created, like *op* and *ap*, enabling creation of CVC words like *top* and *cap*. So the primary teaching sequence is to show a child some GPCs ("sounds"), to blend these GPCs into rimes, to blend onsets onto these rimes to create words (the term "onset" was not used in the game, onsets were called "sounds"), and then to segment the words back into GPCs.

The use of rhyme families enables GG Rime to highlight the higher-level statistical consistencies in the English orthography that are present when GPCs are considered in the context of the orthographic rime unit. The rhyme family format means that in GG Rime, GPC information is always linked to oral rhyming patterns (hence rhyme awareness is trained at the same time as phoneme awareness). Rhyme families are not taught exhaustively, rather 4 – 8 members of a particular family are introduced, and the child is then left to infer for herself that words with analogous orthographic rimes that might be subsequently encountered during classroom reading and spelling activities would be similar. The streams in the game begin with CVC items from the most consistent and most densely populated rime phonological neighbourhoods of English (De Cara and Goswami, 2002), taking into account word frequency and orthographic consistency. Later streams introduce CCVC and CVCC words (e.g., 'bring,' 'sting,' Stream 7; 'best,' 'quest,' Streams 8–10).

BAU Interventions

As the independent assessors carried out the group assignments, we did not have any input into the phonics tuition methods offered to the BAU children by classroom teachers or teaching assistants during Literacy Hour. We did ask the schools to inform us concerning the teaching methods in use, and received a wide range of responses (28 methods), including Reading Recovery, One-to-One Reading, Guided Reading and Writing, Acceleread and Literacy A-Z. Methodologically, the key point is that the chosen method/s of phonics tuition in each school were delivered for matched lengths of time to GG Rime.

Fidelity to the Intervention

Fidelity to the GG Rime intervention programme was measured by the Finnish GraphoGame team, who provided detailed logs including the time spent by each participant in playing GraphoGame and their progress through the game. The gaming log feature enables individualised assessment of learning, and is intended to allow the teacher to identify streams in the game which are causing difficulty and to decide whether to provide extra (game-based or non-game) reinforcement. The log provides a measure both of time spent on the computer, and "active trial time," the portion of that logged period on the computer when the child is actually responding to the gaming challenges rather than listening to instructions or checking out the shop for their avatar. Accordingly, active trial time was used as a control measure for the statistical analyses.

RESULTS

Descriptive Statistics

Potential pre-existing group differences between the intervention group and the BAU control group for the baseline and neurocognitive measures were explored using independent samples *t*-tests. The group comparison methods used by Worth et al. (2018) were replicated exactly, in order to enable direct comparison between their report and the current report. To create a numerical scale to measure response to intervention with GG Rime, the method adopted by Ahmed et al. (2020) was followed. The 25 streams of phonic knowledge taught by the game and their constituent levels were given numerical values ranging from 0 (game not begun) to a maximum of 181 (game completed). The mean progression point reached by the whole GG Rime group ($N = 195$) was 114.5, which corresponds to Stream 16, level 5, just over half-way through the game. The range of scores ran from 11 to 181, s.d. 46.1. To analyse predictive relations between the neurocognitive measures and response to intervention with GG Rime, exploratory correlation and multiple regression analyses were carried out. Following Worth et al. (2018), raw scores were used for all outcome analyses involving the reading and spelling measures, while standard scores were used for the phonological awareness and cognitive measures administered by the current authors. Raw scores were used for the WISC DSF and DSB measures, as a standard score can only be generated if the two measures are combined. The Hearts and Flowers measures were scored as % correct. Following a series of initial exploratory longitudinal multiple regression analyses controlling for active trial time and non-verbal IQ, we created further exploratory multiple regression equations including the child's score on the Phonics Check prior to commencement of the study as the autoregressor (see Boets et al., 2011). In order to estimate the independent effect of a longitudinal predictor such as phonological awareness on growth in phonic decoding skills, the autoregressive effect of prior responsiveness to tuition in phonic decoding skills must be controlled. The Phonics Check score provides an index of pre-existing differences in phonics learning, against which the role of the different neurocognitive measures of interest regarding response to the intervention can be compared. All statistical analyses were carried out using IBM SPSS Statistics versions 25 and 27.

GG Rime Players vs. BAU Controls: Pre-test Performance

To ascertain whether there were any pre-existing differences between the two groups prior to the commencement of the intervention, and following Worth et al. (2018), independent samples *t*-tests (uncorrected) were used. The data are shown in **Table 1**. Inspection of the table shows that the two groups were matched for their performance on the neurocognitive predictors administered prior to the commencement of the intervention. They were also matched on reading ability and on the Phonics Check. However, inspection of the baseline scores in **Table 1** shows that the children were of relatively low ability, with standard scores consistently below the expected population mean. Furthermore, both groups were performing at chance levels on the NGRT at pre-test, scoring on average nine items (chance = 8.6 items, GG Rime group, $t[1,94] = 0.7$; Control group, $t[1,196] = 0.8$). As the TOWRE is a timed task, there is no chance level performance.

Time-Lagged Relations Between the Neurocognitive Predictors and Progression Through GG Rime

The primary research questions were whether phonological awareness skills, EF skills and rhythmic synchronisation abilities would predict response to intervention with GG Rime. Also of interest was whether vocabulary and non-verbal intelligence skills would be associated with response to intervention with GG Rime. **Table 2** (column 1) shows the time-lagged relations between the predictor variables measured at pre-test and progression through the game for the 195 children receiving the GG Rime intervention. Non-parametric correlations by ranks were computed as not all variables were normally distributed. Inspection of the first column of **Table 2** shows significant time-lagged correlations between progression through the game and all the neurocognitive measures of interest, with the exception of performance in the first rhythmic synchronisation trial ($r = -0.02$) and the Incongruent Hearts and Flowers measure ($r = 0.12$). Children's vocabulary development (BPVS, $r = 0.12$) and phonological memory (DSF, $r = 0.11$) also showed no association with progression through the game. As some of the children receiving GG Rime had English as an additional language (EAL, $N = 48$), the correlation for BPVS was also run excluding these children. There was still no significant correlation between response to the GG Rime intervention and vocabulary development ($r = 0.09$). Regarding the neurocognitive predictors, both phonological awareness measures showed significant relations with progression through the game, as did the other EF measures (DSB, H&F Congruent). The relation between the second synchronisation score and response to the intervention was also significant. The significant correlation indicates that children who have better rhythmic synchronisation on the second run of the tapping task show a better response to a phonological intervention ($r = 0.17$, $p < 0.05$).

