

Language counts when learning mathematics with interactive apps

Laura A. Outhwaite, Anthea Gulliford and Nicola J. Pitchford

Laura A. Outhwaite is a Research Fellow at the Centre for Education Policy and Equalising Opportunities at UCL Institute of Education. Her research focuses on educational technology, mathematics, and evaluation methods. Anthea Gulliford is an Educational Psychologist at the University of Nottingham. Her expertise includes applied research methods, educational achievement, and bilingual learners. Nicola J. Pitchford is a Professor of Developmental Psychology at the University of Nottingham. Her expertise lies at the intersection of theory and practice in developmental psychology and educational technology. Address for correspondence: Professor Nicola J. Pitchford, School of Psychology, University of Nottingham, University Park, Nottingham NG7 2RD, United Kingdom. Email: nicola.pitchford@nottingham.ac.uk

Abstract

When available in multiple languages, educational apps can deliver the same mathematics instruction in the child's first language or different language of instruction. This pilot study examined the feasibility of a maths app intervention compared to standard mathematical practice with 61 children aged 5–6 years attending a bilingual immersion school in Brazil. The apps were delivered in either the child's first (L1, Brazilian Portuguese) or second (L2, English) language, while app content, time on task, and school setting were consistent across the two treatment groups. Time on task in the comparison group was less. After 10 weeks, results showed children made significant mathematical learning gains with the apps, for both languages of instruction, compared to standard practice: the finding is discussed in relation to the threat of time spent on task. Children using the apps in Brazilian Portuguese (L1) completed more app topics than children using the apps in English (L2) and proficiency in language of instruction correlated positively with app progress. As children's L1 was significantly stronger than their L2 proficiency, this suggests the apps were most effectively implemented in their L1. This study advances the conjoining of bilingual educational theories to educational apps in a controlled, real-world, bilingual immersion setting and highlights important directions for improving app-based mathematics instruction for bilingual children.

Introduction

Educational applications (apps), available in multiple languages and delivered on touch-screen tablet devices (eg, iPads) are increasingly prevalent in children's early learning experiences (Department for Education, 2019; Holloway, Green, & Livingstone, 2013) across a wide variety of educational and economic contexts (Drinkwater, 2013; Hubber *et al.*, 2016). To inform educational policy and practice, an evidence-base that considers factors that impact learning outcomes is imperative (Connolly, Keenan, & Urbanska, 2018; Outhwaite, Gulliford, & Pitchford, 2019). This study focuses on the influence of children's language proficiency in app-based mathematics instruction in a bilingual setting in Brazil.

Practitioner Notes

What is already known about this topic

- Educational maths apps, available in multiple languages are increasingly popular.
- Emerging evidence demonstrates the benefits of maths apps for supporting children's mathematical development.
- To understand "what works" in the use of maths apps we need to consider factors that may impact outcomes, including children's proficiency in the language of instruction.

What this paper adds

- When delivered in the child's first or second language, maths apps can support the acquisition of basic maths skills.
- To maximise engagement and learning with maths apps, children should have a sufficient level of proficiency in the language of instruction.

Implications for practice and/or policy

- When deciding to implement maths apps with young children, educational practitioners and parents should consider the individual child's proficiency in the languages spoken.

There is growing evidence demonstrating the learning benefits of different educational maths apps for young children (Berkowitz *et al.*, 2015; Schacter & Jo, 2016, 2017). When delivered in the child's local language (L1 for the majority of participating children) the apps at the focus of this study have been shown to be highly effective at supporting the acquisition of basic mathematical skills in radically different educational contexts, including the UK, a high-income Western country (Outhwaite, Faulder, Gulliford, & Pitchford, 2018) and Malawi, a low-income country in Sub-Saharan Africa (Pitchford, 2015; Pitchford, Chigeda, & Hubber, 2019). In addition, a small-scale pilot study found children with English as an Additional Language (EAL) made as much progress in mathematics as their non-EAL peers (Outhwaite, Gulliford, & Pitchford, 2017) after interacting with the maths apps delivered in English; the language of classroom instruction. The availability of the apps in these studies in different languages gives rise to the opportunity to examine the effects of language of instruction in different language contexts. To date, no studies have explicitly evaluated the influence of children's proficiency in their first (L1) and second language (L2)¹ in the context of app-based mathematics instruction. This is vital for supporting decisions about the appropriate implementation of maths apps with young children (Outhwaite *et al.*, 2019).

Bilingual education and app design theory

To examine the impact of a child's proficiency in the language of instruction in app-based mathematical learning it is important to understand the range of contextual support and the degree of cognitive involvement afforded by interactive maths apps, such as those examined in this study (see Figure S1). A child's communicative proficiency can be modelled as being related to two dimensions (Cummins, 1981). One dimension distinguishes between communication that is context embedded (e.g. inclusion of situational and paralinguistic supports, such as concrete cues and intonation) contrasting with context reduced communication (e.g. the absence of contextual or additional linguistic cues, which might occur reading a difficult text). Another dimension maps the degree to which a task presents cognitive demand for the learner, ranging from cognitively undemanding, (e.g. a situation where information can be processed relatively easily) to

tasks with high cognitive demand (e.g. involving the need to process multiple challenging pieces of information quickly).

Along the first dimension, these apps afford a relatively high degree of contextual support for children to derive meaning about the learning activity. For example, the apps provide one-to-one mathematics instruction through an on-screen teacher that gives clear visual task demonstrations and linked verbal instructions, potentially supportive of the child's development of cognitive academic language proficiency (CALP; Cummins, 2008). The apps include congruent auditory and visual information, particularly through interactive virtual objects, verbal labels, and numerical representations, engendering multisensory learning (Carr, 2012; Lindahl & Folkesson, 2012).

Along the second dimension, cognitive demand, the maths apps used in this study enable appropriate levels of cognitive demand in the learning activity. For example, the apps include clear learning objectives and simple, stepwise task instructions that children can repeat as often as desired, which may be particularly supportive when instruction is delivered in the child's L2 (see Figure S1). To complete a topic, children need to achieve 100% pass rate on each topic quiz included in the apps, securing their learning gains. The child also has an individual in-app profile, which saves their progress throughout the apps and can be monitored by teachers and the research team. Overall, these app design features may provide strong contextual support whilst optimising cognitive processing demands (Baker, 2011).

In addition, the apps at the focus of this study are consistent with other learning science principles (Hirsh-Pasek *et al.*, 2015). Meaningful learning is promoted through a stepped curriculum that builds on a child's previous knowledge (Magliaro, Lockee, & Burton, 2005) and extends learning beyond their current attainment level (Inal & Cagiltay, 2007; Vygotsky, 1978). The maths apps provide continuous assessment of knowledge acquired, engenders retrieval-based learning, known to enhance learning outcomes (Dunlosky, Rawson, Marsh, Nathan, & Willingham, 2013; Grimaldi & Karpicke, 2012). Furthermore, motivation and engaged learning (Couse & Chen, 2010) is facilitated through early reward and immediate feedback (positive or negative), which is given after every interaction with the app. Collectively, these features enable the child to regulate their pace of learning and help to create an individual and scaffolded learning environment (Slavin & Lake, 2008).

Current study

Firstly, this study examined the influence of language of instruction in an app-based mathematics learning environment in a bilingual immersion setting in Brazil for children aged 5-6 years. This study asked if app-based mathematics instruction was more effective at supporting the acquisition of basic mathematical skills when delivered in the child's L1 or L2 and examined if proficiency in language of instruction was influential on children's mathematical learning gains and progress through the apps.

Secondly, this study explored if the maths app intervention could be feasibly implemented compared to standard mathematical practice in this bilingual setting. The study asked if children made greater learning gains with the maths apps compared to standard mathematical practice, which in this context was delivered in the child's L2.

Methods

Design

This quasi-experimental study evaluated an educational maths app delivered in either the child's L1 (Brazilian Portuguese) or L2 (English)² compared to standard mathematical practice (delivered in L2) with children aged 5-6 years. Children attended a bilingual immersion private school

in Recife, Brazil. For practical reasons, a non-randomized, between-groups, design was adopted with three predetermined groups based on the three Kindergarten classes.

To address the first aim and examine if the maths apps were more effective when delivered in the child's L1 compared to their L2, the maths apps were implemented across two classes: one class received the app intervention in Brazilian Portuguese (treatment group 1) and one in English (treatment group 2). To address the second aim, a third class continued to receive mathematics teaching practice as usual, which for this school was delivered in English (L2). This formed a control group to explore feasibility of using the maths apps instead of standard practice. Group allocation was determined by senior teaching staff at the school and was based on the school time table and tablet device availability.

Children were assessed on mathematical skills, before (pre-test) and immediately after (post-test) the 10-week intervention period. Language proficiency in L1 and L2 for each child was assessed with a teacher-rated questionnaire at pre-test. Opt-in parental consent was gained for all participating children and was provided by 92% of all available children across the three classes.

Participants

Sixty two children aged 5–6 years across three Kindergarten classes took part in the study: 23 children in Class 1 were allocated to receive the maths apps in Brazilian Portuguese (treatment group 1), 20 children in Class 2 were allocated to receive the maths apps in English (treatment group 2), and 19 children in Class 3 received standard mathematics teaching practice only (control group). All children completed the pre-test (before) and post-test (after) assessments.

Exclusion criteria

To enable children to show learning gains in the region of previous research (Outhwaite *et al.*, 2017, 2018) any child with a pre-test score on the primary outcome measure that was equal to or above 75% was excluded from the analysis. Similarly, time on task has been shown to impact learning gains (Outhwaite *et al.*, 2017). Any child that missed equal to or above 20% of the intervention sessions was excluded from the analysis. According to these criteria only one child (from Group 2) was excluded from the analysis; they achieved 83% on the primary outcome measure at pre-test and was absent for 30% of the intervention sessions due to family travel. Descriptive data for the final sample structure ($n = 61$) is summarized in Table 1.