Regarding response to the GG Rime intervention and progress in reading and spelling (see **Table 2**, bottom row), significant relations between progression through the game and post-test performance were found for both reading and spelling outcome measures (SWE, 0.27; PDE, 0.28, both p's < 0.001; SWS, 0.23, $p < 0.01$). The exception was the NGRT measure, which showed no significant relation with progression through the game ($r = 0.13$).

TABLE 2 | Time-lagged Spearman's correlations between the neurocognitive predictors measured prior to intervention and (a) the progression measure for GG Rime children ($N = 195$), and (b) the post-intervention literacy measures ($N = 391$).

	(a) Progression	(b) NGRT	(b) SWE	(b) PDE	(b) SWS
PhAB Rhyme SS	0.16*	0.35***	0.34***	0.29***	0.32***
PhAB Phoneme SS	0.23**	0.45***	0.43***	0.40***	0.44***
DSB	0.27***	0.33***	0.30***	0.35***	0.30***
DSF	0.11	0.20***	0.19***	0.21***	0.17***
H&F Congruent	0.18*	0.17**	0.19***	0.22***	0.20***
H&F Incongruent	0.12	0.15**	0.08	0.15**	0.10
Sync Score 1	−0.02	0.19***	0.15**	0.10*	0.16**
Sync Score 2	0.17*	0.16**	0.18***	0.18***	0.16**
BPVS Vocabulary SS	0.12	0.19**	0.10	0.06	0.11
WISC Blocks	0.18*	0.22***	0.08	0.12*	0.11*
WISC Similarities	0.17*	0.33***	0.32***	0.23***	0.27***
Progression		0.13	0.27***	0.28***	0.23**

Progression, progression through GG Rime; NGRT, post-test raw score on National Group Reading Test; SWE, post-test raw score on the TOWRE Sight Word Efficiency Scale; PDE, post-test raw score on the TOWRE Phonic Decoding Efficiency Scale; SWS, post-test raw score on the Single Word Spelling Test; DSB, Digit Span Backwards raw score; DSF, Digit Span Forwards raw score; H&F Congruent, Hearts and Flowers task, % correct in the Congruent condition; H&F Incongruent, Hearts and Flowers task, % correct in the Incongruent condition.
*$p < 0.05$, **$p < 0.01$; ***$p < 0.001$.

Time-Lagged Relations Between the Neurocognitive Predictors and Literacy Outcomes for GG Rime and BAU Children Combined

It was also of interest to ascertain whether the selected neurocognitive predictors were associated with progress in literacy for the whole cohort of 391 children. Accordingly, relations between the neurocognitive predictor and baseline measures and reading and spelling development for the full cohort ($N = 391$) are shown in columns 2–5 of **Table 2**. As would be expected on the basis of the prior literature, individual differences in phonological awareness, rhythmic synchronisation (both Trial 1 and Trial 2 measures) and phonological memory (DSF) were significantly associated with literacy outcomes for the whole cohort of children over the year of the study. However, individual differences in receptive vocabulary as measured by the BPVS did not show significant time-lagged relations with literacy, with the exception of progress on the NGRT (the only multiple choice measure). To check whether this null result for vocabulary was due to the relatively large number of EAL children failing the Phonics Check in the full sample ($N = 108$), the correlations for BPVS were recomputed taking only those children in the full sample who did not have EAL ($N = 283$). When EAL children were excluded, then each literacy outcome measure was significantly associated with vocabulary development (NGRT, $r = 0.28$, $p < 0.001$; SWE, $r = 0.17$, $p < 0.01$; PDE, $r = 0.15$, $p < 0.05$; SWS, $r = 0.19$, $p < 0.01$). Overall, therefore, literacy outcomes were associated with pre-existing vocabulary skills for children with English as their first language.

Multiple Regression Analyses Comparing the Strength of the Neurocognitive Predictors With Respect to Progression Through GG Rime

As hypothesised therefore, phonological awareness skills, EF skills and rhythmic synchronisation abilities all showed time-lagged associations with response to intervention with GG Rime. To analyse the strength of these predictive relations, a series of exploratory multiple regression analyses were then carried out. Following the significant associations between progression through the game and active trial time ($r = 0.47$, $p < 0.001$), and progression through the game and IQ (shown in **Table 2**), active trial time and non-verbal IQ were controlled in these exploratory analyses. In each case the dependent variable was progression through the game, and a fixed entry method was used, so that individual differences in active trial time were controlled at Step 1 and non-verbal cognitive ability (WISC Blocks) controlled at Step 2, before the strength of the different neurocognitive predictors was explored. Cooks Distance scores were always less than 0.065. The data are shown in **Table 3A**. Inspection of the table shows that once individual differences in active trial time and non-verbal cognitive ability had been controlled, all the neurocognitive predictors except for the first time that children tapped to the beat (Sync Score 1) accounted for significant unique variance in progression through GG Rime. The data show that each neurocognitive measure was a unique predictor. Gender also approached significance in these analyses ($p = 0.05$).

To compare the predictive strength of the different neurocognitive predictors with respect to each other regarding response to intervention with GG Rime, and to check for interactions between them, the measure accounting for the most unique variance in each case was selected (phoneme awareness, DSB and Sync Score 2). Six 5-step fixed entry exploratory multiple regression equations were then computed. The dependent variable was always progression through the game. Each equation entered active trial time and non-verbal cognitive ability at Steps 1 and 2, and then entered either (a) DSB at Step 3 and PhAB phoneme awareness at Step 4; or (b) DSB at Step 3 and Sync Score 2 at Step 4; or (c) phoneme awareness at Step 3 and DSB at Step 4; or (d) phoneme awareness at Step 3 and Sync Score 2 at Step 4; or (e) Sync Score 2 at Step 3 and DSB at Step 4; or (f) Sync Score 2 at Step 3 and phoneme awareness at Step 4. Finally, in each case the interaction between the two selected neurocognitive measures was entered at Step 5. If the interaction term were significant, this would suggest that EF, rhythmic synchronisation and phonological awareness abilities affect each other, which in turn affect children's response to GG Rime. For example, having good EF skills may compensate for having poor phonological awareness.

These further equations showed that EF, phoneme awareness and rhythmic synchronisation did not interact with each other. Further, all three measures remained significant predictors of progression through the game after controlling for the other neurocognitive predictors (**Tables 3B,C** for examples). The single exception was rhythmic synchronisation once DSB was controlled (equation [b] above). Accordingly, it can be concluded that pre-existing phonological skills and EF skills are the best predictors of children's response to intervention with GG Rime. Perhaps surprisingly, having both good EF skills and good phonological skills does not exert any extra effect on progression through the game.

TABLE 3 | Unique variance (R^2 change) in progression through GG Rime explained by the different neurocognitive predictors measured prior to the intervention.