School context

The study was conducted in a Brazilian Portuguese-English bilingual immersion private school in Recife, Brazil. The school followed a one-way dual language programme with 50/50 immersion (Gomez, Freeman, & Freeman, 2005). Separate class teachers delivered L1 and L2 instruction, and time allocated to instruction in each language was split equally. In the first half of the school day all subjects, including mathematics, were taught in the child's L2. The second half of the school day focused on language skills, including reading and writing, only in the child's L1. Table S1 illustrates the standard daily school routine (see Group 3) and how the app intervention was embedded into this routine (see Group 1 & 2).

Standard mathematics practice

The school curriculum followed an enquiry-based pedagogical approach organised around interdisciplinary themes. Standard mathematics instruction was delivered in the child's L2. Mathematical concepts were integrated with other subject areas and learning activities were largely developed in whole class or small groups. For example, in an observed lesson from the control group (Class 3), the data-handling concept of sorting and matching was taught in a

Table 1: Descriptive data, including mean age in months (SD, min-max), gender (female: male) and mean L1 and L2 proficiency (SD, min-max) for the final study sample and each participating group

Descriptive data	Total (n = 61)	Group 1 (n = 23) Portuguese (L1 Maths App)	Group 2 (n = 19) English (L2 Maths App)	Group 3 (n = 19) Control (Standard Practice in English)
Age (months)	67.79 (3.49) 62.00–73.00	67.74 (3.54) 62.00–73.00	68.89 (3.14) 64.00–73.00	66.74 (3.59) 62.00–71.00
Gender (F:M)	27:34	10:13	8:11	9:10
L1 proficiency	16.07 (3.83) 8.00–21.00	15.65 (3.07) 10.00–21.00	17.16 (3.83) 8.00–21.00	15.47 (4.57) 8.00–21.00
L2 proficiency	13.31 (5.29) 0.00–21.00	13.96 (3.70) 8.00–21.00	15.53 (2.50) 12.00–19.00	10.32 (7.42) 0.00–21.00

science-based practical activity. First, the class teacher introduced the concepts of sorting and matching with concrete examples. Children were then required to individually categorize living and non-living things and create a poster communicating their findings, which was then discussed with the whole class. These embedded mathematics activities were typically implemented 2-3 times a week, for approximately 20 minutes a session.

Maths apps

Children assigned to the treatment groups used two maths apps, “Maths 3-5” and “Maths 4-6”, developed by *onebillion*, an educational, not-for-profit organization (www.onebillion.org). The maths apps are designed to support the acquisition of basic mathematical skills, including Number, Shape, Space, and Measure. Details about the topics covered in each app and how the app content maps onto the mathematical concepts covered in the school curriculum can be found in Table S2.

Overall the apps are designed to deliver one-to-one, child-centred instruction through interactive picture, audio, and animation formats with clear objectives, instructions, and immediate formative feedback, consistent for all users. Children work through the apps with headphones individually, at their own pace, repeating instructions and activities as often as needed. The topic quizzes, noted above, are designed to assess children’s knowledge of the mathematical concepts. The number of quizzes passed was recorded as an indication of children’s progress throughout the maths apps (Pitchford, Kamchedzera, Hubber, & Chigeda, 2018).

There are 28 topics in total across the two maths apps (see Table S2). The first 22 topics are available in Brazilian Portuguese; the last six topics are available only in English. The difference in the number of available topics in each language could not be controlled, or topics restricted. However, none of the content from the last six topics was included in the primary outcome variable. As the apps are self-paced children could choose the order in which they worked through the app topics, but they were encouraged to complete the app topics systematically as presented in the apps. The first author recorded children’s progress through the apps and access to the different topics. No child in either app group completed the full 22 topics available in both languages over the 10-week intervention period (see Table 2). Two children in treatment group 2 (English maths app) completed one of the six topics that was available only in English.

The first author (native English speaker) with the support of a native Brazilian Portuguese speaker and speaker of English evaluated the adequacy of the translation between the app content in

Table 2: Descriptive data for the EGMA assessment of mathematical ability (max score 54) and number of app topics completed (max 22 topics Brazilian Portuguese; 28 topics English)

<i>Outcome variable</i>	<i>Group 1 (n = 23) Brazilian Portuguese (L1 Maths App)</i>	<i>Group 2 (n = 19) English (L2 Maths App)</i>	<i>Group 3 (n = 19) Control (Standard Practice in English)</i>
<i>Maths score (EGMA)</i>			
Pre-test mean (SD)	15.70 (6.15)	20.53 (7.94)	16.68 (6.90)
Post-test mean (SD)	24.83 (6.35)	30.32 (10.31)	22.53 (7.28)
Gain mean (SD)	9.13 (4.18)	9.79 (5.94)	5.84 (4.13)
Gain range (min-max)	1.00–17.00	0.00–21.00	–1.00–15.00
%Gain	16.9	18.1	10.8
Cohen’s <i>d</i> (95% CI)	1.46 (0.54–2.38)	1.06 (0.10–2.03)	0.83 (–0.11–1.77)
<i>Progress through apps</i>			
App topics completed	14.18 (3.16)	12.33 (2.70)	n/a
Mean (SD)			

Brazilian Portuguese and English. For each of the 22 topics available in Brazilian Portuguese and English, one activity was randomly selected from the seven activities available in each topic in “Maths 3-5” and 6 activities available in each topic in “Maths 4-6.” The Brazilian Portuguese and English transcripts of the first trial for each of the randomly selected activities were compared, such as the one illustrated in Figure S1. All selected app trials were judged to accurately convey the same meaning in Brazilian Portuguese and English, except for one, where the opposite meaning was conveyed. In this trial, the visual information positioned the numbers “10, 20, 30, 40” in ascending order and the English verbal task instructions asked children to continue placing the numbers in the “right order.” However, the Brazilian Portuguese verbal task instructions asked children to continue placing the numbers in “descending order.”

Primary outcome variable: Mathematics attainment

Children’s mathematical knowledge was assessed using the Early Grade Maths Assessment (EGMA); a paper-based, age-appropriate, and internationally developed measure of mathematics attainment (Brombacher, 2010). It assesses number recognition, quantity discrimination, pattern completion, and numerical operations (addition and subtraction). The majority of EGMA items required a non-verbal response from the child (eg, pointing) but for items that required a verbal response (eg, stating a number answer) the child could answer in Brazilian Portuguese or English, as they chose. The maximum score was 54. This assessment was independent from the maths app intervention and the school curriculum so was not biased in favour of the treatment groups or the control group. Content areas of the maths apps and school curriculum that are covered in EGMA are outlined in Table S2. Reliability analysis of the EGMA assessment showed high internal consistency between pre-test and post-test scores, $r = 0.811$, Cronbach’s $\alpha = 0.888$.

Secondary outcome variable: Progression through the maths apps

Children’s progression through the maths apps was captured by the number of topics completed (maximum 22 topics across the two apps for Brazilian Portuguese and 28 topics across the two apps for English). For a child to complete an app topic they had to achieve 100% correct on the topic quiz. Each topic quiz included ten questions designed to assess knowledge of the set of activities and concepts the child had been working on within the topic using new content. Pitchford *et al.* (2018) showed that topics passed within the maths apps correlated significantly with learning gains measured by EGMA, demonstrating that progression through the apps is a valid indicator of mathematics learning. Monitoring data (topics passed) stored within the apps was extracted by the first author and was available for 22 of the 23 children in treatment group 1 (Brazilian Portuguese) and 18 of the 19 children in treatment group 2 (English). Missing data was due to hardware corruption on one device, which was shared across the two treatment groups and so impacted one child from each treatment group.

Language proficiency

Children’s language proficiency in Brazilian Portuguese and English was assessed using a 7-item questionnaire. The child’s Brazilian Portuguese teacher completed the Brazilian Portuguese section of the questionnaire and the same procedure was followed by the child’s English teacher (native Brazilian Portuguese speaker and speaker of English) for the English section of the questionnaire. The questionnaire was designed to measure children’s abilities in speaking (4 items), reading (1 item), writing (1 item) and listening (1 item) in each language. The questionnaire was developed specifically for this study and was based on the Alberta Language and Development Questionnaire, a non-language and non-culturally specific parental questionnaire of children’s language competencies in their first and second language (Paradis, Emmerzael, & Duncan, 2010).

Items were scored between 0 and 3 and the maximum score for each language was 21. The total score for each language was used as an indication of the child's competencies in their L1 and L2. Reliability analysis of the language proficiency questionnaire showed high internal consistency between Brazilian Portuguese (L1) and English (L2) scores, $r = 0.708$, Cronbach's $\alpha = 0.804$.

Procedure

Study invitation and consent procedures

The participating school were invited to take part due to the school's location in Brazil and pre-existing access to the required tablet devices (iPads). The school were given continued access to the maths apps, free of charge, by *onebillion*, meaning that all children, including children in the control group, had access to the maths apps after study completion, at the discretion of school staff. Parents of children were sent information sheets and an information evening was held at the school by senior teaching staff and the first author to fully inform parents about the study.

Teacher training

Prior to study commencement, participating teachers met with the research team to discuss the practicalities of the proposed research design within their daily school time table and tablet device availability. All teaching staff were trained on how to implement the maths app intervention by the first author and a specialist teacher from the UK who has experience of using the maths apps in his class with children aged 4–5 years.

Mathematics and language assessments

Children's mathematics abilities were assessed using EGMA immediately before (pre-test) and after (post-test) the 10-week intervention period. Language assessments were conducted at pre-test only. All assessment instructions were delivered in Brazilian Portuguese using audio recordings playing a standardised script. For items that required a non-verbal response, children responded by pointing. For tasks that required a verbal response, children could respond in Brazilian Portuguese or English, as they chose. Assessments typically lasted approximately 15–20 minutes per child and were administered by the first author on a one-to-one basis in a quiet area, free from distraction, in the child's familiar school environment.