(A)	**Standardised beta**	**R^2 change**
Step 1		
Active trial time	0.434	0.188***
Step 2		
WISC NV IQ (Blocks)	0.204	0.042**
Step 3		
PhAB Rhyme	0.287	0.073**
PhAB Phoneme	0.324	0.093***
DSB	0.241	0.054***
H&F Congruent	0.153	0.023*
H&F Incongruent	0.137	0.017*
Sync Score 1	0.021	0.000
Sync Score 2	0.157	0.024*
Gender	−0.125	0.015^{+}
(B)		
Step 1		
Active trial time	0.434	0.188***
Step 2		
WISC NV IQ (Blocks)	0.204	0.042**
Step 3		
DSB	0.241	0.054***
Step 4		
PhAB Phoneme	0.275	0.061***
Step 5		
DSB × PhAB Phoneme	0.096	0.000
(C)		
Step 1		
Active trial time	0.434	0.188***
Step 2		
WISC NV IQ (Blocks)	0.204	0.042**
Step 3		
PhAB Phoneme	0.324	0.093***
Step 4		
Sync Score 2	0.147	0.021*
Step 5		
PhAB Phoneme × Sync 2	−0.261	0.001

*DSB, Digit Span Backwards; H&F Congruent, Hearts and Flowers task, % correct in the Congruent condition; H&F Incongruent, Hearts and Flowers task, % correct in the Incongruent condition. **p < 0.01; *p < 0.05; $^{+}p = 0.05$. ***p < 0.001; **p < 0.01; *p < 0.05.*

As will be recalled, a secondary research question of interest was whether boys may benefit from the GG Rime CARI more than girls. Consistent with the hypothesis that boys may benefit more from educational technology, gender approached significance in the first set of exploratory multiple regressions reported above (p = 0.05). Accordingly, a 6-step exploratory multiple regression equation was computed controlling for active trial time and non-verbal cognitive ability at Steps 1 and 2, then controlling for the three neurocognitive predictors Sync Score 2, DSB and phoneme awareness at Steps 3, 4, and 5, and then finally entering gender at Step 6. The dependent variable was progression through the game. The results are shown in **Table 4A**. As can be seen, gender was still a significant predictor of progress through the game even after this set of very stringent controls. Accordingly, there is clear evidence that boys respond to this technological software intervention better than girls.

Multiple Regression Analyses Regarding Progression Through GG Rime Controlling for the Autoregressor

As noted earlier, a further research question addressed in the current study was whether pre-existing phonological awareness, EF and rhythmic synchronisation skills would predict response to intervention with GG Rime after participants' prior sensitivity to phonics tuition had been controlled using an autoregressor. To address this question, three further exploratory multiple regression equations were created to include the autoregressor, thereby including an estimate of children's responsiveness to tuition in phonics prior to receiving the intervention (Boets et al., 2011). Including an autoregressor isolates the specific effects of the neurocognitive predictors (phoneme awareness, rhythmic synchronisation and DSB, respectively) on response to the GG Rime phonics intervention. To control for pre-existing phonics skills, the children's Phonics Check score at the end of the first school year was used as a measure of pre-existing differences in responsiveness to tuition in phonics. The Phonics Check score was entered as Step 1 in these further equations, active trial time and non-verbal cognitive ability were Steps 2 and 3, and a neurocognitive predictor was entered at Step 4. The dependent variable was progression through the game. The three equations (**Table 4B** for the equation with phoneme awareness at Step 4) showed that both phoneme awareness and DSB (standardised $\mathcal{B}$= 0.166, p = 0.012, not in table) remained significant predictors of progression through the game after controlling for prior phonics skills, while rhythmic synchronisation did not (standardised $\mathcal{B}$= 0.106, p = 0.09, not in table). Accordingly, both EF skills and phonological awareness exert specific effects on progression through a game designed to teach children phonic skills, even when children's pre-existing responsiveness to phonics tuition is taken into account.

Comparing the Neurocognitive Predictors for the Adequate and Inadequate Responders, and Girls Versus Boys, Regarding Progression Through GG Rime

As a final assessment of response to the intervention, we analysed performance in the different neurocognitive predictor and baseline measures for those children who played above the mean progression point in the game (the 95 children scoring > 114 on the progression measure), versus those children who played below the mean progression point in the game (N = 100). This comparison provides a second method of assessing which pre-existing differences between the children who were randomly assigned to play GG Rime may be related to response to the intervention. Note that the neurocognitive predictor measures have not previously been included in prior reports of the RCT (Worth et al., 2018; Ahmed et al., 2020). If the adequate responders had better pre-existing phonological awareness, EF or rhythmic synchronisation skills than the inadequate responders, this would support the hypothesis that these are important neurocognitive skills for responding to intervention with a CARI. The data are shown in **Table 5**. Interestingly, inspection of the table shows that the two sub-groups of GG Rime children did not differ significantly in their phonological awareness prior to beginning intervention with GG Rime. They also did not differ in their EF skills as measured by the Hearts and Flowers measure, but they did differ in their EF skills as measured by DSB. Children who progressed further through the intervention had better DSB scores ($p < 0.01$). The two GG Rime sub-groups (adequate versus inadequate responders) also differed regarding their rhythmic synchronisation skills on the second tapping trial (Sync Score 2). Children with a better ability to synchronise to the beat made more progress through the game. As noted, prior work suggests that individual differences in rhythmic synchronisation

TABLE 4 | Unique variance (R^2 change) in progression through GG Rime explained by gender **(A)** and when controlling for the autoregressor Phonics Check **(B)**.

	Standardised beta	R^2 change
(A)		
Step 1		
Active trial time	0.434	0.188***
Step 2		
WISC NV IQ (Blocks)	0.204	0.042**
Step 3		
Sync Score 2	0.157	0.024*
Step 4		
DSB	0.218	0.043**
Step 5		
PhAB Phoneme	0.278	0.062***
Step 6		
Gender	0.085	0.014*
(B)		
Step 1		
Phonics Check	0.235	0.055**
Step 2		
Active trial time	0.468	0.216***
Step 3		
WISC NV IQ (Blocks)	0.184	0.034**
Step 4		
PhAB Phoneme	0.256	0.051***

*DSB, Digit Span Backwards; PhAB Phoneme, Phonological Assessment Battery phoneme task. **p < 0.01; *p < 0.05; ***p < 0.001; **p < 0.01; *p < 0.05.*

TABLE 5 | Means and standard deviations in parentheses for group characteristics prior to the intervention for (a) GG Rime players who were adequate versus inadequate responders as measured by their progression through the game, and (b) boys and girls.