Maths app implementation

Children allocated to receive the maths app intervention in Brazilian Portuguese (L1 treatment group) and English (L2 treatment group) used the maths apps in sessions lasting approximately 20 minutes, four times a week, for 10 consecutive weeks. Sessions were implemented instead of the small group, embedded mathematics activities, used in standard practice.

Children used the same iPad each session, accessing their individual profile within the apps, which saved their progress. To aid classroom organisation, iPads were labelled with the children's name. Prior to the session, the teaching assistant set up the children's profile within the app, ensuring the correct language was selected. All children started using the apps in their allocated language on topic one in "Maths 3-5." Once children had completed all ten topics, they progressed to "Maths 4-6." Teachers kept a register of which children had completed the "Maths 3-5" app to support daily organisation. The maths app intervention was implemented in the children's normal classroom and children were sat in their regular seating plan of 5-7 children per table. Children worked through the mathematics content within their individual profile independently with headphones. Teaching staff provided technical support and behaviour management when needed; all mathematical instruction was delivered through the apps. Implementation of the maths apps by teachers optimised ecological validity.

Time on task

While the maths apps were implemented instead of the embedded mathematics activities typical of standard practice (see Table S1) the amount of time on task was not equivalent across the treatment and control groups. Specifically, standard mathematics practice was typically implemented in the control group 2–3 times a week. This equated to approximately 6–10 hours of mathematics instruction over the 10-week intervention period. In contrast, children in the two treatment groups used the maths apps 4 times a week, equalling approximately 13 hours of mathematics instruction. Therefore, the two treatment groups received greater exposure to mathematics instruction over the 10-week intervention period than the control group (see Table S1) but time on task was equated across the two treatment groups.

Results

Shapiro-Wilk tests identified some measures (ie, L1 proficiency, post-test EGMA, app progress) to deviate significantly from normality ($p < .05$). However, more conservative non-parametric analyses gave the same pattern of results as the parametric analyses reported in the text.

Bilingual status of sample

Paired samples *t*-tests comparing L1 and L2 proficiency across the whole sample and for each group (see Table 1) showed children's L1 proficiency was significantly higher than their L2 proficiency, $t(60) = 5.76, p < .001, d = 0.60$, identifying the sample as unbalanced bilinguals (Baker, 2011). This pattern was consistent across each group, Group 1 $t(18) = 2.28, p = .035, d = 0.50$, Group 2 $t(22) = 3.84, p < .001, d = 0.51$, Group 3 $t(18) = 4.74, p < .001, d = 0.84$.

Primary outcome variable: Mathematics gains

Table 2 shows mean EGMA performance at pre-test and post-test for each group. In accordance with the quasi-experimental research design and differences in pre-test EGMA scores (see Table 2) across the three predetermined groups, a change-score analysis was used (Thomas & Zumbo, 2012; Van Breukelen, 2006). Group mean gain scores (post-test minus pre-test), percentage gains, and within-group effect sizes with 95% confidence intervals (CI) reflecting the progress of each group are also reported in Table 2 along with mean number of topics passed for the two treatment groups.

To examine if the maths app intervention supported the acquisition of basic mathematical skills when delivered in either the child's L1 Brazilian Portuguese (treatment group (1) or the child's L2, English (treatment group (2) compared to standard mathematics practice (control group), the mean gain in EGMA scores (post-test EGMA minus pre-test EGMA; see Table 2) for the three groups were compared using a one-way ANOVA.

Results revealed significant group differences in EGMA gain scores, $F(2, 58) = 3.78, p = .029$. Post hoc independent samples *t*-tests showed that children using the maths apps in Brazilian Portuguese (treatment group 1) made significantly greater gains than the control group, $t(40) = 2.55, p = .015$, between group effect size $d = 0.79$, 95% CI = 0.16–1.42. Similarly, children using the maths apps in English (treatment group 2) made significantly greater gains than the control group, $t(36) = 2.38, p = .023$, between group effect size $d = 0.77$, 95% CI = 0.11–1.43. In contrast, there was no significant difference in gains across the two maths app treatment groups, $t(40) = 0.42, p = .676$, between group effect size $d = 0.13$, 95% CI = -0.48–0.74.

Secondary outcome variable: Progression through the Maths apps

Table 2 reports the mean number of topics completed (number of topic quizzes passed) for children in treatment group 1 (Brazilian Portuguese) and 2 (English). An independent samples

t-test showed children using the maths apps in Brazilian Portuguese (treatment group 1) made more progress throughout the apps compared to children using the apps in English (treatment group 2), which trended towards significance, $t(38) = 1.96$, $p = .057$, between group effect size $d = 0.63$, 95% CI = 0.002–1.25.

Language of instruction proficiency

Pearson correlations showed a small but not statistically significant correlation between proficiency in language of instruction and EGMA gains across the two treatment groups who used the maths apps, $r = 0.25$, $p = .113$. However, a moderate and statistically significant relationship between proficiency in language of instruction and progress through the maths apps measured by topics passed was identified, $r = 0.36$, $p = .022$.