	(a) GG Rime adequate	(a) GG Rime inadequate	(b) Boys	(b) Girls
N	95	100	117	78
Age (years; months)	6;7 (3.2)	6;7 (3.4)	6;7 (3.3)	6;7 (3.5)
Phonics Check	19.6 (8.6)	16.3 (8.8)**	17.6 (8.9)	18.4 (8.8)
WISC Blocks	9.2 (2.7)	8.2 (2.6)**	8.8 (2.7)	8.6 (2.7)
WISC Similarities	9.8 (3.2)	8.6 (3.1)**	9.0 (3.3)	9.5 (3.0)
BPVS Vocabulary	89.9 (11.4)	86.6 (9.5)*	89.1 (10.8)	87.2 (10.30
PhAB Rhyme	91.4 (11.9)	88.5 (11.7)	89.9 (12.0)	89.8 (11.6)
PhAB Phoneme	92.3 (9.1)	89.8 (9.0)	91.6 (9.1)	90.3 (9.1)
WISC DSF	6.4 (1.8)	6.0 (1.8)	6.2 (12.0)	6.1 (11.6)
WISC DSB	4.0 (1.7)	3.3 (1.8)**	3.6 (1.9)	3.7 (1.5)
H&F Congruent	72.7 (24.5)	67.4 (22.9)	71.7 (21.9)	67.4 (26.3)
H&F Incongruent	41.1 (25.4)	36.7 (22.1)	41.1 (24.1)	35.4 (23.1)
Sync Score T1/6	2.8 (1.4)	3.0 (1.3)	2.8 (1.4)	3.1 (1.3)
Sync Score T2/6	3.3 (1.4)	2.8 (1.3)*	2.9 (1.4)	3.2 (1.2)
Playing days	34.9 (9.8)	24.0 (6.7)**	30.1 (9.6)	28.0 (10.4)
Active trial time minutes	270.1 (88.5)	198.7 (75.9)***	240.0 (97.2)	223.1 (75.8)
Progression score (range 0–187)	155.3 (23.9)	75.7 (22.3)**	120.9 (46.0)	104.8 (44.7)*

*WISC, Wechsler Intelligence Scale for Children; BPVS, British Picture Vocabulary Scale SS; PhAB, Phonological Awareness Battery SS; DSF, Digit Span Forwards raw score; DSB, Digit Span Backwards raw score; H&F Congruent, Hearts and Flowers task, % correct in the Congruent condition; H&F Incongruent, Hearts and Flowers task, % correct in the Incongruent condition; Sync Score, tapping synchronisation score; T1, Trial 1; T2, Trial 2; Progression score, progression through the levels of GG Rime. **$p < 0.01$; *$p < 0.05$. ***$p < 0.001$; **$p < 0.01$; *$p < 0.05$.*

may provide an index of individual differences in neural rhythmic entrainment, and therefore contribute to individual differences in the development of phonological awareness and subsequent progress in literacy (Colling et al., 2017).

Regarding the baseline measures, **Table 5** also shows that the sub-group of children who progressed further through the intervention had significantly higher cognitive and verbal skills (higher group scores for WISC Blocks, WISC Similarities and receptive vocabulary [BPVS]). They did not, however, show better phonological memory as measured by the DSF task, but they did have significantly higher Phonics Check scores. Hence phonics ability, vocabulary and cognitive ability prior to being assigned to their RCT grouping also determined which children benefitted more from playing GG Rime. Overall, this analysis shows that the adequate responders were in general a higher-ability group than the inadequate responders. Two differences from the correlational analyses (**Table 2**) should be noted. First, the finding that pre-existing phonological skills as measured by the PhAB did not differ between the two GG Rime sub-groups, but that individual differences in phonological skills *within* the whole GG Rime cohort was the strongest predictor of response to the intervention, is consistent with the view that the relationship between response to literacy interventions and phonological awareness is a linear one. These data support the wider finding in the literature that pre-existing phonological skills are a critical factor regarding individual differences in children's ability to learn phonics (Galuschka et al., 2014). Second, pre-existing vocabulary skills did differ between the two GG Rime groups, but individual differences in vocabulary skills within the whole GG Rime cohort did not predict response to the intervention. This suggests that GG Rime offers phonic learning benefits even to children with lower language skills.

Next, the same sub-group analyses were performed by gender (boys, $N = 117$; girls, $N = 78$). This comparison enables contextualisation of the gender effect (for example, despite the random assignment, it may be that boys had significantly better pre-existing phonological skills than girls). Independent samples t-tests showed that boys made significantly more progress through the game than girls, attaining on average 120 out of the 181 progression points compared to 104 for girls ($p = 0.016$). None of the neurocognitive and baseline measures differed significantly by gender, however, and no significant differences were found for active trial time either. As girls persevered with the GG Rime game for an equal amount of time to boys, the data do not suggest that the game is less motivating for girls. Accordingly, playing a CARI seems to help boys to respond more to a literacy intervention than girls.

Multiple Regression Analyses Entering the Neurocognitive Measures as Predictors of Progress in Reading and Spelling for GG Rime Versus BAU Children

Our final research question was whether the selected neurocognitive predictors would be predictive of progress in literacy for the whole cohort of 391 children. To supplement the correlational data reported in **Table 2** for all 391 children, we ran these analyses separately for the children who received GG Rime ($N = 195$) versus BAU ($N = 196$). Four exploratory multiple regression analyses were carried out in each case, taking the NGRT raw score, the TOWRE SWE raw score, the TOWRE PDE raw score or the SWS raw score as the dependent variable, respectively. A fixed entry method was used, with individual differences in general non-verbal cognitive ability (WISC Blocks) entered first, then gender as a binary variable at step 2, then the sensory measure (rhythmic synchronisation) at step 3, an EF measure (DSB) at step 4, and then phonological awareness (the phoneme measure) at Step 5. Cooks Distance scores were less than 0.05. The results are shown in **Table 6**, which reports the unique variance accounted for by each predictor measure for the four different literacy outcomes, respectively.

Inspection of the data for the GG Rime players (top half of **Table 6**) shows that for each literacy measure the different neurocognitive predictors (rhythmic synchronisation, EF/attention skills and phonological awareness) accounted for unique variance in the development of literacy skills. There were two exceptions, both for the rhythmic synchronisation score, which did not account for significant unique variance in the SWS and the NGRT. For the BAU children, the data show that the different neurocognitive predictors accounted for significant unique variance in all of the literacy measures with one exception (rhythmic synchronisation and PDE). Interestingly, non-verbal IQ was also a significant predictor of literacy outcomes for the BAU children, for all measures. This was not the case for the GG Rime children. This pattern suggests that children's reading and spelling development will benefit from phonics tuition via GG Rime irrespective of their non-verbal cognitive abilities, an important result. Regarding gender, significant unique variance was only added to the equations for spelling, for both GG Rime and BAU children. Hence while gender is important for responding to intervention with a CARI, gender does not affect reading development, only spelling development (girls in the cohort were better spellers, with a mean standard score of 94.6 compared to 90.6 for the boys, $p < 0.001$). The largest absolute amounts of variance in all analyses for both groups were accounted for by the EF/attention measure (DSB) and the phonological awareness measure (phoneme awareness). Interestingly, EF skills accounted for more unique variance in literacy outcomes for the BAU children, while phonological awareness accounted for more unique variance in literacy outcomes for the GG Rime children.

It can be concluded that the neurocognitive predictors administered to the sample were generally significant predictors of literacy outcomes, irrespective of whether children received GG Rime or BAU. Pre-existing non-verbal cognitive skills did not affect the literacy benefits accrued from GG Rime, however, but they did affect the literacy benefits accrued from the other phonics tuition methods that formed BAU. EF skills also contributed more unique variance toward children's benefit from phonics tuition methods that were not technology-based, a surprising result. Finally, the role of vocabulary as a predictor was also explored. As will be recalled, vocabulary was significantly correlated with all the literacy outcome measures when data analysis was limited to those children in the full sample who were not EAL (those children for whom English was the first language, N = 283, see discussion of **Table 2**). Accordingly, to estimate the possible role of vocabulary in literacy development over the year, we restricted the children in each group to non-EAL children (GG Rime, N = 147; BAU, N = 136). We then ran the same regression equations for the GG Rime children and the BAU children, respectively, but entering BPVS at Step 2 instead of gender. When English was the child's first language, and the child was playing GG Rime, vocabulary was never a significant predictor of literacy outcomes (NGRT, standardised $\mathcal{B}$= 0.146, p = 0.102; SWE, standardised $\mathcal{B}$= 0.038, p = 0.674; PDE, standardised $\mathcal{B}$= 0.061,

TABLE 6 | Unique variance (R^2 change) in progression in the different literacy outcome measures taken at the end of the school year for GG Rime and BAU groups explained by the neurocognitive predictors measured at the beginning of the school year.

	NGRT	TOWRE SWE	TOWRE PDE	SWS
GG Rime				
Step 1				
WISC NV IQ (Blocks) Std $\mathcal{B}$	0.119	0.038	0.084	0.107
Unique R^2	0.014	0.001	0.007	0.011
Step 2				
Gender Std $\mathcal{B}$	0.031	0.079	−0.031	0.175
Unique R^2	0.001	0.006	0.001	0.030*
Step 3				
Sync Score 2 Std $\mathcal{B}$	0.108	0.193	0.204	0.094
Unique R^2	0.011	0.036**	0.040**	0.008
Step 4				
DSB Std $\mathcal{B}$	0.273	0.218	0.236	0.256
Unique R^2	0.068***	0.043**	0.051**	0.060**
Step 5				
Phoneme aw. Std $\mathcal{B}$	0.411	0.382	0.360	0.396
Unique R^2	0.142***	0.123***	0.109***	0.132***
BAU				
Step 1				
WISC NV IQ (Blocks) Std $\mathcal{B}$	0.281	0.198	0.157	0.180
Unique R^2	0.079***	0.039**	0.025*	0.033*
Step 2				
Gender Std $\mathcal{B}$	0.078	0.113	0.018	0.152
Unique R^2	0.006	0.013	0.000	0.023*
Step 3				
Sync Score 2 Std $\mathcal{B}$	0.178	0.152	0.080	0.167
Unique R^2	0.031**	0.023*	0.006	0.027*
Step 4				
DSB Std $\mathcal{B}$	0.302	0.382	0.352	0.358
Unique R^2	0.077***	0.123***	0.104***	0.108***
Step 5				
Phoneme aw. Std $\mathcal{B}$	0.340	0.329	0.299	0.324
Unique R^2	0.090***	0.084***	0.069***	0.082***

*Std $\mathcal{B}$, standardised beta; DSB, Digit Span Backwards; Phoneme aw, phoneme awareness; NGRT, post-test raw score on National Group Reading Test; SWE, post-test raw score on the TOWRE Sight Word Efficiency Scale; PDE, post-test raw score on the TOWRE Phonic Decoding Efficiency Scale; SWS, post-test raw score on the Single Word Spelling Test. ***p < 0.001; **p < 0.01; *p < 0.05.*

p = 0.500; and SWS, standardised $\mathcal{B}$= 0.006, p = 0.945). When English was the child's first language, and the child experienced BAU, however, vocabulary was a significant predictor of real word reading and spelling (NGRT, standardised $\mathcal{B}$= 0.230, p = 0.013; SWE, standardised $\mathcal{B}$= 0.220, p = 0.017; and SWS, standardised $\mathcal{B}$= 0.258, p = 0.006), but not of non-word reading (PDE, standardised $\mathcal{B}$= 0.090, p = 0.337). Accordingly, for children receiving phonics tuition via GG Rime, individual differences in vocabulary prior to receiving the intervention did not affect reading and spelling outcomes a year later. This provides converging evidence that even children with limited vocabularies can benefit from intervention with a CARI. By contrast, at least for the range of phonics teaching materials offered to BAU children in this study, having better cognitive ability and better vocabulary skills prior to intervention enhanced literacy outcomes.

DISCUSSION

The majority of studies exploring the best predictors of young children's response to literacy interventions have identified phonological awareness as the most important factor (Vaughn et al., 2003; Compton et al., 2006; Vellutino et al., 2006, 2008; Fletcher et al., 2011; Denton et al., 2013; Catts et al., 2015; van der Kleij et al., 2017). The current study, which explored the neurocognitive predictors of response to intervention with the CARI GG Rime, also found that individual differences in phonological awareness were the strongest predictor of response to intervention, at least in terms of the absolute amount of unique variance explained (see **Table 3A**). However, rhythmic synchronisation skills and EF skills were also significant predictors of response to intervention with GG Rime, even when controlling for phonological awareness, and so was gender. Boys showed a significantly greater response to intervention with this CARI than girls (**Tables 4A**, **5**). This is an important result. Boys and girls spent similar amounts of time actively playing GG Rime, and so it does not appear that this CARI is less engaging for girls. Indeed, the game content was designed to be equally motivating for girls and boys. Nor did boys differ from girls regarding any neurocognitive predictor. Yet the boys who played GG Rime progressed significantly further through the game, thereby receiving a greater exposure to the phonics content. To our knowledge, this is the first demonstration of a response to a reading intervention showing greater effects for boys. Given that boys are more likely to have reading disabilities, this is a very important finding.