Discussion

This quasi-experimental, pilot study examined the role of proficiency in the language of instruction in the context of app-based mathematics instruction for bilingual children aged 5–6 years old in Brazil. To our knowledge this is the first study to apply bilingual education theory to technology-based mathematical interventions in a controlled, real-world, bilingual immersion setting.

Language of instruction

In this study, the same maths app intervention was implemented within the same Kindergarten classroom environment; only the language of instruction was manipulated. Results showed children aged 5–6 years, who used the maths apps made significant progress in mathematics regardless of whether the apps were delivered in their L1 or L2. However, a larger effect size for gain scores was observed for children who used the apps in Brazilian Portuguese (within-group effect size 1.46, 95% CI = 0.54–2.38) compared to English (within-group effect size 1.06, 95% CI = 0.10–2.03). This corroborates the results found for progression through the apps, as measured by topics passed. Children who received instruction with the apps in Brazilian Portuguese made more progress through the maths apps than children who received instruction in English, a difference that approached significance ($p = .057$), and was characterised by a medium effect size ($d = 0.63$, 95% CI = 0.002–1.25). This indicative finding is consistent with previous research with these maths apps, which showed the level of language processing difficulties of children with special educational needs predicted their learning gains (Pitchford *et al.*, 2018). As the children in this study were identified as unbalanced bilinguals (Baker, 2011), with L1 stronger than L2, these results may suggest the maths apps were implemented most effectively in the child's first language.

This evidence may suggest that children need to have sufficiently developed language proficiency to access curriculum content and to respond to instruction (Cummins, 2008). In the current study, some of the app tasks may have required vocabulary that extended beyond the children's current level of proficiency, specifically their cognitive academic language proficiency (ie, CALP) in the language of instruction. For example, in one quiz question, the on-screen teacher asks the child "how many pencils are there?" from an array of pencils, footballs, cups, bowls, and shoes. If the child cannot identify which items are pencils, their performance may reflect their low CALP rather than their mathematical ability. As progress through the apps was measured by number of topics completed (100% pass rate needed in the topic quiz), these vocabulary restrictions may have hindered children's progress, particularly as the maths apps are self-paced in nature.

Furthermore, the maths apps may not provide sufficiently context embedded communication needed for some learners to adequately access and interact with the mathematical content.

Although the apps are identified to provide relatively high levels of contextual support, through the congruent auditory and visual information, interactive pictures, audios, and animations, combined with an on-screen teacher providing clear task demonstrations, this may be insufficient, depending on the language proficiency of the learner, given these vocabulary restrictions. Context embedded communication typically also includes supportive cues, such as gestures, body language, and intonation, or even concrete cues. These cues allow the learner to draw upon their basic interpersonal communication skills (BICS; Cummins, 2008) and contribute to the understanding of conveyed information (Church, Ayman-Nolley, & Mahootian, 2004; Singer & Goldin-Meadow, 2005). BICS therefore may support understanding when children have not reached a sufficient threshold in their CALP (Cummins, 2008). While some BICS features can be identified within these apps, such as intonation and concrete cues, others including gestures and body language are not.

To explore these two proposed explanations, direct measures of children's L1 and L2 CALP specific vocabulary skills are needed in future research. Without these additional measures these explanations are speculative. Further studies are required to develop research-based design frameworks for app-based mathematics instruction in a bilingual setting, beyond the specific apps evaluated in this study that draw upon the CALP/ BICS distinction (Cummins, 2008) and supports children's developing language proficiencies through supportive scaffolding.

Comparisons to standard mathematical practice

When comparing the maths app intervention to standard practice, results showed children who used the maths apps in either Brazilian Portuguese (treatment group 1) or in English (treatment group 2) made significantly greater mathematical learning gains (16.9% and 18.1% respectively) compared to children who received standard mathematics practice only (control group; 10.8%). However, whilst time on task was equated for the two treatment groups, the control group received less mathematics instruction across the study (see Table S1). Specifically, children in the two treatment groups had approximately 13 hours of exposure to mathematics instruction, compared to approximately 6–10 hours for the children in the control group. Although this procedure was based on practical factors due to the school's capacity and routine, the additional time learning mathematics, may, in part, have accounted for the greater learning gains seen in the two treatment groups compared to control group (Cheung & Slavin, 2013). It is not possible to disentangle the impact of increased time learning maths from the unique impact of app-based mathematical learning in the current study. However, these apps have been shown to significantly raise learning outcomes compared to standard mathematics practice even when time on task was equated across the treatment and control groups in a high-income, Western context (Outhwaite *et al.*, 2018). Future research should build on this current pilot study by ensuring fully time-equivalent treatment and control groups (Holmes & Dowker, 2013). This will help to establish if the current findings are corroborated in a bilingual immersion setting, such as the Brazilian context presented in this study. At the very least, the current results indicate that incorporating additional, structured mathematics instruction into a weekly time table significantly improves early learning outcomes for mathematics, corroborating previous proof of concept studies (Outhwaite *et al.*, 2017).