When exploring specific effects of neurocognitive predictors, it is important to control for the child's previous response to literacy instruction before computing the effects of different factors on children's current response to literacy interventions. This was achieved here for intervention with GG Rime by using the child's Phonics Check score at the end of the previous school year as the autoregressor (**Table 4B**). Both phonological awareness and EF skills remained significant predictors of children's progression through GG Rime when the autoregressor of prior phonic learning skills was controlled. Interestingly, the children allocated to play GG Rime who were adequate versus inadequate responders in terms of progression through the game did not differ in their phonological awareness skills at the group level, although they did differ in their EF (DSB) and rhythmic synchronisation (Sync Score 2) skills at the group level (see **Table 5**). Nevertheless, analyses showed that it was the level of an individual child's neurocognitive skills that governed their progression through the game (**Tables 3**, **4**). Accordingly, the current study replicates the previous findings of non-CARI studies in finding a linear relationship between a child's phonological awareness skills and their ability to respond to a literacy intervention. However, in contrast to the previous findings reported by Miciak et al. (2019) for EF, here EF measures did determine response to literacy intervention with a CARI. This may reflect the fact that the intervention was technology-based, since EF skills such as attention and cognitive flexibility are required for efficient gaming. Further, EF skills were unique predictors of response to intervention with GG Rime, as the interaction terms in the regression equations (which enable assessment of whether, for example, poor EF skills can be compensated for by good phonological skills) were never significant. EF skills are known to be associated with reading outcomes in the classroom (Yeniad et al., 2013). Consistent with this finding, in the current study, EF skills predicted literacy outcomes for the BAU children (**Table 6**).

The current study also showed little relationship between a child's vocabulary and their ability to progress through GG Rime. Vocabulary development was not associated with response to the intervention for either the whole sample of GG Rime players, nor for those who did not have EAL, nor indeed for those who were EAL children. Nevertheless, the sub-groups analyses for GG Rime presented in **Table 5** showed that the adequate responders had higher standardised vocabulary scores (89.9 versus 86.6 for the inadequate responders). While a few studies have found significant effects of vocabulary as a predictor of response to literacy interventions (Compton et al., 2006; Vellutino et al., 2006), many other studies have not found vocabulary to be a significant predictor. This finding is important regarding the real-world use of CARIs such as GG Rime. The data showed that even children with limited vocabularies derived some benefit from playing GG Rime. As will be recalled, children were entered into the RCT on the basis of achieving a low score on the United Kingdom Phonics Check at age 5 to 6 years. Low scores may be expected if English is not the child's first language. Some participants in the study had only recently arrived in the United Kingdom. Yet to their teachers' astonishment, some of these children, who had little spoken English as far as their teachers were aware, took to GG Rime quickly and were able to play successfully through a number of levels. While none of the EAL children were in the "top half" of responders to the game, the data show that it is nevertheless educationally valuable for these children to play a CARI like GG Rime. Indeed, GG Rime has also been shown to promote English phonics learning in a study in India, where none of the children had English as a first language (Patel et al., 2018).

Our final research question was whether the selected neurocognitive predictors would be predictive of literacy outcomes for the whole cohort of 391 children. The data in **Table 6** shows that the neurocognitive predictors chosen here were appropriate to our research questions, as analyses with both the GG Rime and BAU children, respectively, revealed that all three types of neurocognitive predictor showed significant relations with reading and spelling outcomes in the end-of-year analyses. Rhythmic synchronisation accounted for significant but small amounts of unique variance for most literacy outcome measures, while both EF and phonological awareness (DSB and phoneme awareness measure) accounted for significant and larger amounts of unique variance in all the literacy outcome measures (see **Table 6**). This replicates the prior literature (e.g., Bus and van Ijzendoorn, 1999; Yeniad et al., 2013; Bonacina et al., 2018). By contrast, having better vocabulary skills and better non-verbal IQ prior to intervention only enhanced literacy outcomes for the BAU children, who received a broad range of different phonics interventions. For children who received GG Rime, having better vocabulary skills and better non-verbal IQ prior to intervention did not enhance literacy outcomes. This suggests that children with low vocabulary and low cognitive abilities still derive benefit in terms of literacy outcomes from playing GG Rime.

Accordingly, the current study provides useful information regarding the potential educational value of supplementing initial literacy teaching about phonics with educational technology via gaming Apps such as GG Rime. The data show that solo gaming with GG Rime can be very beneficial for some children's literacy outcomes, especially boys, and that children with low vocabulary skills also derive benefit from the game. High-quality educational technology can thus enable cost-effective learning for children who are struggling with phonics (Richardson and Lyytinen, 2014; Ahmed et al., 2020). GraphoGame has been developed on the basis of systematic research in over 20 languages. The Finnish and English GraphoGames are available as downloadable Apps, while research-based versions of GraphoGame are available in a range of languages including German, French, Indonesian and Chinese (Brem et al., 2010; Borleffs et al., 2017; Li et al., 2017; Ruiz et al., 2017). Regarding further development of GraphoGame, designing the learning technology in ways that benefit children with more limited EF skills could be considered.

The current study has a number of important limitations. Firstly, progress through the game is only one index of response to intervention with GG Rime, and the results could have been different if a different measure had been selected. Secondly, our participants were children who played GG Rime during the *second* year of reading tuition at school, despite GG Rime having been developed for initial reading instruction. This was an unavoidable feature of the RCT due to current United Kingdom government policy. Finally, the total intervention received was limited (a median of 3.45 h spent in active trial time), which may have introduced inadvertent bias. Very few participants were able to complete all 25 streams in the playing time available and hence to learn the majority of the phonic "rules" of English. If the children had played the game for longer and had been able to progress through more gaming streams, different outcomes may have been found. A longer trial could also enable assessment of how long children remain motivated to play a CARI. In future work, it would be optimal to ensure that children play daily until they have played their way through the entire game, and then assess the best predictors of response to the intervention, for example by using the active trial time needed by children to progress through the entire game as the dependent variable.

CONCLUSION

The current study suggests that phonological awareness, EF skills and rhythmic synchronisation skills all predict response to intervention with the GG Rime CARI. Vocabulary development did not predict response to this intervention, and boys responded better to the CARI than girls. Phoneme awareness, EF skills and the ability to synchronise finger tapping to a rhythmic beat were also significant predictors of literacy outcomes, both for the GG Rime players and for the BAU children. However, non-verbal cognitive ability was only predictive of literacy outcomes for the BAU children, suggesting that even children with lower levels of cognitive ability will benefit from playing GG Rime. Overall the data support the view that digital learning apps such as CARIs hold great promise for improving the literacy learning environments of young children (Bus et al., 2020), perhaps particularly boys. Nevertheless, the data suggest that design technology for CARIs like GG Rime may benefit from considering how to support children with low EF skills. Further, supplementary and explicit instruction in phonological awareness may enhance the benefits that can accrue from CARIs, particularly for children who have lower phonological skills to begin with.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation to any qualified researcher.

ETHICS STATEMENT

The study was reviewed by the Psychology Research Ethics Committee of the University of Cambridge.

AUTHOR CONTRIBUTIONS

AW contributed to the investigation, data curation, formal analysis, and writing – original draft. HA contributed to the investigation, data curation, and formal analysis. NM and HN contributed to the investigation and data curation. UR contributed to the software and resources. MW contributed to the conceptualisation and funding acquisition. UG contributed to the conceptualisation, methodology, funding acquisition, supervision, project administration, formal analysis, and writing –

original draft. All authors contributed to the article and approved the submitted version.

FUNDING

This research was funded by the Education Endowment Foundation/Wellcome Trust Education and Neuroscience scheme, grant number RG 78836.

ACKNOWLEDGMENTS

We thank the developing team of the "GraphoGame" at the University of Jyväskylä, Bianca Armitage for her help in scoring the standardised measures, and the Finnish support team for the RCT, especially Iivo Kapanen and Mika Halttunen. We also thank the independent evaluators (National Foundation for Educational Research) and the children, families and schools who participated in the study.

REFERENCES

Ahmed, H., Wilson, A., Mead, N., Noble, H., Richardson, U., Wolpert, M. A., et al. (2020). An evaluation of the efficacy of GraphoGame Rime for promoting English phonics knowledge in poor readers. *Front. Educ.* 5:132.

Al Otaiba, S., and Fuchs, D. (2002). Characteristics of children who are unresponsive to early literacy intervention: a review of the literature. *Remedial Spec. Educ.* 23, 300–316. doi: 10.1177/07419325020230050501

Boets, B., Vandermosten, M., Poelmans, H., Luts, H., Wouters, J., and Ghesquiere, P. (2011). Preschool impairments in auditory processing and speech perception uniquely predict future reading problems. *Res. Dev. Disabil.* 32, 560–570. doi: 10.1016/j.ridd.2010.12.020

Bonacina, S., Krizman, J., White-Schwoch, T., and Kraus, N. (2018). Clapping in time parallels literacy and calls upon overlapping neural mechanisms in pre-readers. *Ann. N. Y. Acad. Sci.* 1423, 338–348. doi: 10.1111/nyas.13704

Borleffs, E., Glatz, T. K., Daulay, D. A., Richardson, U., Zwarts, F., and Maassen, B. A. M. (2017). GraphoGame SI: the development of a technology-enhanced literacy learning tool for Standard Indonesian. *Eur. J. Psychol. Educ.* 19, 595–613. doi: 10.1007/s10212-017-0354-9

Brem, S., Bach, S., Kucian, K., Guttorm, T. K., Martin, E., Lyytinen, H., et al. (2010). Brain sensitivity to print emerges when children learn letter-speech sound correspondences. *Proc. Natl. Acad. Sci. U.S.A.* 107, 7939–7944. doi: 10.1073/pnas.0904402107

Bus, A. G., and van Ijzendoorn, M. H. (1999). Phonological awareness and early reading: a meta-analysis of experimental training studies. *J. Educ. Psychol.* 91, 403–414. doi: 10.1037/0022-0663.91.3.403

Bus, A. G., Neuman, S. B., and Roskos, K. (2020). Screens, apps and digital books for young children: the promise of multimedia. *AERA Open* 6, 1–6.

Catts, H. W., Nielsen, D. C., Bridges, M. S., Liu, Y. S., and Bontempo, D. E. (2015). Early identification of reading disabilities within an RTI framework. *J. Learn. Disabil.* 48, 281–297. doi: 10.1177/0022219413498115

Colling, L. J., Noble, H. L., and Goswami, U. (2017). Neural entrainment and sensorimotor synchronization to the beat in children with developmental dyslexia: an EEG study. *Front. Neurosci.* 11:360.

Compton, D. L., Fuchs, D., Fuchs, L. S., and Bryant, J. D. (2006). Selecting at-risk readers in first grade for early intervention: a 2-year longitudinal study of decision rules and procedures. *J. Educ. Psychol.* 98, 394–409. doi: 10.1037/0022-0663.98.2.394

Corriveau, K., Goswami, U., and Thomson, J. (2010). Auditory processing and early literacy skills in a preschool and kindergarten population. *J. Learn. Disabil.* 43, 369–382.

Cumming, R., Wilson, A., Leong, V., Colling, L. J., and Goswami, U. (2015). Awareness of rhythm patterns in speech and music in children with specific language impairments. *Front. Hum. Neurosci.* 9:672.

Davidson, M. C., Amso, A., Anderson, L. C., and Diamond, A. (2006). Development of cognitive control and executive functions from 4 to 13 years: evidence from manipulations of memory, inhibition, and task switching. *Neuropsychologia* 44, 2037–2078. doi: 10.1016/j.neuropsychologia.2006.02.006

De Cara, B., and Goswami, U. (2002). Similarity relations among spoken words: the special status of rimes in English. *Behav. Res. Methods* 34, 416–423. doi: 10.3758/bf03195470

Dellatolas, G., Watier, L., Le Normand, M.-T., Lubart, T., and Chevrie-Muller, C. (2009). Rhythm reproduction in kindergarten, reading performance at second grade, and developmental dyslexia theories. *Arch. Clin. Neuropsychol.* 24, 555–563. doi: 10.1093/arclin/acp044

Denton, C. A., Tolar, T. D., Fletcher, J. M., Barth, A. E., Vaughn, S., and Francis, D. J. (2013). Effects of Tier 3 intervention for students with persistent reading difficulties and characteristics of inadequate responders. *J. Educ. Psychol.* 105, 633–648. doi: 10.1037/a0032581

Dunn, L. M., Dunn, D. M., Whetton, and Burley. (2009). *British Picture Vocabulary Scale*, 3rd Edn. London: GL Assessment.

Fletcher, J. M., Stuebing, K. K., Barth, A. E., Denton, C. A., Cirino, P. T., Francis, D. J., et al. (2011). Cognitive correlates of inadequate response to reading intervention. *School Psychol. Rev.* 40, 3–22. doi: 10.1080/02796015.2011.12087725

Frederickson, N., Frith, U., and Reason, R. (1997). *Phonological Assessment Battery: Standardised Edition*. NFER-Nelson.