In addition, as the maths apps were self-paced, it may be possible that children in the intervention groups accessed additional and potentially more advanced mathematical content (see Table S2) than children in the control group who received standard mathematical practice at the pace established by the classroom teacher. This makes direct comparisons between the intervention and control groups challenging. However, the standardised outcome measure (EGMA)

was independent of both the intervention and control learning content, thus limiting potential threats to internal validity.

Finally, it is important to acknowledge the error observed in the analysis of translation adequacy, which was based on a random sample of app topics (see above). While the single error observed in this sample is unlikely to significantly impact the results, it is imperative that future research achieves and maintains linguistic accuracy in the presentation of the same mathematical content in different languages.

Implications & conclusion

This study provides initial proof of concept for the use of app-based mathematics instruction for children aged 5–6 years in a dual immersion bilingual setting. Current findings indicatively suggest that to maximise interactions and learning with the maths apps, children need to have a sufficient level of language proficiency in the language of instruction. As children in the current study were shown to have a stronger L1 proficiency, it suggests that for these children the apps are most effectively implemented in Brazilian Portuguese. However, for similarly high-quality educational apps that are not available in multiple languages, app-based learning in L2 can be effective. Further research is needed to explore the influence of specific language proficiencies, particularly CALP and BICS, as well as children's language responses during implementation.

Critically, this study provides initial evidence towards advancing our understanding of bilingual learners' experiences of app-based learning in a controlled, real-world, bilingual setting. For teaching professionals, this study demonstrates this particular maths app intervention, implemented through the reported arrangements, can positively benefit children's acquisition of basic mathematical skills and that implementation considerations should be made that account for individual children's proficiencies in the language of instruction. Now that proof of concept has been established in this context, further research is needed to support more detailed recommendations to teaching professionals and parents for using educational maths apps with young bilingual children.

Acknowledgements

This work was supported by the Overseas Fieldwork Fund from the Economic and Social Research Council [Grant number ES/J500100/1]. We would like to thank the staff, children and parents at ABA Global School, Recife for participating in this research. We would also like to thank *onebillion* for supporting this research through in-kind donations of the app software. Lastly, we would like to thank Marie Neves for their support in translation and recruiting the participating school and Marc Faulder for their support in training the participating teachers.

Statements on open data, ethics and conflict of interest

The data presented in this study is available upon reasonable request to the first author.

Ethical approval for this study was granted by the Federal University of Pernambuco, the regional ethics committee, Comitês de Ética em Pesquisa (CEP) and the national ethics committee, Comissão Nacional de Ética em Pesquisa (CONEP).

All authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Notes

¹This categorization of *first* and *second* language is used to support this initial investigation. It is not supposed that languages are acquired fully sequentially or that children do not also use other languages.

²The typical language of mathematics instruction in this study was English, which for all participating children was their second language (L2).

References

- Baker, C. (2011). *Foundations of bilingual education and bilingualism*. Bristol, UK: Multilingual Matters.
- Berkowitz, T., Schaeffer, M. W., Maloney, E. A., Peterson, L., Gregor, C., Levine, C. S., & Beilock, S. L. (2015). Maths at home adds up to achievement in schools. *Science*, 350(6257), 196–198.
- Brombacher, A. (2010). *Malawi early grade mathematics assessment (EGMA): National baseline report*. Malawi: USAID/Malawi and the Ministry of Education, Science and Technology.
- Carr, J. (2012). Does Maths achievement h'APP'en when iPads and game-based learning are incorporated into fifth-grade Mathematics instruction? *Journal of Information Technology Education*, 11, 269–286.
- Cheung, A. C. K., & Slavin, R. E. (2013). The effectiveness of educational technology applications for enhancing Maths achievement in K-12 classrooms: A meta-analysis. *Educational Research Review*, 9, 88–113.
- Church, R. B., Ayman-Nolley, S., & Mahootian, S. (2004). The role of gesture in bilingual education: Does gesture enhance learning? *International Journal of Bilingual Education and Bilingualism*, 7(4), 303–319.
- Connolly, P., Keenan, C., & Urbanska, K. (2018). The trials of evidence-based practice in education: A systematic review of randomised controlled trials in education research 1980–2016. *Educational Research*, 60(3), 276–291.
- Couse, L. J., & Chen, D. W. (2010). A tablet computer for young children? Exploring its viability for early childhood education. *Journal of Research on Technology in Education*, 43(1), 75–96.
- Cummins, J. (1981). The role of primary language development in promoting educational success for language minority students. In California State Department of Education (Eds.), *Schooling and language minority students. A theoretical framework* (pp. 3–51). California, USA: California State Department of Education.
- Cummins, J. (2008). BICS and CALP: Empirical and theoretical status of the distinction. In B. Street & N. H. Hornberger (Eds.), *Encyclopedia of language and education* (pp. 71–83). New York, USA: Springer.
- Department for Education. (2019). *Realising the potential of technology in education*. Retrieved from <https://www.gov.uk/government/publications/realising-the-potential-of-technology-in-education>
- Drinkwater, D. (2013). *Brazilian government spends big on nearly 500,000 tablets for local teachers*. Retrieved from <http://tabtimes.com/news/education/2013/11/27/brazilian-government-spends-big-nearly-500000-tabletslocal-teachers>
- Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). Improving students' learning with effective learning techniques: Promising directions from cognitive and educational psychology. *Psychological Science in the Public Interest*, 14(1), 4–58.
- Gomez, L., Freeman, D., & Freeman, Y. (2005). Dual language education: A promising 50–50 model. *Bilingual Research Journal*, 29(1), 145–164.
- Grimaldi, J. D., & Karpicke, P. J. (2012). Retrieval-based learning: A perspective for enhancing meaningful learning. *Educational Psychology Review*, 24(3), 401–418.
- Hirsh-Pasek, K., Zosh, J. M., Michnick-Golinkoff, R., Gray, J. H., Robb, M. B., & Kaufman, J. (2015). Putting education in “Educational” apps: Lessons from the science of learning. *Psychology Science*, 16(1), 3–34.
- Holloway, D., Green, L., & Livingstone, S. (2013). *Zero to eight: Young children and their internet use*. London, UK: LSE.
- Holmes, W., & Dowker, A. (2013). Catch up numeracy: A targeted intervention for children who are low-attaining in mathematics. *Research in Mathematics Education*, 15(3), 249–265.
- Hubber, P. J., Outhwaite, L. A., Chigeda, A., McGrath, S., Hodgen, J., & Pitchford, N. J. (2016). Should touch screen tablets be used to improve educational outcomes in primary school children in developing countries? *Frontiers in Psychology*, 7, 839.