Galuschka, K., Ise, E., Krick, K., and Schulte-Korne, G. (2014). Effectiveness of treatment approaches for children and adolescents with reading disabilities: a meta-analysis of randomized controlled trials. *PLoS One* 9:e89900. doi: 10.1371/journal.pone.0089900

Goswami, U. (2011). A temporal sampling framework for developmental dyslexia. *Trends Cogn. Sci.* 15, 3–10. doi: 10.1016/j.tics.2010.10.001

Goswami, U., Huss, M., Mead, N., Fosker, T., and Verney, J. (2013). Perception of patterns of musical beat distribution in phonological developmental dyslexia: significant longitudinal relations with word reading and reading comprehension. *Cortex* 49, 1363–1376. doi: 10.1016/j.cortex.2012.05.005

Hughes, C. (1998). Executive function in preschoolers: links with theory of mind and verbal ability. *Br. J. Dev. Psychol.* 16, 233–253. doi: 10.1111/j.2044-835X.1998.tb00921.x

Huss, M., Verney, J. P., Fosker, T., Mead, N., and Goswami, U. (2011). Music, rhythm, rise time perception and developmental dyslexia: perception of musical meter predicts reading and phonology. *Cortex* 47, 674–689. doi: 10.1016/j.cortex.2010.07.010

Li, Y., Li, H., De, X., Sheng, X., Richardson, U., and Lyytinen, H. (2017). An evidence-based research on facilitating students' development of individualized learning by game-based learning-pinyin graphogame as an example. *China Educ. Technol.* 364, 95–101.

Linan-Thompson, S., Vaughn, S., Prater, K., and Cirino, P. T. (2006). The response to intervention of english language learners at risk for reading problems. *J. Learn. Disabil.* 39, 390–398. doi: 10.1177/00222194060390050201

Lipsey, M. W., Nesbitt, K. T., Farran, D. C., Dong, N., Fuhs, M. W., and Wilson, S. J. (2017). Learning-related cognitive self-regulation measures for prekindergarten children: a comparative evaluation of the educational relevance of selected measures. *J. Educ. Psychol.* 109, 1084–1102. doi: 10.1037/edu0000203

Livingstone, S., and Helsper, E. (2007). Gradations in digital inclusion; children, young people and the digital divide. *New Media Soc.* 9, 671–696. doi: 10.1177/1461444807080335

Lundetrae, K., and Thomson, J. M. (2018). Rhythm production at school entry as a predictor of poor reading and spelling at the end of first grade. *Read. Writ.* 31, 215–237. doi: 10.1007/s11145-017-9782-9

Miciak, J., Cirino, P. T., Ahmed, Y., Reid, E., and Vaughn, S. (2019). Executive functions and response to intervention: identification of students struggling with reading comprehension. *Learn. Disabil. Quar.* 42, 17–31. doi: 10.1177/0731948717749935

Nelson, J. R., Benner, G. J., and Gonzalez, J. (2003). Learner characteristics that influence the treatment effectiveness of early literacy interventions: a meta-analytic review. *Learn. Disabil. Res. Pract.* 18, 255–267. doi: 10.1111/1540-5826.00080

Ozernov-Palchik, O., Wolf, M., and Patel, A. D. (2018). Relationships between early literacy and nonlinguistic rhythmic processes in kindergarteners. *J. Exp. Child Psychol.* 167, 354–368. doi: 10.1016/j.jecp.2017.11.009

Patel, P., Torppa, M., Aro, M., Richardson, U., and Lyytinen, H. (2018). GraphoLearn India: the effectiveness of a computer-assisted reading intervention in supporting struggling readers of english. *Front. Psychol.* 9:1045.

Richardson, U., and Lyytinen, H. (2014). The GraphoGame method: the theoretical and methodological background of the technology-enhanced learning environment for learning to read. *Hum. Technol.* 10, 39–60. doi: 10.17011/ht/urn.201405281859

Rios-Lopez, P., Molinaro, N., and Lallier, M. (2019). Tapping to a beat in synchrony predicts brain print sensitivity in pre-readers. *Brain Language* 199:104693. doi: 10.1016/j.bandl.2019.104693

Ruiz, J., Lassault, J., Sprenger-Charolles, L., Richardson, U., Lyytinen, H., and Ziegler, J. (2017). GraphoGame: un outil numérique pour enfants en difficultés d'apprentissage de la lecture. *ANAE:Actualtites de Neuropsychologie de L'enfant* 148, 333–343.

Torgesen, J. K., Wagner, R. K., and Rashotte, C. A. (1999). *Test of Word Reading Efficiency*, 2nd Edn. Austin: TX: Pro-Ed.

Treiman, R., Mullennix, J., Bijeljac-Babic, R., and Richmond-Welty, E. D. (1995). The special role of rimes in the description, use, and acquisition of english orthography. *J. Exp. Psychol. Gen.* 124, 107–136. doi: 10.1037/0096-3445.124.2.107

van der Kleij, S. W., Segers, E., Groen, J. A., and Verhoeven, L. (2017). Response to intervention as a predictor of long-term reading outcomes in children with dyslexia. *Dyslexia* 23, 268–282. doi: 10.1002/dys.1562

Vaughn, S., Linan-Thompson, S., and Hickman-Davis, P. (2003). Response to treatment as a means of identifying students with reading/learning disabilities. *Exceptional Children* 69, 391–410. doi: 10.1177/001440290306900401

Vellutino, F. R., Scanlon, D. M., Small, S., and Fanuele, D. P. (2006). Response to intervention as a vehicle for distinguishing between children with and without reading disabilities: evidence for the role of kindergarten and first-grade interventions. *J. Learn. Disabil.* 39, 157–169. doi: 10.1177/00222194060390020401

Vellutino, F. R., Scanlon, D. M., Zhang, H., and Schatschneider, C. (2008). Using response to kindergarten and first-grade intervention to identify children at risk for long-term reading difficulties. *Read. Writ.* 21, 437–480. doi: 10.1007/s11145-007-9098-2

Wechsler, D. (2016). *Wechsler Intelligence Scale for Children*, 5th Edn. London, UK: Pearson Assessment.

Woodruff Carr, K., White-Schwoch, T., Tierney, A. T., Strait, D. L., and Kraus, N. (2014). Beat synchronization predicts neural speech encoding and reading readiness in pre-schoolers. *Proc. Natl. Acad. Sci. USA* 111, 14559–14564. doi: 10.1073/pnas.1406219111

Worth, J., Nelson, J., Harland, J., Bernardinelli, D., and Styles, B. (2018). *GraphoGame Rime: Evaluation Report and Executive Summary*. London: Wellcome Trust.

Yeniad, N., Malda, M., Mesman, J., van Ijzendoorn, M. H., and Pieper, S. (2013). Shifting ability predicts math and reading performance in children: a meta-analytic study. *Learn. Indiv. Differ.* 23, 1–9. doi: 10.1016/j.lindif.2012.10.004

Conflict of Interest: The University of Cambridge has licensed GraphoGame Rime to Grapho Group Oy for commercial distribution, and UG will be one beneficiary of monies received by the University.