- Inal, Y., & Cagiltay, K. (2007). Flow experiences of children in an interactive social game environment. *British Journal of Educational Technology*, 38(3), 455–464.
- Lindah, M. G., & Folkesson, A. M. (2012). ICT in preschool: Friend or foe? The significance of norms in a changing practice. *International Journal of Early Years Education*, 20(4), 422–436.
- Magliaro, S. G., Lockee, B. B., & Burton, J. K. (2005). Direct instruction revisited: A key model for instructional technology. *Educational Technology Research and Development*, 53(4), 41–55.
- Outhwaite, L. A., Faulder, M., Gulliford, A., & Pitchford, N. J. (2018). Raising early achievement in math with interactive apps: A randomized control trial. *Journal of Educational Psychology*, 111(2), 284–298.
- Outhwaite, L. A., Gulliford, A., & Pitchford, N. J. (2017). Closing the gap: efficacy of a tablet intervention to support the development of early mathematical skills in UK primary school children. *Computers & Education*, 108, 43–58.
- Outhwaite, L. A., Gulliford, A., & Pitchford, N. J. (2019). A new methodological approach for evaluating the impact of educational intervention implementation on learning outcomes. *International Journal of Research & Method in Education*. <https://doi.org/10.1080/1743727X.2019.1657081>.
- Paradis, J., Emmerzael, K., & Duncan, T. S. (2010). Assessment of English language learners: Using parent report on first language development. *Journal of Communication Disorders*, 43(6), 474–497.
- Pitchford, N. J. (2015). Development of early mathematical skills with a tablet intervention: A randomized control trial in Malawi. *Frontiers in Psychology*, 6, 485.
- Pitchford, N. J., Chigeda, A., & Hubber, P. J. (2019). Interactive apps prevent gender discrepancies in early grade mathematics in a low-income country in Sub-Saharan Africa. *Developmental Science*, 22(5), e12864.
- Pitchford, N. J., Kamchedzera, E., Hubber, P. J., & Chigeda, A. L. (2018). Interactive apps promote learning of basic mathematics in children with special educational needs and disabilities. *Frontiers in Psychology*, 9, 262.
- Schacter, J., & Jo, B. (2016). Improving low-income pre-schoolers Maths achievement with Maths Shelf, a preschool tablet computer curriculum. *Computers in Human Behavior*, 55, 223–229.
- Schacter, J., & Jo, B. (2017). Improving preschoolers' mathematics achievement with tablets: a randomized controlled trial. *Mathematics Education Research Journal*, 29(3). <https://doi.org/10.1007/s13394-017-0203-9>.
- Singer, M. A., & Goldin-Meadow, S. (2005). Children learn when their teacher's gestures and speech differ. *Psychological Science*, 16(2), 85–89.
- Slavin, R. E., & Lake, C. (2008). Effective programs in elementary mathematics: A best-evidence synthesis. *Review of Educational Research*, 78(3), 427–515.
- Thomas, D. R., & Zumbo, B. D. (2012). Difference scores from the point of view of reliability and repeated-measures ANOVA: In defense of difference scores for data analysis. *Educational and Psychological Measurement*, 72(1), 37–43.
- Van Breukelen, G. J. (2006). ANCOVA versus change from baseline had more power in randomized studies and more bias in nonrandomized studies. *Journal of Clinical Epidemiology*, 59(9), 920–925.
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, USA: Harvard University Press.

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.